Surveys in Geophysics
https://doi.org/10.1007/s10712-026-09952-9

SeisDPS-3D: A Diffusion Model-Based Approach for 3D Seismic Data Deconvolution

Dawei Liu[1,2] · **Hongzhi Yu**[1] · **Mauricio D. Sacchi**[2]

Received: 6 May 2025 / Accepted: 25 April 2026

Abstract
Deconvolution is a critical technique for improving seismic data resolution. It eliminates the filtering effects of the seismic wavelet to obtain reflection coefficients. However, deconvolution is a highly ill-posed problem, making accurate recovery of reflection coefficients challenging. Traditional methods obtain a stable solution to ill-posed problems by incorporating explicit model-based priors based on simplified assumptions. However, they often struggle to produce satisfactory results in complex geological environments or with noisy data. Although deep learning methods can effectively enhance deconvolution results by learning implicit priors from training data, they also suffer from certain limitations, such as generalization problems. Diffusion models offer a novel approach to acquiring prior information, known as generative priors. They directly model the approximate probability distribution of reflection coefficients by training a denoiser and learning to generate reflection coefficients from random noise. This denoiser provides the diffusion model with inherent robustness to random noise. We propose a 2D seismic deconvolution diffusion model (SeisDPS) that combines generative priors with model-based priors to achieve more accurate and high-fidelity reflection coefficients. However, existing diffusion-based deconvolution methods typically focus on processing 2D data, and even a few recent 3D approaches are computationally expensive. Then, building upon SeisDPS, we propose an algorithm for 3D seismic deconvolution (SeisDPS-3D). This 3D algorithm applies SeisDPS to one direction (e.g., inline profiles using SeisDPS) while applying unconditional sampling to the other (e.g., crossline profiles for continuous sampling correction) and vice versa. This unconditional sampling uses the prior information learned by the 2D diffusion model as 3D lateral constraints in orthogonal directions, thereby effectively improving the spatial continuity of the 3D deconvolution results. Under the same conditions, compared to the 2D diffusion model using slice processing, our method only increases the memory consumption by 5%. Our method is trained solely on 2D synthetic reflection coefficients. It demonstrates superior results to the sparse-spike inversion method on 3D synthetic and field datasets. Experimental results show that our method exhibits greater robustness to noise and improved spatial continuity.

Keywords Diffusion model · 3D seismic deconvolution · High-resolution seismic data · 3D algorithm

Extended author information available on the last page of the article

1 Introduction

Convolution models represent seismic data as the convolution of seismic wavelets and reflection coefficients, effectively providing a fundamental framework for representing seismic traces (Robinson 1957). Deconvolution aims to eliminate the filtering effects of seismic wavelets to obtain reflection coefficients, thereby enhancing seismic data resolution. Moreover, deconvolution is also often called equalization as it guarantees that all seismograms represent responses with equal residual source functions. However, deconvolution is often a highly ill-posed problem. First, due to the filtering effect, seismic data inherently possesses a limited bandwidth (Ricker 1953). The loss of high-frequency information makes it challenging to uniquely recover reflection coefficients (Backus and Gilbert 1968; Claerbout 1971; Claerbout and Muir 1973; Claerbout 1976). Second, seismic data are invariably contaminated by noise. Noise introduces uncertainty and can be amplified during deconvolution, making the deconvolution solution unstable. Third, approximate forward operators simplify the complex physics of seismic wave propagation to improve computational efficiency. However, this introduces modeling errors, which further exacerbate the ill-posedness. Introducing accurate prior information provides an effective way to mitigate ill-posedness.

Traditional methods rely on explicit, simplified mathematical models, introducing explicit prior information to mitigate the ill-posed nature of deconvolution. Early deconvolution methods rely on assumptions like reflection coefficients following a white noise distribution and the seismic wavelet being minimum-phase. Robinson (1957) first proposes deconvolution methods; building upon Robinson's work, several other deconvolution methods were proposed, including predictive deconvolution (Peacock and Treitel 1969), least-squares deconvolution (Berkhout 1977), and minimum/maximum entropy deconvolution (Burg 1972; Wiggins 1978; Sacchi et al. 1994), homomorphic deconvolution (Ulrych 1971), and l_p $(0 \leq p \geq 1)$ sparsity constraints (Debeye and van Riel 1990; Taylor et al. 1979; Sacchi 1997; Chopra et al. 2006; Zhang and Castagna 2011; Gholami and Sacchi 2012). However, the simplified assumptions underlying these traditional methods often lead to incomplete or inaccurate prior information, consequently limiting their performance. Therefore, acquiring more accurate prior information is crucial for better deconvolution results. Furthermore, traditional methods typically operate trace-by-trace, neglecting the inherent spatial continuity of seismic data and resulting in poor lateral continuity in deconvolution results. Although subsequent studies introduce multi-trace deconvolution methods (Zhang et al. 2013; Kazemi and Sacchi 2014; Hamid and Pidlisecky 2015; Gholami 2016), these approaches typically rely on adjacent or a limited number of seismic traces, such as three or five traces (Porsani and Ursin 2007; Li et al. 2016). While this strategy improves lateral continuity to some extent, it remains insufficient for capturing long-range correlations between traces, especially those separated by significant distances or involving complex relationships across multiple traces. Additionally, traditional deconvolution frameworks are primarily designed for 2D seismic data, and algorithms capable of direct application to high-dimensional data remain scarce.

In recent years, deep learning techniques have driven the development of numerous deep learning-based techniques for seismic data deconvolution (Araya-Polo et al. 2018; Chai et al. 2021). Unlike traditional methods relying on explicit mathematical formulations, deep learning-based deconvolution methods acquire prior knowledge directly from data. These learning-based approaches can be categorized based on their learning paradigms as supervised, unsupervised, and semi-supervised, as summarized in Table 1.

Table 1 Comparison of different deep learning methods for seismic deconvolution

Method family	Data consistency	Handcrafted prior	Generative prior	3D consistency	Inference time	Memory usage
2D supervised	Δ	✗	✗	NA	Fast	Low
3D supervised	Δ	✗	✗	Δ	Fast	Medium
Semi-supervised	✓	✓	✗	NA	Fast	Low
Unsupervised:						
CNN-based	✓	✓	✗	NA	Moderate	Medium
3D Diffusion	✗	✗	3D Prior	Native 3D	Slow	High
SeisDPS*	✓	✓	2D Prior	NA	Moderate	Medium
SeisDPS-3D*	✓	✓	2D Prior	Strategic 3D	Moderate	Medium

△ Denotes optional usage.

*Denotes the method proposed in this study

Supervised learning learns an implicit prior from paired training samples and can learn complex non-linear mappings, achieving satisfactory performance when trained on large and representative datasets (Zhang et al. 2021; Wu et al. 2021). However, the scarcity of high-quality labels in real field seismic datasets often leads to significant generalization issues. To mitigate the reliance on labeled data, unsupervised learning methods aim to implicitly learn prior knowledge from the inherent statistical properties and structural patterns present within the unlabeled training seismic data itself (Wang et al. 2022). However, the absence of labels leads to a lack of clear evaluation criteria, which can cause the training process to converge to non-global optima, potentially resulting in suboptimal performance. Consequently, the performance of unsupervised learning may be lower than that of supervised methods. Self-supervised learning is a specialized form of unsupervised learning. It generates pseudo-labeled data from large amounts of unlabeled data through carefully designed pretext tasks and subsequently trains the model in a manner analogous to supervised learning (Wu et al. 2025). The quality of the automatically generated pseudo-labeled data plays a crucial role in determining the final performance of the model. Semi-supervised learning offers a balanced approach by utilizing a small amount of labeled data with a larger quantity of unlabeled data. They often incorporate domain-specific prior knowledge, such as information derived from well-logging data, to constrain the deconvolution process (Alfarraj and AlRegib 2019; Chen et al. 2021; Liu et al. 2025). The labeled data trains an initial model, while abundant unlabeled data allows the model to capture the underlying data distribution and characteristics better, thus enhancing the robustness of the deconvolution model. Beyond leveraging real field seismic and well-logging data, synthetic data is also explored as a valuable source for acquiring prior knowledge (Torres and Sacchi 2023). Although leveraging synthetic data effectively enhances generalization, the model's distribution must be re-generalized to adapt to different real-world datasets. Despite the diverse strategies employed by these learning-based methods to acquire prior information, they consistently demonstrate superior accuracy and robustness compared to traditional approaches (Chai et al. 2021; Gao et al. 2021; Chen et al. 2023; Rahman et al. 2024). Nevertheless, most existing deep learning approaches are also applied to 2D seismic data.

Diffusion models implicitly learn generative priors by directly modeling the approximate probability distribution of reflection coefficients through training a denoiser (Sohl-Dickstein et al. 2015; Song and Ermon 2019; Ho et al. 2020). This paradigm shift in acquiring prior information offers significant advantages, notably the ability to train the diffusion model using solely clean reflection coefficient data, thereby circumventing the requirement for paired data-reflectivity training samples often needed in supervised learning. Furthermore, this denoiser guides random noise during the generation process towards regions more consistent with the prior distribution, endowing the diffusion model with inherent robustness to random noise. This generative process allows for the introduction of additional prior information through the use of specific sampling algorithms (Daras et al. 2024), such as ILVR (Choi et al. 2021) and diffusion posterior sampling (DPS) (Chung et al. 2022), effectively alleviating the result non-uniqueness in unconditional sampling approaches such as denoising diffusion probabilistic models (DDPM) (Ho et al. 2020). For example, Yu et al. (2025b) successfully applies diffusion models to 2D seismic data deconvolution by employing DPS. This model effectively integrates observed data into the sampling process, providing guidance and constraints on the deconvolution results and demonstrating enhanced generalization capabilities with satisfactory outcomes. Although DPS introduces additional prior information into the sampling process of diffusion models, diffusion models excel at learning complex nonlinear features and therefore tend to fit such complexities when constructing approximate prior distributions. This tendency

can consequently lead to potentially generating artifacts in regions characterized by simpler, more regular structures, such as Fig. 5a. Therefore, building upon Yu et al. (2025b), we propose a novel 2D seismic deconvolution diffusion posterior sampling model (SeisDPS). This method introduces a generative prior learned by the diffusion model, which can achieve more accurate and high-fidelity reflection coefficients when combined with an explicit prior. However, even with these improvements, SeisDPS remains fundamentally a 2D method that can only handle 2D seismic data.

Although diffusion models have shown potential in deconvolution, the effective application remains limited to 2D data. The application of 3D seismic data remains a largely unexplored area. Three-dimensional data provides detailed mapping and characterization of intricate geological features from various perspectives, which is essential for accurately assessing the size, shape, orientation, and internal heterogeneity of reservoirs, faults, and other structures of interest. In contrast, 2D data only provides a limited 2D perspective, and using simple 2D processing may lead to a loss of crucial information about the true 3D geometry and spatial relationships of subsurface features. Therefore, given the inherently 3D nature of subsurface geological structures, a method capable of directly handling 3D data is essential.

There are two main approaches for applying diffusion models to 3D data: 2D diffusion applied to 3D data and 3D diffusion that directly learn the prior distribution of the 3D space. Although current technologies have made 3D data modeling easier, they require significant time and computational costs. For example, Yu et al. (2025a) employed diffusion models for 3D data modeling, and even with 8 NVIDIA A100 GPUs for training, it still took 21 days. Furthermore, testing with 6 GPUs also required 30 min, highlighting the significant challenges in terms of computational resources and time. In contrast to the direct 3D approach, the second strategy involves initially slicing the 3D data into a series of 2D slices, processing each slice independently using a 2D diffusion model, and subsequently reassembling the processed slices in their original order. Although this method offers a pathway to address 3D problems with lower computational demands, it inherently neglects the spatial continuity of 3D data, often resulting in artifacts and diminished lateral coherence in the results. To address the discontinuity issue between processed slices, Chung et al. (2023) used total variation to enhance the continuity of adjacent slices in the orthogonal dimension. While this strategy can effectively smooth the processed 3D data and improve local consistency, it primarily considers dependencies between neighboring slices. It does not fully capture the intrinsic spatial continuity of 3D data. Therefore, the development of an effective new solution is essential to overcome these limitations and enable the efficient processing of 3D data in the field of seismic deconvolution.

Therefore, based on SeisDPS, we propose an algorithm for 3D deconvolution (SeisDPS-3D). This method performs SeisDPS on the inline profile and uses unconditional sampling on the crossline profiles probabilistically or with fixed frequency to correct the sampling process continuously and vice versa. The inline profile is typically chosen for SeisDPS because the crossline spacing is often smaller than the inline spacing in real-world data, allowing for finer spatial information to enhance the overall 3D deconvolution. Compared to the simple total variation constraint (Chung et al. 2023), the prior information-based constraints help the model produce more geologically realistic solutions. By adding minimal additional memory to the 2D slice processing, we achieved 3D data deconvolution, effectively addressing the memory consumption challenge of processing 3D problems and significantly improving the spatial continuity of the deconvolution results. Moreover, we train only on 2D synthetic reflection coefficient data and can achieve robust performance across multiple datasets by fine-tuning the model's hyperparameters. In addition, multiple

experimental results further demonstrate the proposed method's robustness to noise, superior generalization capability, and satisfactory spatial continuity of the inversion results.

2 Method

2.1 Score-Based Diffusion Model

Diffusion models, as one of the most advanced generative models currently, learn to model the probability distribution of training data via the forward process. Following training, they leverage the reverse process to generate data resembling the training data distribution.

The currently popular diffusion model was first proposed by Ho et al. (2020), and later, Song et al. (2020) extended this model by defining it through stochastic differential equations (SDE). In the forward process, given a target data $\boldsymbol{x}_0 \sim p_0$, noise is progressively added over T time steps, gradually transforming the data into a simple standard Gaussian distribution $\boldsymbol{x}_T \sim \mathcal{N}(0, \mathbf{I})$. The forward process is defined as follows:

$$\mathrm{d}\boldsymbol{x} = -\frac{\beta_t}{2}\boldsymbol{x}dt + \sqrt{\beta_t}\mathrm{d}\boldsymbol{w} \tag{1}$$

where β_t is the noise schedule of the process. It is typically defined as a linear function that increases monotonically with respect to time t. Generally, β_t is set to rise linearly from 10^{-4} to 0.02 over time t (Ho et al. 2020). Additionally, $\boldsymbol{w}$ denotes the standard Wiener process (a.k.a., Brownian motion).

We aim to reconstruct the approximate distribution of the original data from Gaussian noise. This process is represented by the corresponding inverse SDE as follows (Anderson 1982):

$$\mathrm{d}\boldsymbol{x} = \left[-\frac{\beta_t}{2}\boldsymbol{x} - \beta_t \nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)\right]\mathrm{d}t + \sqrt{\beta_t}\mathrm{d}\overline{\boldsymbol{w}}, \tag{2}$$

where $\mathrm{d}t$ corresponds to time running backward and $\mathrm{d}\overline{\boldsymbol{w}}$ to the standard Wiener process running backward.

To solve the inverse time stochastic differential equation similar to Eq. 2, it is important to determine the time-dependent score function $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)$. This term can be approximated using the neural network $\boldsymbol{s}_\theta$, which is trained using denoising score matching (DSM) (Vincent 2011):

$$\theta^* = \arg\min_{\boldsymbol{\theta}} \mathbb{E}_{\boldsymbol{x}_t|\boldsymbol{x}_0,\boldsymbol{x}_0}[\|\boldsymbol{s}_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t) - \nabla_{\boldsymbol{x}_t} \log p(\boldsymbol{x}_t|\boldsymbol{x}_0)\|_2^2]. \tag{3}$$

By substituting $\boldsymbol{s}_{\theta^*}(\boldsymbol{x}_t, t)$ for $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)$ in Eq. 2, the sampling results can be obtained recursively.

2.2 Diffusion Posterior Sampling

Diffusion posterior sampling (DPS), proposed by Chung et al. (2022), effectively introduces additional prior information by incorporating observed data into the sampling process, thus

further constraining the solution space. The general forward model of an inverse problem is defined as follows:

$$\boldsymbol{y} = \mathcal{A}(\boldsymbol{x}) + \boldsymbol{\epsilon}, \tag{4}$$

where $\mathcal{A}(\cdot)$ represents the forward measurement operator and ϵ denotes the measurement noise, which is assumed to follow a Gaussian distribution. According to Bayes'theorem, the term $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)$ in Eq. 2 can be reorganized and rewritten as:

$$\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t|\boldsymbol{y}) = \nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t) + \nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{y}|\boldsymbol{x}_t). \tag{5}$$

Since $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)$ is determined solely by the distribution of the training data, the pre-trained $\boldsymbol{s}_{\theta^*}(\boldsymbol{x}_t, t)$ is used to substitute the unconditional score term $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)$. Thus, Eq. 5 can be rewritten as:

$$\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t|\boldsymbol{y}) \simeq \boldsymbol{s}_{\theta^*}(\boldsymbol{x}_t, t) + \nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{y}|\boldsymbol{x}_t). \tag{6}$$

Since the term $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{y}|\boldsymbol{x}_t)$ depends on time t and there is no clear direct relationship between $\boldsymbol{y}$ and $\boldsymbol{x}_t$, the primary challenge is to approximate the term $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{y}|\boldsymbol{x}_t)$. To address this issue, Chung et al. (2022) proposed an approximation method:

$$\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{y}|\boldsymbol{x}_t) \simeq \nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{y}|\hat{\boldsymbol{x}}_0(\boldsymbol{x}_t)), \tag{7}$$

where $\hat{\boldsymbol{x}}_0(\boldsymbol{x}_t)$ emphasizes that $\hat{\boldsymbol{x}}_0$ is a function of $\boldsymbol{x}_t$. Under Gaussian noise, based on Tweedie's denoising estimate (Efron 2011), equations 6 and 7 are integrated as follows:

$$\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t|\boldsymbol{y}) \simeq \boldsymbol{s}_{\theta^*}(\boldsymbol{x}_t, t) - \zeta_t \nabla_{\boldsymbol{x}_t} \left\| \boldsymbol{y} - \mathcal{A}(\hat{\boldsymbol{x}}_0) \right\|_2^2, \tag{8}$$

where $\zeta_t = \eta / \left\| \boldsymbol{y} - \mathcal{A}(\hat{\boldsymbol{x}}_0) \right\|_2$ is the step size, and η is a hyperparameter that controls the step size, allowing for adjustments in the iterative process. The algorithm samples from the posterior distribution, with the algorithmic flow as shown in Algorithm 1. To simplify the calculation, define $\alpha_t = 1 - \beta_t$ and $\bar{\alpha}_t = \prod_{s=0}^{t} \alpha_s$ (Ho et al. 2020).

Algorithm 1 DPS

Require: $T, \boldsymbol{y}, \{\zeta_t\}_{t=1}^N, \{\sigma_t\}_{t=1}^N, s_\theta$
1: $\boldsymbol{x}_T \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$
2: **for** $t = T, \ldots, 1$ **do**
3: $\quad \boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$ **if** $t > 1$, **else** $\boldsymbol{z} = 0$
4: $\quad \hat{\boldsymbol{x}}_0 \leftarrow \frac{1}{\sqrt{\bar{\alpha}_t}} \left(\boldsymbol{x}_t + (1 - \bar{\alpha}_t) \boldsymbol{s}_\theta (\boldsymbol{x}_t, t) \right)$
5: $\quad \boldsymbol{x}'_{t-1} \leftarrow \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t} \boldsymbol{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t} \hat{\boldsymbol{x}}_0 + \sigma_t \boldsymbol{z}$
6: $\quad \boldsymbol{x}_{t-1} \leftarrow \boldsymbol{x}'_{t-1} - \zeta_t \nabla_{\boldsymbol{x}_t} \| \boldsymbol{y} - \mathcal{A}(\hat{\boldsymbol{x}}_0) \|_2^2$
7: **end for**
8: **return** $\hat{\boldsymbol{x}}_0$

2.3 Improved DPS for Seismic Deconvolution

Deconvolution is based on the assumption of the convolution model (Robinson 1957), where the seismic trace can be represented as the convolution of the reflection coefficients and a seismic wavelet. The forward model is generally expressed as follows:

$$\mathbf{D} = \mathcal{L}(\mathbf{M}) + \boldsymbol{\epsilon}. \tag{9}$$

where $\mathbf{M}$ represents the reflection coefficients, $\mathbf{D}$ is the observed seismic data, ϵ denotes an unknown data error represented as additive noise, and $\mathcal{L}(\cdot)$ is a forward operator.

In the convolution model, no specific assumption is made about the seismic wavelet, which may either be time-invariant or time-varying. In this paper, the seismic wavelet is assumed to be time-invariant and known or at least well approximated as a preliminary step (Ulrych et al. 1995). While a more realistic model entails an unknown and nonstationary propagating wavelet, this simplification serves as a good approximation. As a specific realization of the general forward operator $\mathcal{A}(\cdot)$ in Eq. 4, the forward operator $\mathcal{L}(\cdot)$ is represented as a Toeplitz matrix formed by the seismic wavelet, where L_{ij} is given as follows:

$$L_{ij} = \begin{cases} r_{i-j+1} & i \geq j \\ 0 & i < j \end{cases}, \tag{10}$$

where r_{i-j+1} refers to the $(i-j+1)$th sample point of the seismic wavelet. This means that we need to estimate the seismic wavelet in advance.

Although DPS provides more prior information to the diffusion model's sampling process, its application to certain datasets may inadvertently introduce artifacts, such as Fig. 5a. A plausible explanation is that diffusion models, owing to their proficiency in learning intricate nonlinear features, tend to overfit these complexities when constructing an approximate prior distribution, potentially generating artifacts in regions characterized by simpler, more regular structures.

To address these issues, we introduce a sparse explicit prior based on the assumption that subsurface interfaces are typically sparsely distributed. Combining this sparsity-explicit prior with the complex implicit prior learned by the diffusion model ensures that the generated reflection coefficients are more realistic. This approach enables high-resolution deconvolution in both complex and simple subsurface regions. Specifically, we introduce ℓ_1 regularization to implement the sparse constraint. The ℓ_1 regularization facilitates the reconstruction of sparse subsurface structures, such as subsurface interfaces or stratified features. Furthermore, introducing this additional constraint helps alleviate the ill-posedness inherent in deconvolution and reduces the risk of overfitting. The ℓ_1 regularization introduces a constraint on the reflection coefficients by incorporating their ℓ_1-norm into the objective function, as defined below:

$$\min_{M} \|\mathbf{D} - \mathcal{L}(\mathbf{M})\|_2^2 + \xi \|\mathbf{M}\|_1, \tag{11}$$

where ξ denotes the regularization parameter.

To enhance the sparsification process, a soft threshold operation is introduced. Soft thresholding is a nonlinear method for promoting sparsity. Moreover, this operation can

also remove small-amplitude random noise. For a signal $\mathbf{M}$, the soft thresholding operation is defined as follows

$$\mathcal{S}(\mathbf{M}, \lambda) = \text{sign}(\mathbf{M}) \cdot \max(|\mathbf{M}| - \lambda, 0), \tag{12}$$

where λ represents the threshold, sign $(\cdot)$ denotes the sign function. Finally, we redefine the sampling process at each step based on DPS, starting with soft thresholding:

$$\hat{\mathbf{M}}_0' \leftarrow \mathcal{S}(\hat{\mathbf{M}}_0, \lambda). \tag{13}$$

Then, the processed $\hat{\mathbf{M}}_0'$ is used as the new $\mathbf{M}_0$ for the next step, resulting in $\mathbf{M}_{t-1}$:

$$\mathbf{M}_{t-1} \leftarrow \mathbf{M}_{t-1}' - \zeta_t \nabla_{\mathbf{M}_t} \|\mathbf{D} - \mathcal{L}(\hat{\mathbf{M}}_0')\|_2^2 - \xi \nabla_{\mathbf{M}_t} \|\hat{\mathbf{M}}_0'\|_1, \tag{14}$$

The improved algorithm is called SeisDPS, and the algorithmic flow is illustrated in Algorithm 2.

Algorithm 2 SeisDPS

Require: $T, \mathbf{D}, \mathcal{L} \in \mathbb{R}^{x \times t}, \{\zeta_t\}_{t=1}^N, \{\boldsymbol{\sigma_t}\}_{t=1}^N, \lambda, \xi, s_\theta$
$\mathbf{M}_T \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}) \in \mathbb{R}^{x \times t}$
for $t = T, ..., 1$ **do**
  $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$ **if** $t > 1$, **else** $\mathbf{Z} = 0$
  $\hat{\mathbf{M}}_0 \leftarrow \frac{1}{\sqrt{\bar{\alpha}_t}} (\mathbf{M}_t + (1 - \bar{\alpha}_t) \boldsymbol{s}_\theta (\mathbf{M}_t, t))$
  $\mathbf{M}_{t-1}' \leftarrow \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t} \mathbf{M}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t} \hat{\mathbf{M}}_0 + \boldsymbol{\sigma_t} \mathbf{Z}$
  $\hat{\mathbf{M}}_0' \leftarrow \mathcal{S}(\hat{\mathbf{M}}_0, \lambda)$
  $\mathbf{M}_{t-1} \leftarrow \mathbf{M}_{t-1}' - \zeta_t \nabla_{\mathbf{M}_t} \|\mathbf{D} - \mathcal{L}(\hat{\mathbf{M}}_0')\|_2^2 - \xi \nabla_{\mathbf{M}_t} \|\hat{\mathbf{M}}_0'\|_1$
end for
return $\mathbf{M}_0'$

2.4 Solving 3D Deconvolution with 2D Diffusion Models

For a 3D seismic data $\mathcal{D} \in \mathbb{R}^{x \times y \times t}$, the considerable computational demand of direct 3D deconvolution, particularly in terms of GPU memory, presents a significant challenge. A widely adopted strategy involves slicing the data and independently applying 2D deconvolution along the inline or crossline profiles, followed by sequential reconstruction of the data. However, this approach neglects the intrinsic spatial continuity of the 3D data, leading to suboptimal results with degraded continuity in the other orthogonal spatial dimension. Chung et al. (2023) uses total variation (TV) to smooth adjacent slices in the orthogonal axis. Although this strategy can effectively smooth the processed 3D data and improve local consistency, it primarily considers dependencies between neighboring slices. It does not fully capture the intrinsic spatial continuity of 3D data.

Inspired by the work of Lee et al. (2023), we adopt an alternating sampling approach to process 3D data using a 2D diffusion model. Distinct from the approach of Lee et al. (2023), which requires two separate models, our framework leverages a single unified model for sampling. This is attributed to the inherent characteristics of seismic data, where the inline and crossline profiles share consistent distributions, enabling the

employment of a single underlying model to unify them. Moreover, a and b induce weighting between the two distributions according to their importance. In the case of unconditional sampling, the process can be formally expressed as follows:

$$\begin{aligned}\nabla_{\mathcal{X}_t}\log p(\mathcal{X}_t) &= a\nabla_{\mathcal{X}_t}\log q^{(I)}(\mathcal{X}_t) + b\nabla_{\mathcal{X}_t}\log q^{(C)}(\mathcal{X}_t)\\ &= a\sum_{j=1}^{y}\nabla_{\mathcal{X}_t}\log q^{(I)}(\mathcal{X}_{t,[:,j,:]}) + b\sum_{i=1}^{x}\nabla_{\mathcal{X}_t}\log q^{(C)}(\mathcal{X}_{t,[i,:,:]})\\ &\simeq a\sum_{j=1}^{y}s_\theta^{(I)}(\mathcal{X}_{t,[:,j,:]}) + b\sum_{i=1}^{x}s_\theta^{(C)}(\mathcal{X}_{t,[i,:,:]}),\end{aligned} \tag{15}$$

where $\mathcal{X}_{t,[:,j,:]}$ and $\mathcal{X}_{t,[i,:,:]}$ denote the j-th inline slice and the i-th crossline slice of $\mathcal{X}_t$, respectively. Both a and b are weights, while $q^{(I)}$ and $q^{(C)}$ represent the modeling of slices along different profiles (inline and crossline, respectively), but both are derived from the same underlying model. The corresponding sampling equation is as follows:

$$\begin{cases}\Sigma s_\theta^{(I)}(\mathcal{X}_{t,[:,j,:]}), & P = a/(a+b)\\ \Sigma s_\theta^{(C)}(\mathcal{X}_{t,[i,:,:]}), & Q = b/(a+b)\end{cases}, \tag{16}$$

where P and Q denote the probability of performing the sampling step. Notably, P and Q can be defined using either a stochastic mechanism or a structured interval-based approach to balance computational efficiency with spatial continuity.

We introduce a dual strategy (SeisDPS-3D) to improve generation quality. This algorithm applies SeisDPS constraints along inline profiles while using unconditional diffusion in crossline profiles. This unconditional sampling leverages the prior information learned by the 2D diffusion model as 3D lateral constraints in orthogonal directions, addressing the issue of poor spatial continuity in the crossline profile caused by the inline sampling emphasis. In this paper, the inline profile is selected as the main profile because the crossline spacing is often smaller than the inline spacing in real-world data, allowing for the use of finer spatial information to enhance the overall 3D deconvolution. Sampling is performed using a structured interval-based approach, and a hyperparameter K is introduced to control the frequency of this improvement. We define the remainder when the time step t is divided by K as ϕ. For example, when $K = 3$, every three SeisDPS steps along the inline profile are corrected by one unconditional sampling step in the crossline profile. This approach balances quality across profiles while maintaining computational efficiency. Specifically, in contrast to the profile-by-profile method, where complete sampling is performed on a single slice before stitching, the enhanced algorithm performs consistent sampling steps across multiple slices, followed by stitching them into complete 3D data before initiating the next iteration of sampling. The complete procedure of SeisDPS-3D is detailed in Algorithm 3, while the sampling process workflow is illustrated in Fig. 1.

Algorithm 3 SeisDPS-3D

Require: $T, \mathcal{D} \in \mathbb{R}^{x\times y\times t}, \mathcal{L} \in \mathbb{R}^{x\times t}, \{\boldsymbol{\sigma_t}\}_{t=1}^{N}, \{\zeta_t\}_{t=1}^{N}, \xi, K, \lambda, s_\theta$

$\mathcal{M}_T \sim \mathcal{N}(0, \boldsymbol{I}) \in \mathbb{R}^{x\times y\times t}$
for $t = T$,..., 1 **do**
  $\mathbf{Z} \sim \mathbf{N}(\mathbf{0}, \boldsymbol{I})$ **if** $t > 1$, **else** $\mathbf{Z} = 0$
  $\phi \leftarrow t \bmod K$
  if $\phi \neq 0$ **then**
    for i in $1 : x$ **do** ▷ processing (y, t) slices
      $\mathbf{M}_t \leftarrow \mathcal{M}_t[i, :, :]$
      $\mathbf{D}_i \leftarrow \mathcal{D}[i, :, :]$
      $\hat{\mathbf{M}}_0 \leftarrow \frac{1}{\sqrt{\bar{\alpha}_t}}\left(\mathbf{M}_t + (1-\bar{\alpha}_t)\boldsymbol{s}_\theta\left(\mathbf{M}_t, t\right)\right)$
      $\mathbf{M}'_{t-1} \leftarrow \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{M}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}\hat{\mathbf{M}}_0 + \boldsymbol{\sigma_t}\mathbf{Z}$
      $\hat{\mathbf{M}}'_0 \leftarrow \mathcal{S}(\hat{\mathbf{M}}_0, \lambda)$
      $\mathbf{M}_{t-1} \leftarrow \mathbf{M}'_{t-1} - \zeta_t \nabla_{\mathbf{M}_t} \|\mathbf{D}_i - \mathcal{L}(\hat{\mathbf{M}}'_0)\|_2^2 - \xi \nabla_{\mathbf{M}_t} \|\hat{\mathbf{M}}'_0\|_1$ ▷ SeisDPS
      $\mathcal{M}_{t-1}[i, :, :] \leftarrow \mathbf{M}_{t-1}$
    end for
  else
    for j in $1 : y$ **do** ▷ processing (x, t) slices
      $\mathbf{M}_t \leftarrow \mathcal{M}_t[:, j, :]$
      $\mathbf{M}_{t-1} \leftarrow \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{M}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{s}_\theta(\mathbf{M}_t, t)\right) + \boldsymbol{\sigma_t}\mathbf{Z}$ ▷ DDPM
      $\mathcal{M}_{t-1}[:, j, :] \leftarrow \mathbf{M}_{t-1}$
    end for
  end if
end for
return $\mathcal{M}_0$

3 Experiment

This section first demonstrates the feasibility of the proposed two methods using 2D and 3D synthetic data, respectively. Then, we evaluate the effectiveness of our method using field data sets. In addition, we choose the sparse spike deconvolution method for comparison to verify our method, which is implemented using the fast iterative shrinkage thresholding algorithm (FISTA) (Beck and Teboulle 2009). We will refer to this deconvolution method as FISTA. FISTA is a more general method for ℓ_1 regularization of linear inverse problems.

3.1 Synthetic Data Experiment

3.1.1 Synthetic Training Data Construction

We first generated a noise-free 2D reflectivity training data set containing 5000 samples, each with dimensions of 288 × 992, representing 288 traces, each containing 992 time points with a sampling interval of 2 ms. This data set is carefully modeled to capture the

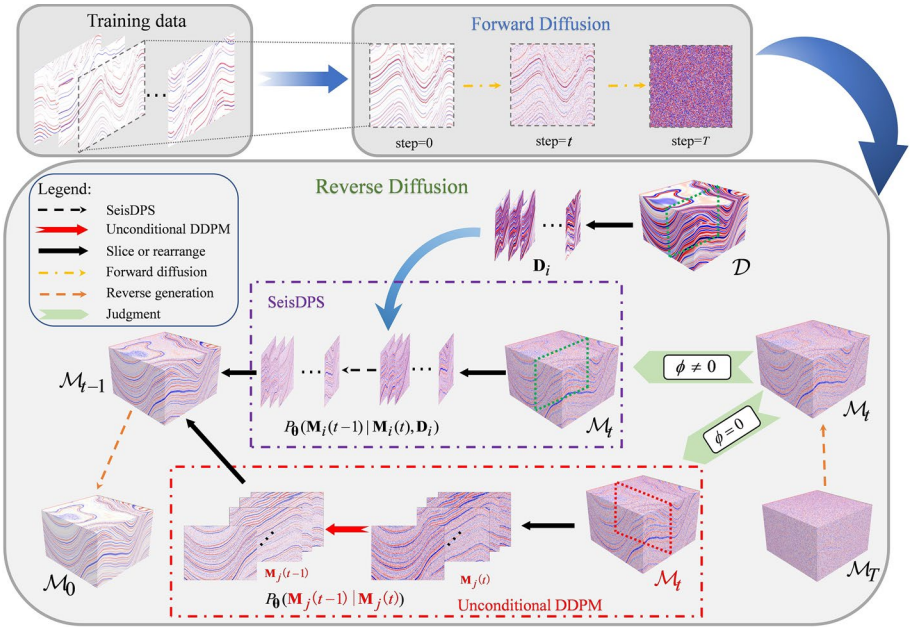

Fig. 1 The workflow of SeisDPS-3D. In the forward process, the training data is transformed by adding noise to a random Gaussian distribution. In the reverse generation process, a 3D Gaussian noise volume ($\mathcal{M}_T$) of dimensions identical to the original data ($\mathcal{D}$) is initialized. Following each sampling, a conditional branching is executed: 1) If $\phi \neq 0$, $\mathcal{M}_t$ and $\mathcal{D}$ are sliced along the inline direction. Then, perform a single SeisDPS step. Upon completion of sampling, the inline slices are reconstructed in their original sequence to produce $\mathcal{M}_{t-1}$. 2) If $\phi = 0$, $\mathcal{M}_t$ is sliced along the crossline direction, and a single unconditional DDPM step is applied to mitigate the issue of poor continuity in the crossline profiles due to the use of the inline profiles as the primary profiles. After sampling, the crossline slices are reconstructed in their original sequence. Subsequently, sampling iterations are conducted to obtain the final result $\mathcal{M}_0$

features of actual geological structures, including fault structures, horizontal layers, and curved strata. A portion of the data from the dataset is presented in Fig. 2.

In addition, a reflection coefficient with richer information is selected to generate seismic records for subsequent tests, as it better represents real-world data. The 2D synthetic reflection coefficients have dimensions of 256×512, including 256 traces, each containing 512 time points. It is then convolved with a 30 Hz Ricker wavelet to generate synthetic seismic data. We utilize the Ricker wavelet (Hosken 1988; Gholamy and Kreinovich 2014) because of its versatility. Figure 3 shows the process of generating synthetic seismic data, where (a), (b), and (c) represent reflection coefficients, Ricker wavelet, and synthetic data, respectively.

3.1.2 Training and Parameter Setting

In the experiment, 90% of the data set is used for training, and the remaining 10% are reserved for testing. Before training, each dataset is randomly cropped to a patch of 256×256 samples. The number of iterations, T, in the forward process is set to 1000. The Adam optimizer (Kingma et al. 2014) is selected as the optimization algorithm with a learning rate of 10^{-4}. The training phase is conducted on an NVIDIA GeForce RTX 3090 GPU

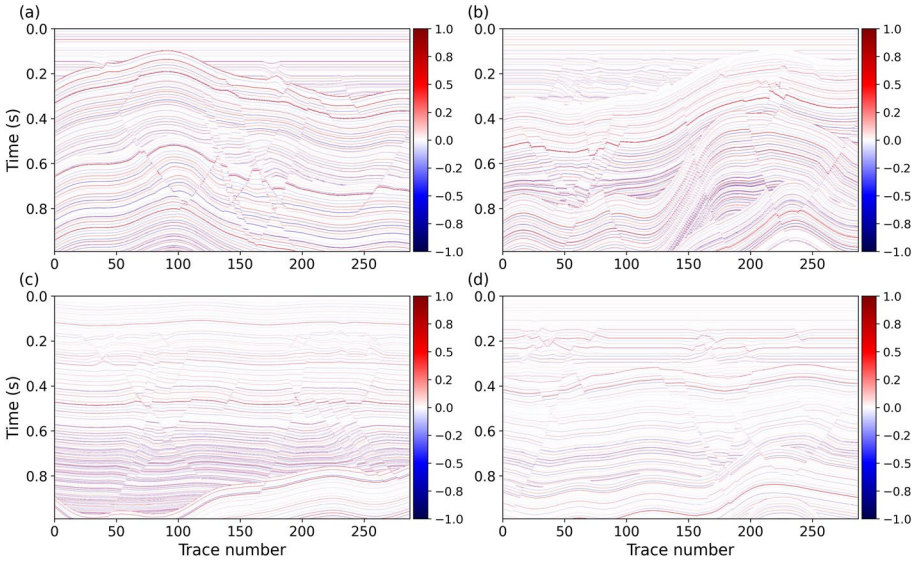

Fig. 2 A portion of the data in the training dataset

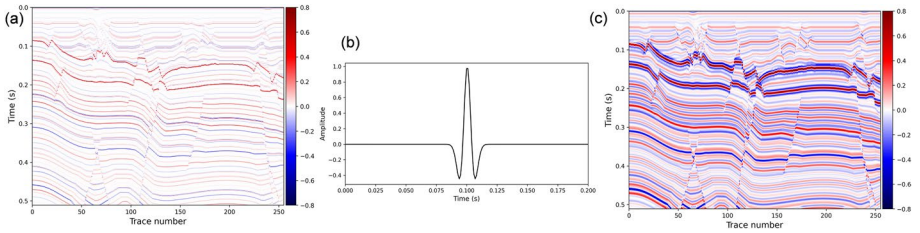

Fig. 3 Seismic data synthesis process. **a**, **b** and **c** represent the reflection coefficients, the Ricker wavelet, and the synthetic data, respectively

for 1,000,000 steps. All models used in subsequent experiments are based on this data set (including 2D, 3D, synthetic, and field data).

3.1.3 2D Data Experimental Results

Our initial evaluation is performed on noise-free synthetic data, with results detailed in the Supplementary Materials due to space limitations. Figure 28 demonstrates the efficacy of the SeisDPS method. Single-trace results and multi-trace normalized average amplitude spectra are further presented in Figs. 26 and 27, respectively.

Although real poststack seismic data generally exhibit a high signal-to-noise ratio (SNR), they inevitably contain residual noise. To evaluate the robustness of the proposed method, we augment clean synthetic data with noise at a fixed SNR. We add a noise matrix $\boldsymbol{N}$ to the clean signal $\mathbf{D}$ such that $\mathbf{D}_{\boldsymbol{N}} = \mathbf{D} + \boldsymbol{N}$. For this, we adopt the definition

$$\mathrm{SNR} = \frac{||\mathbf{D}||_2^2}{\gamma^2 ||\boldsymbol{N}||_2^2}, \tag{17}$$

where γ is a scalar that determines the desired SNR, and $\boldsymbol{N}$ is sampled from a standard normal distribution. For a quantitative evaluation of the results, we calculate the reconstruction Structural Similarity Index Measure (SSIM), Mean Squared Error (MSE), and accuracy

$$\text{Accuracy (dB)} = 10 \log_{10} \frac{||\mathbf{M}||_2^2}{||\mathbf{M} - \mathbf{M}^*||_2^2}, \tag{18}$$

where $\mathbf{M}$ and $\mathbf{M}^*$ are the true and inverted reflectivity models, respectively.

In this section, due to the added noise, the step size η is set to 0.5, and the hyperparameters of λ and ξ are both set to 0.01 to encourage the generation of sparse solutions. In the FISTA algorithm, the regularization parameter μ is determined using the χ^2 criterion (Hennenfent et al. 2008) to ensure the data residual aligns with the noise level, rather than being set to a fixed value. The iterative process terminates when either the maximum number of iterations K_{iter} of 500 is reached, or the convergence tolerance falls below 10^{-5}. The choice of hyperparameters for the SeisDPS method is discussed in the discussion section.

The sensitivity of the proposed method to noise is evaluated by varying SNRs and injecting random noise based on the SNR definition. The results are summarized in Table 2. The FISTA algorithm is observed to be highly sensitive to noise, with accuracy deteriorating significantly as the SNR decreases. In contrast, the SeisDPS method exhibits only a slight decrease in accuracy, maintaining consistently high SSIM values and consistently low MSE values. These findings are further supported by Fig. 4, which shows that the SeisDPS method can reconstruct complete reflectivity coefficients even under extreme SNR conditions as low as 3 dB, albeit with minor detail loss. In contrast, the FISTA algorithm shows a marked increase in noise as the SNR decreases. Therefore, quantitatively and visually, the proposed SeisDPS method exhibits significantly greater noise robustness than the FISTA algorithm.

Moreover, to demonstrate the effectiveness of incorporating sparse regularization and soft thresholding, we compared the results obtained using only the DPS method and those achieved with SeisDPS. As outlined in Algorithm 1, DPS relies on a single hyperparameter. We ensure that this hyperparameter equals SeisDPS ($\eta = 0.5$) for a fair comparison. The only distinction between SeisDPS and DPS is that SeisDPS includes sparse regularization and soft thresholding. The comparison results are presented in Fig. 5. Although the DPS method can reconstruct the results well, it contains more artifacts, especially within the yellow box. Compared to the true reflection coefficients shown in Fig. 3a, the results of SeisDPS are notably superior to those obtained by DPS alone. This further illustrates that the introduction of sparse regularization and soft thresholding can significantly improve the deconvolution results, producing more satisfactory outcomes and validating the effectiveness of the proposed SeisDPS method.

Finally, we further demonstrated the robustness of the SeisDPS method to different seismic wavelets and varying wavelet frequencies. We ensure that only the seismic wavelet or its frequency varies in each test. Results are presented in Figs. 29–32 of the Supplementary Materials. These results indicate that the SeisDPS method applies to diverse seismic wavelets without specific frequency requirements.

Table 2 The impact of different input SNRs on output Accuracy (dB), SSIM, and MSE

	SeisDPS			FISTA		
SNR (dB)	Accuracy	SSIM	MSE	Accuracy	SSIM	MSE
21	3.40	0.8521	8.97e-4	3.52	0.8329	8.26e-4
18	3.35	0.8477	9.13e-4	3.34	0.8211	1.202e-3
15	3.31	0.8141	9.21e-4	2.02	0.7808	1.238e-3
12	3.22	0.8080	9.40e-4	1.77	0.7490	1.311e-3
9	3.08	0.7973	9.71e-4	1.31	0.7198	1.461e-3
6	2.87	0.7811	1.02e-3	0.52	0.6715	1.746e-3
3	2.58	0.7581	1.09e-3	−0.46	0.5880	2.196e-3

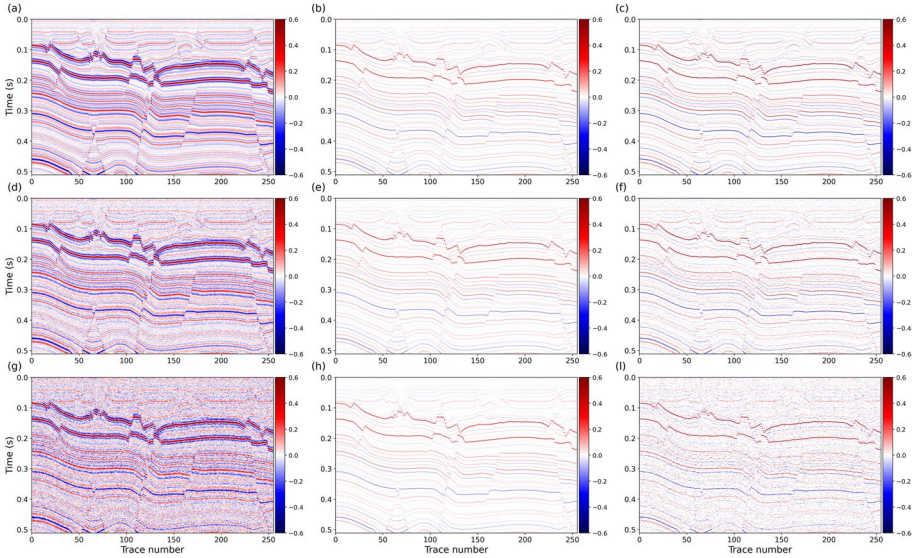

Fig. 4 Deconvolution results under convolution with a 30 Hz **Ricker wavelet** for different SNR. **a**, **d** and **g** show the synthetic data at SNR values of 18 dB, 9 dB, and 3 dB, respectively; **b**, **e**, and **h** present the corresponding SeisDPS results; **c**, **f**, and **l** show the results obtained using FISTA

3.1.4 3D Data Experimental Results

Although the effectiveness of the proposed SeisDPS method has been validated on 2D data, seismic data is inherently 3D. We use the open-source Synthoseis model (Merrifield et al. 2022) to generate 3D synthetic seismic data. This model has a 3D structure and multiple subsurface scenarios, making it particularly suitable for evaluating our 3D method. The generated reflection coefficients are convolved with a Ricker wavelet (Hosken 1988) to obtain synthetic seismic data. The data dimensions are $384 \times 384 \times 288$ (inline, crossline, and time, respectively), with a time sampling interval of 1 ms.

The SeisDPS-3D method has only one additional hyperparameter, K, compared to SeisDPS. In this section, the hyperparameters η, λ, and ξ remain the same as those used in SeisDPS for noisy data, set to 0.5, 0.01, and 0.01, respectively. The hyperparameter K is set

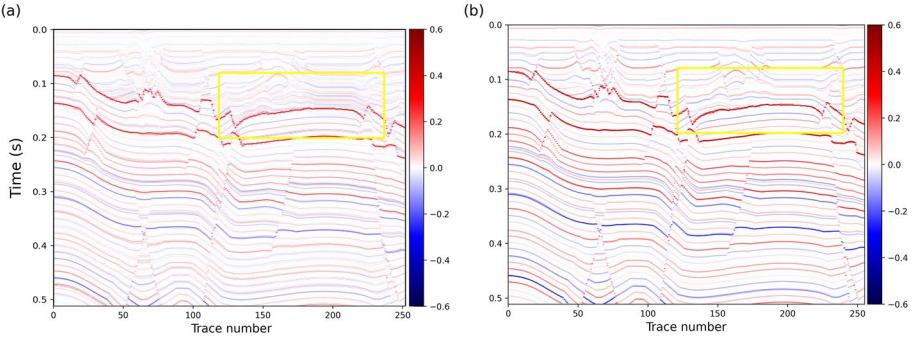

Fig. 5 Comparison of results between the DPS and SeisDPS methods, with **a** representing the results obtained by DPS and **b** representing those obtained by SeisDPS. The yellow box highlights that the SeisDPS method produces superior deconvolution results, while the DPS method exhibits some artifacts

to 5, which means that after every five iterations of the main profile SeisDPS sampling, an unconditional DDPM correction is applied to the results of another profile.

Figure 6 shows the results obtained by SeisDPS and FISTA for conventional 2D slice stitching, along with the results processed by the proposed SeisDPS-3D method. Figure 6b highlights the significant discontinuities in the crossline and time profiles resulting from processing along the inline slices by the SeisDPS method. Since FISTA processes data trace-by-trace, there is no significant difference in performance between the inline and crossline profiles. As shown in Fig. 6d, the FISTA results are better than those from SeisDPS. Figure 6c shows the results of SeisDPS-3D, which exhibits good continuity in both the inline and crossline profiles. To further highlight the effectiveness of the SeisDPS-3D method, the next step is to present the slice results.

Figure 7 shows the results for inline 300. The original synthetic seismic data are shown in Fig. 7a, while Figs. 7b–d display the results obtained using the SeisDPS method (processed slice by slice), the proposed SeisDPS-3D method, and the FISTA algorithm, respectively. Due to the inline-by-inline processing approach, all three methods show commendable performance, as highlighted in the yellow box, displaying excellent detail resolution. Overall, compared to 2D methods, the proposed 3D method produces smoother results. For fault structures, FISTA performs poorly compared to our proposed method, as shown in the green box, while the SeisDPS-3D method provides a more explicit fault description.

Figure 8 shows the result at crossline 150. Figure 8b highlights the inherent flaw of using the 2D slice method for processing 3D data: poor performance in the other dimension, neglecting the inherent spatial continuity of 3D data. The SeisDPS-3D method significantly enhances the continuity of the results while effectively preserving stratigraphic details. The yellow box highlights that the SeisDPS-3D method reveals more distinct fault information, making it easier to identify. Although the FISTA algorithm's results align well with the overall structure of the original data, its poor spatial continuity results in less effective fault information processing compared to SeisDPS-3D. This further demonstrates the effectiveness of the SeisDPS-3D method.

Figure 9 shows the time-slice results at time = 200 ms. The comparison between the SeisDPS-3D method (Fig. 9c) and the FISTA algorithm (Fig. 9d) underscores our

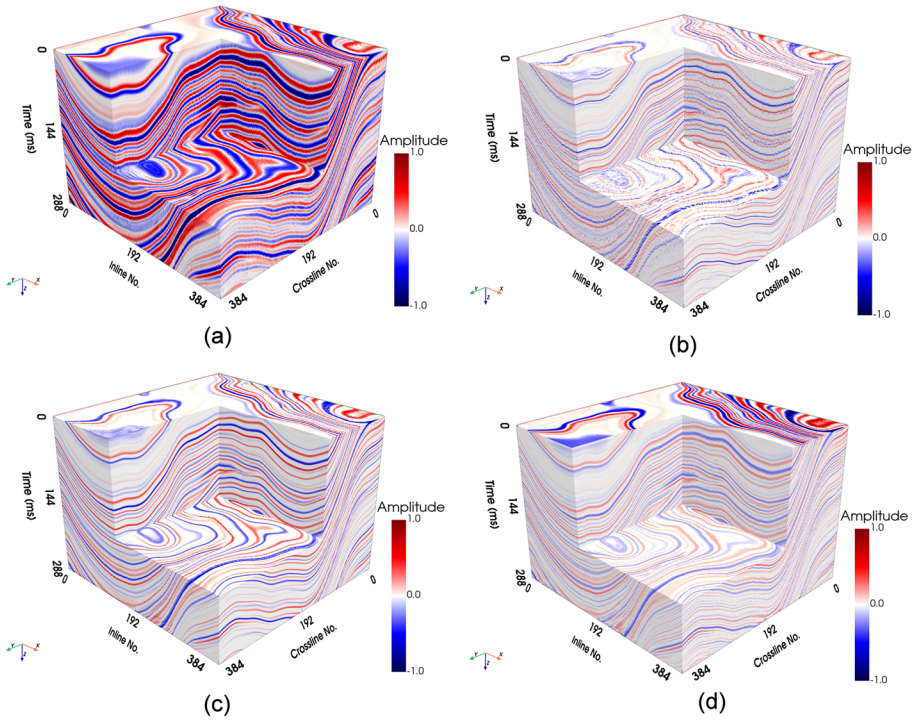

Fig. 6 Deconvolution results for 3D synthetic data. **a** Synthetic seismic data, **b** results by the SeisDPS method, which processes the data slice by slice, **c** results from the proposed SeisDPS-3D method, and **d** results derived from the FISTA algorithm

effectiveness. Our proposed method effectively addresses spatial continuity issues and exhibits superior performance in specific details, highlighting its significant potential.

The SeisDPS-3D method demonstrates satisfactory results from the comprehensive 3D performance evaluation to the slice display. Notably, by performing a single pre-training on 2D data and fine-tuning the hyperparameters, we obtained good results on 3D data, addressing the memory challenges of directly processing 3D data while validating the SeisDPS-3D method's strong generalization ability.

3.2 Field Data Examples

3.2.1 Field 2D Data Experimental Results

This section uses a 2D real seismic data set to validate the effectiveness of the proposed SeisDPS method. The seismic data dimensions are 288×496, with 288 traces and each trace containing 496 time points with a sampling interval of 1 ms. The model used for testing is trained with the synthetic reflection coefficients mentioned earlier. Figure 10 illustrates the estimated statistical wavelet and frequency spectrum. The step size η is set to 0.3, and the hyperparameters λ and ξ are set to 0.01. In the FISTA algorithm, the regularization parameter μ is set to 0.05.

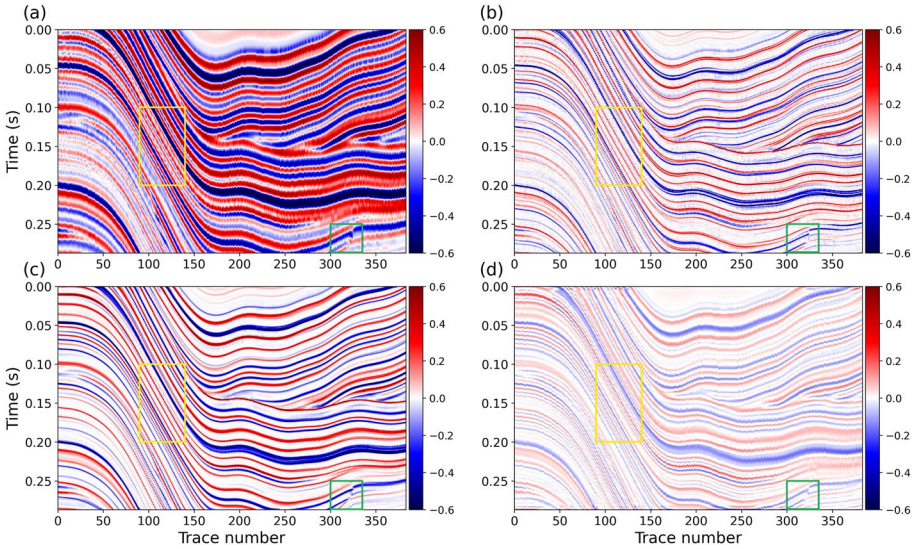

Fig. 7 Inline results (inline = 300) of synthetic seismic data. **a** Original synthetic seismic data, **b** results by the SeisDPS method, which processes the data slice by slice, **c** results from the proposed SeisDPS-3D method, and **d** results derived from the FISTA algorithm. The yellow box highlights that all three methods achieve satisfactory performance on inline profiles, primarily due to the inline-by-inline processing approach. In contrast, the green box emphasizes that the SeisDPS-3D method excels at revealing fault information

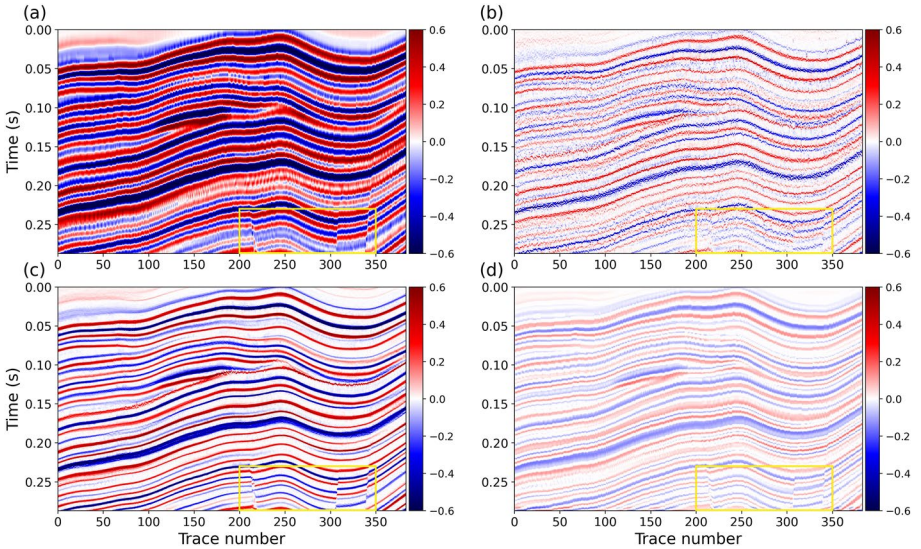

Fig. 8 Crossline results (crossline = 150) of synthetic seismic data. **a** Original synthetic seismic data, **b** results by the SeisDPS method, which processes the data slice by slice, **c** results from the proposed SeisDPS-3D method, and **d** results derived from the FISTA algorithm. The yellow box highlights the visible fault structures in the SeisDPS-3D method's crossline profiles, demonstrating consistency with the original data

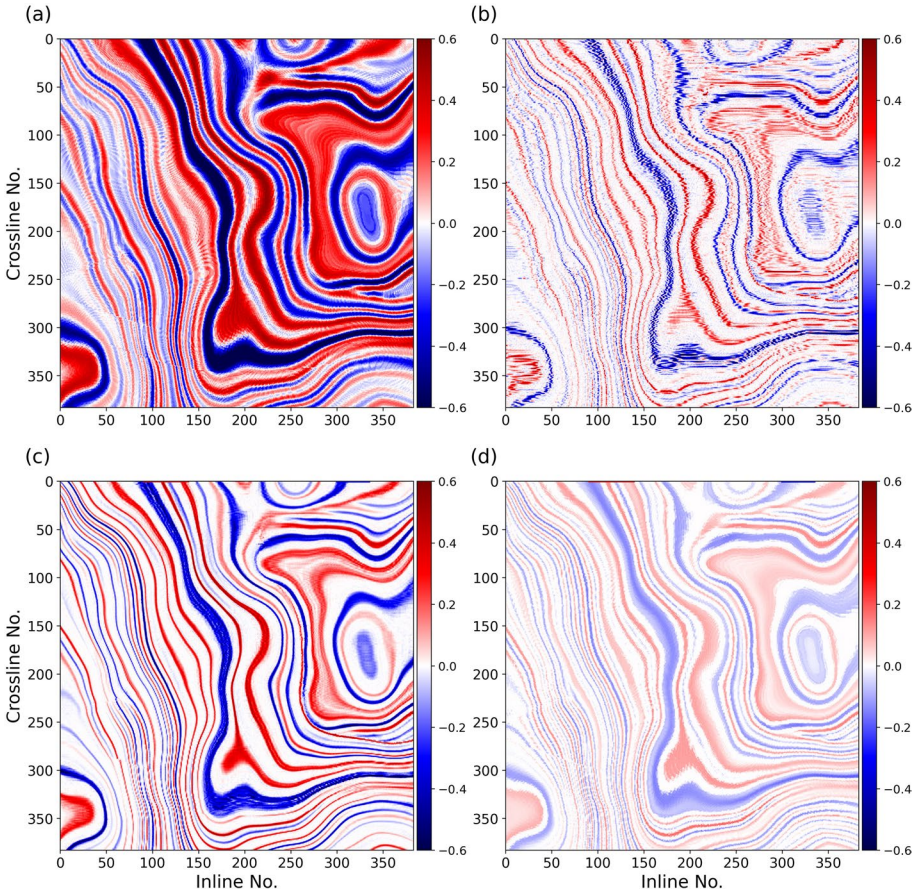

Fig. 9 Time-slice results (time = 200 ms) of synthetic seismic data. **a** Original synthetic seismic data, **b** results by the SeisDPS method, which processes the data slice by slice, **c** results from the proposed SeisDPS-3D method, and **d** results derived from the FISTA algorithm

Figure 11 shows the results for the real data, where Fig. 11a, b, and c represent the real data, the results of the proposed SeisDPS method, and the FISTA results, respectively. Figures 11d, e, and f show magnified sections of the yellow boxes. The overall image demonstrates that both methods successfully reconstruct the reflection coefficients of the target data. However, the zoomed-in image reveals that the SeisDPS method significantly enhances lateral continuity, clearly outperforming the FISTA algorithm. This is because the SeisDPS method processes the data across the entire 2D plane.

Figure 12 presents the Pearson Correlation Coefficient (PCC) results between the synthetic trace data and the original data. Trace number 75 lies precisely within the yellow region highlighted in Fig. 11. Although both methods yield high PCC values between the synthetic trace and original data, the SeisDPS method demonstrates better consistency with the original data in the details, specifically in the area highlighted by the yellow box in Fig. 12.

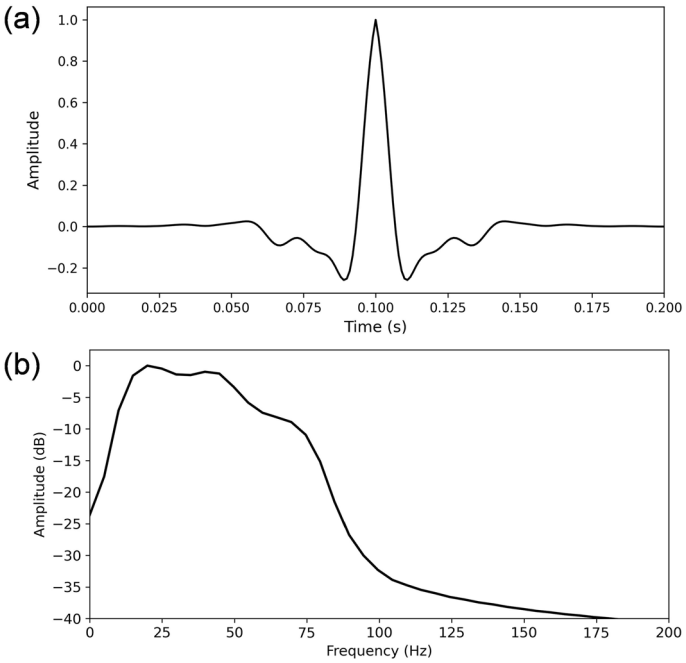

Fig. 10 Characterization of the extracted seismic wavelet. **a** Extracted statistical seismic wavelet from the real field 2D seismic data. **b** Amplitude spectrum

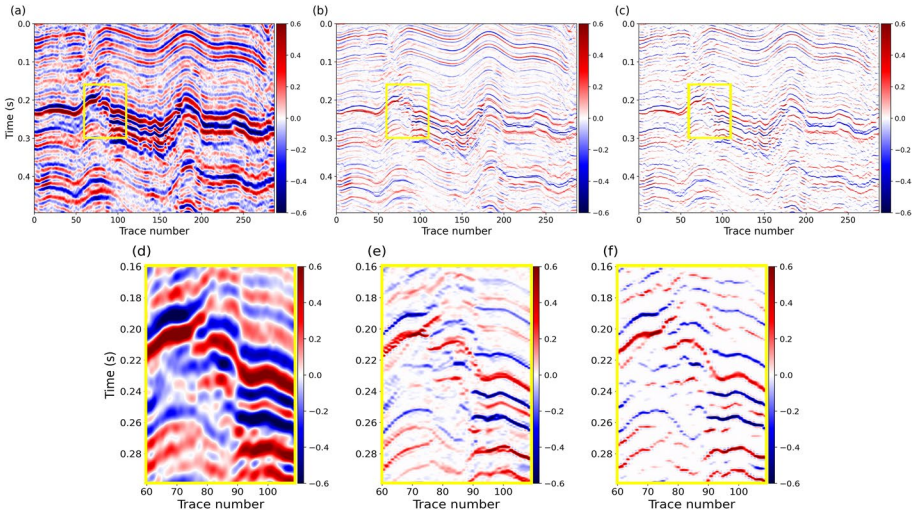

Fig. 11 Deconvolution results for real field 2D data. **a** Real field 2D data, **b** results obtained using the proposed method (SeisDPS), **c** results derived from the FISTA algorithm, **d**– **f** show zoomed-in views of the regions outlined by the yellow boxes in **a**– **c**, respectively. The yellow box demonstrates that the SeisDPS method provides a substantial enhancement in lateral continuity compared to the FISTA algorithm

The normalized multitrace amplitude spectra are shown in Fig. 13 to further compare the FISTA algorithm and the SeisDPS method. The red line represents the original

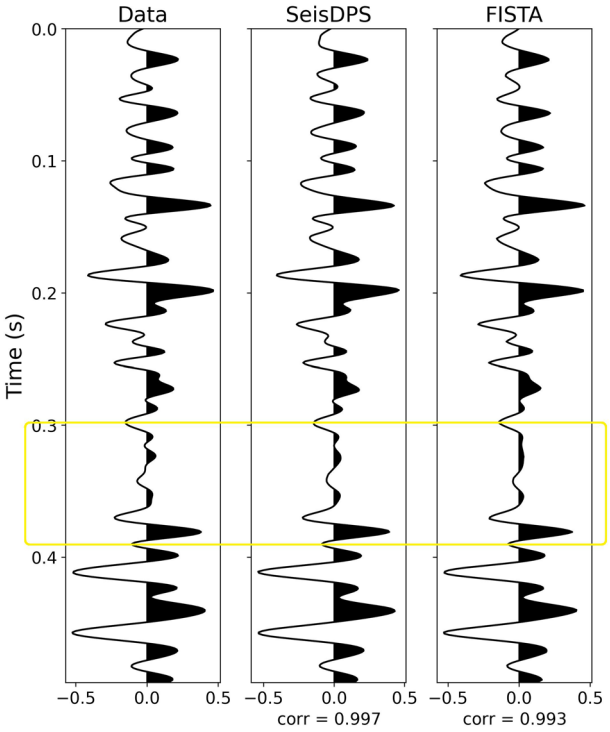

Fig. 12 Representative trace (trace number = 75) from the real 2D data showing the Pearson correlation coefficient (bottom) between the synthetic trace data and the input data. **a** Real field 2D data, **b** results obtained using the proposed method (SeisDPS), **c** results derived from the FISTA algorithm. The yellow box demonstrates that the re-convolved result obtained using the SeisDPS method exhibits better conformity with the original data

data, and the results of SeisDPS and FISTA are represented by the blue and green lines, respectively. It can be observed that both methods effectively enhance high-frequency energy. However, the SeisDPS method maintains higher fidelity in the low frequencies, more accurately preserving their original content. Nevertheless, it shows a slight reduction in high frequencies compared to FISTA.

3.2.2 Field 3D Data Experimental Results

This section demonstrates the effectiveness of our proposed method using a 3D real seismic data set. The spatial dimensions consist of 256 inlines and 256 crosslines, with a bin size of 40 m × 20 m, covering an area of 10.24 km × 5.12 km. The time sampling is 2 ms. Figure 14 illustrates the estimated statistical wavelet and frequency spectrum. The step size η is set to 0.6, and the hyperparameters λ and ξ are set to 0.01. In the FISTA algorithm, according to χ^2 criterion, the regularization parameter μ is set to 0.05.

Figures 15b and c show SeisDPS-3D and FISTA deconvolution processing results, respectively. Compared to the Real Field data (Fig. 15a), the SeisDPS-3D method successfully reconstructs satisfactory results and provides spatial continuity that more closely resembles the Real Field data. We also present inline, crossline, and time profiles to illustrate the SeisDPS-3D method's effectiveness further.

Figure 16 shows the inline profile results at 2 km. Figure 16a shows the original profile. The results of SeisDPS-3D (Fig. 16b) are more satisfactory than those of FISTA (Fig. 16c), significantly enhancing lateral continuity and improving spatial energy consistency. As highlighted in the yellow box, the SeisDPS-3D method achieves higher resolution in strong

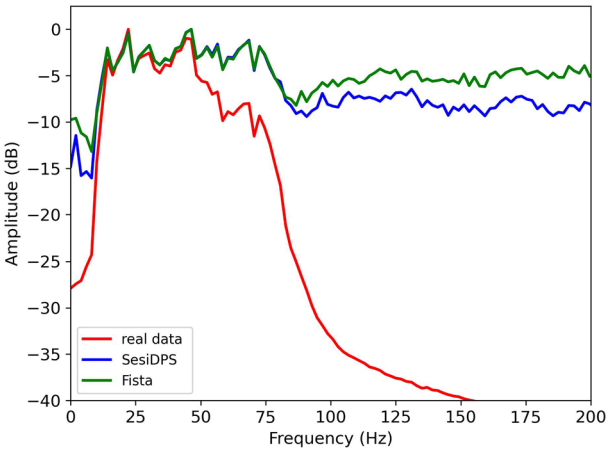

Fig. 13 Multi-trace normalized average amplitude spectrum, where the red, blue, and green lines represent the real field data, the results from the SeisDPS method, and the results from the FISTA algorithm, respectively

reflection areas, whereas the FISTA algorithm exhibits noticeable blurring. Figure 16d, e, and f are magnified views of the black box region. As indicated by the black arrows, compared to the FISTA algorithm, the SeisDPS-3D method achieves higher resolution and better aligns with the actual structure. Additionally, the SeisDPS-3D method exhibits strong fault detection capabilities. As shown by the green arrows in Fig. 16g, h, and l, a fault is visible in the original seismic data. However, in the results, the FISTA algorithm fails to

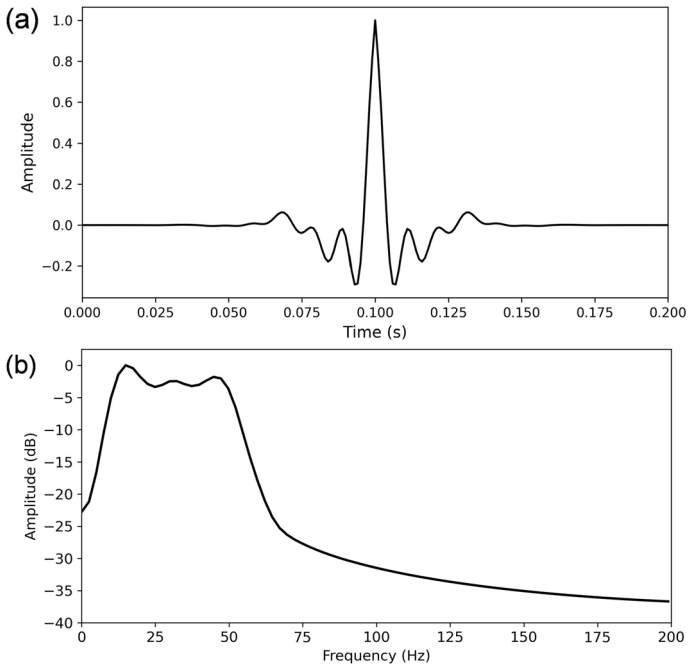

Fig. 14 Characterization of the extracted seismic wavelet. **a** Extracted statistical seismic wavelet from the real field 3D seismic data. **b** Amplitude spectrum

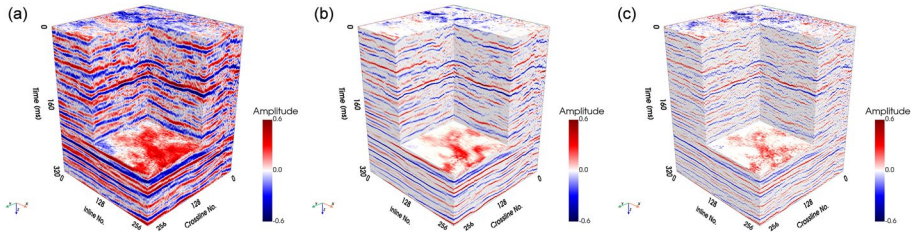

Fig. 15 Deconvolution results for real field 3D seismic data. **a** Real field seismic data, **b** results from the proposed SeisDPS-3D method, and **c** results from the FISTA algorithm

reveal a distinct fault structure, whereas the SeisDPS-3D method produces a fault structure that matches the original data more closely.

Figure 17 shows the crossline profile results at 2.6 km. From the results of SeisDPS-3D (Fig. 17b), the SeisDPS-3D method significantly improves the poor continuity problem of another profile caused by conventional 2D methods, demonstrating superior performance. Looking at the details, the results in the yellow box indicate that the SeisDPS-3D method is superior. Figures 17d–f present magnified views of the green box region. The green arrows highlight artifacts introduced by the FISTA algorithm, which do not align with the structural features of the original data. In contrast, the SeisDPS-3D method more accurately preserves the structural characteristics of the original data. Figures 17g, h, and l show the results within the black box. The black arrows further emphasize the discontinuity in the FISTA algorithm, demonstrating the reliability of our approach.

Further examination of time slices at 1500 ms in Fig. 18 emphasizes the SeisDPS-3D method's superior resolution recovery capabilities. After processing, the depth-time slice reveals results more closely aligned with reality than the FISTA algorithm. The SeisDPS-3D method preserves structural features that align with the original data and demonstrates superior spatial continuity, highlighting finer spatial variations.

To further validate the effectiveness of SeisDPS-3D, we calculated the C3 algorithm-based coherence attributes of the deconvolution results. Since the SeisDPS-3D method significantly improved the results of the crossline profile, we focus on presenting the crossline profile. Figure 19 shows the crossline profile results at 2.6 km. Compared to the C3 coherence attribute of the original data shown in Fig. 19a, the resolution in Fig. 19b shows a definite improvement. Specifically, in the fault zone between 0 km and 2 km, the enhanced resolution in Fig. 19b allows for better differentiation of adjacent faults. However, for the strong reflector layer between 1.8 s and 1.9 s, as indicated by the yellow box in Fig. 17, which should be a continuous stratum, Fig. 19c exhibits numerous fake faults. Moreover, Fig. 19c displays a large amount of fault information between 6 km and 10 km, while the original seismic data in Figure 17a exhibits good continuity in this region. These comparisons demonstrate that the SeisDPS-3D method ensures better fidelity.

Figure 20 presents the local similarity between the reconvolution results and the real field 3D data for the crossline profile at 2.6 km. The SeisDPS-3D method exhibits noticeably lower local similarity at the boundaries. However, since the color bar range is set from 0.9 to 1, the accuracy remains relatively high even at these boundary regions. Furthermore, the SeisDPS-3D method shows low local similarity in 1.5–1.8 s areas, but it can be observed from Fig. 17 that these areas predominantly correspond to regions with more faulting. Overall, the crossline profile results are slightly better for the FISTA algorithm

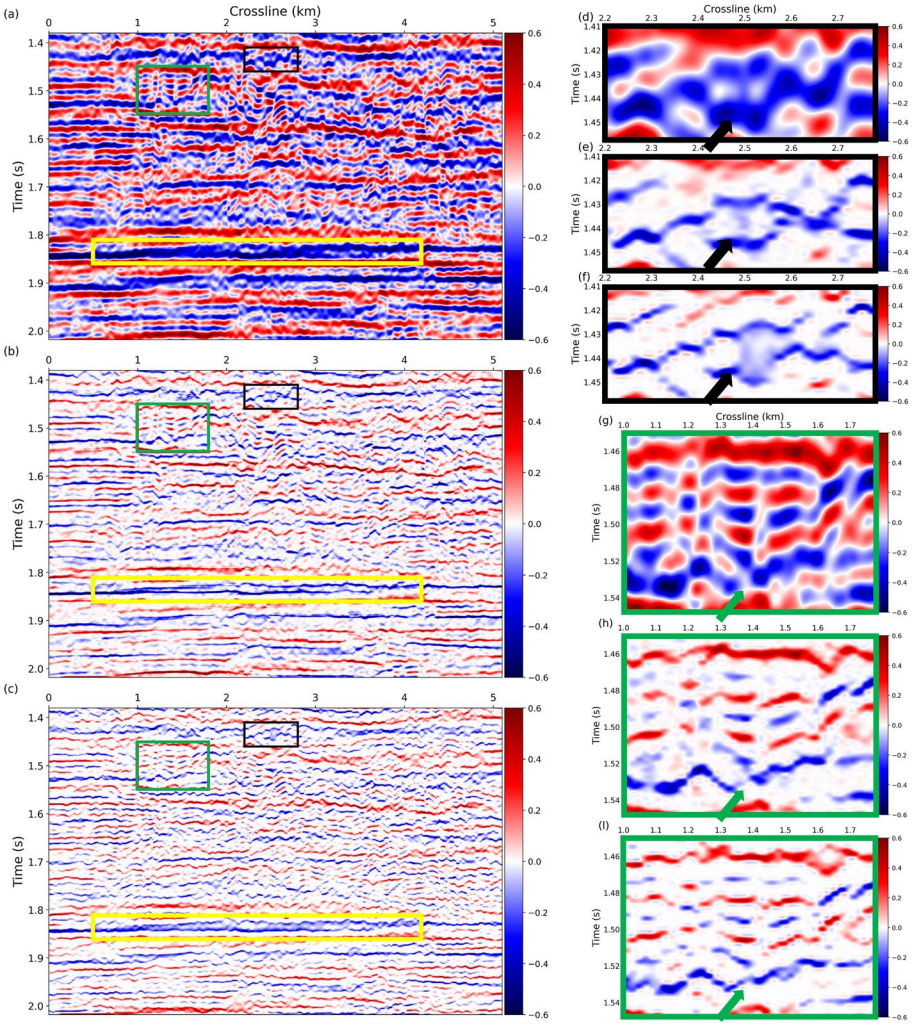

Fig. 16 Inline profile results at 2 km. **a** Real field seismic data, **b** results from the proposed SeisDPS-3D method, and **c** results from the FISTA algorithm. Figures **d**, **e**, and **f** are magnified views of the black box regions in **a**– **c**, and **g**, **h**, and **l** are magnified views of the green box regions in **a**– **c**. The yellow box demonstrates the SeisDPS-3D method's improved resolution in strong reflection zones, contrasting with the blurring observed with FISTA. Further reinforcing the SeisDPS-3D method's high resolution, the black box highlights a region (indicated by the black arrow) where FISTA suffers from artifacts. The green box demonstrates the SeisDPS-3D method's superior fault detection capabilities compared to FISTA. A clear fault is present in the original data at the location of the green arrow, which is identifiable in the SeisDPS-3D method but almost invisible in the FISTA result

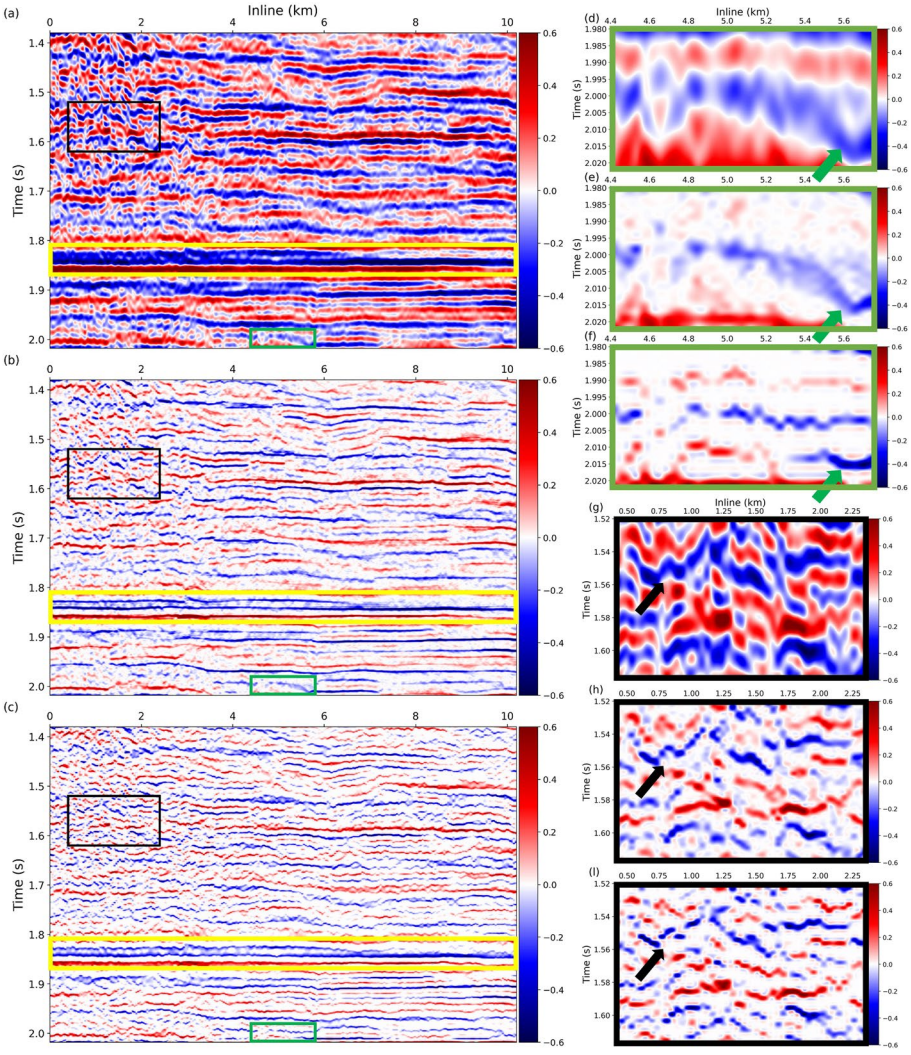

Fig. 17 Crossline profile results at 2.6 km. **a** real field seismic data, **b** results from the proposed SeisDPS-3D method, and **c** results from the FISTA algorithm. Figures **d**, **e**, and **f** are magnified views of the green box regions in **a**- **c**, and **g**, **h**, and **l** are magnified views of the black box regions in **a**–**c**. The yellow box emphasizes the lateral continuity and resolution enhancement provided by the SeisDPS-3D method. This is further illustrated in the black box, where the black arrow highlights the SeisDPS-3D method's significantly superior lateral continuity compared to FISTA. In addition, the green box demonstrates the reliability of the SeisDPS-3D method. The green arrow points to a region where FISTA produces structures inconsistent with the original data, while the SeisDPS-3D method faithfully retains the structural characteristics of the original data

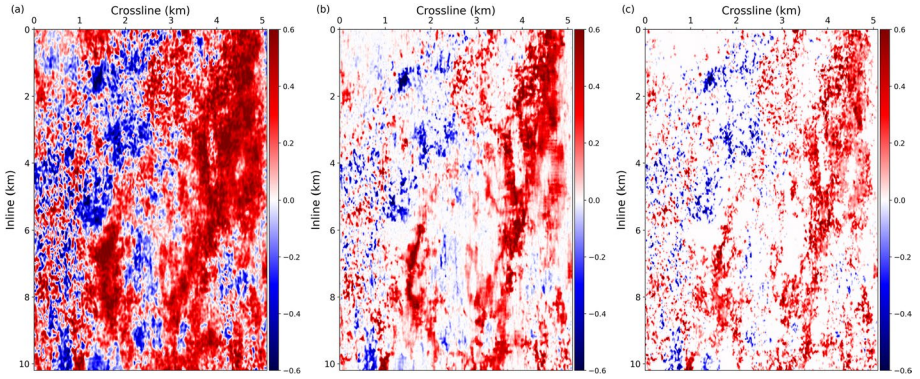

Fig. 18 Time-slice results at 1500 ms. **a** Real field seismic data, **b** results from the proposed SeisDPS-3D method, and **c** results from the FISTA algorithm

than for the SeisDPS-3D method, which may be attributed to the processing being primarily focused along the inline profile.

Finally, we further evaluated the two methods from a spectral perspective. Figure 21a displays the multi-trace normalized average amplitude spectrum for the real field 3D data and the deconvolution results. Both methods effectively broaden the spectrum, with the SeisDPS-3D method showing a slight advantage in the 60 Hz to 80 Hz range. However, both methods exhibit some deviation from the original data in the 20 Hz to 40 Hz range, although their overall trends remain consistent with the original data spectrum. This discrepancy is likely caused by the inaccuracy of the extracted seismic wavelet and the non-stationarity of the seismic data. Furthermore, the spectral ratios in Fig. 21b demonstrate that while both methods follow a similar trend, SeisDPS-3D provides a more significant boost at higher frequencies.

4 Discussion

4.1 Different Hyperparameters

Both SeisDPS and SeisDPS-3D involve several hyperparameters. Specifically, the hyperparameter η determines the step size, λ is the threshold in the soft-thresholding method, and ξ is the regularization coefficient in the coefficient regularization. The selection of hyperparameters is a key factor affecting the reconstruction results. To further illustrate the impact of different hyperparameters, we test on 2D data with noisy data at an SNR of 18 dB. We apply Eq. 18 to analyze the results quantitatively for ease of comparison.

Figure 22 shows the impact of different hyperparameters on the results. First, when $\xi = 0.01$, that is, the regularization parameter is 0.01, the results (Fig. 22a) indicate that increasing the step size η significantly improves the reconstruction results. However, λ and η are not more effective at larger values; both have an optimal value that promotes generating the best results. When λ is 0.005 and 0.01, the quantitative results are not significantly different, so $\lambda = 0.01$ is used in calculating Fig. 22b. In Fig. 22b, it is clear that after adding the regularization term, the accuracy of the results is significantly improved, which also

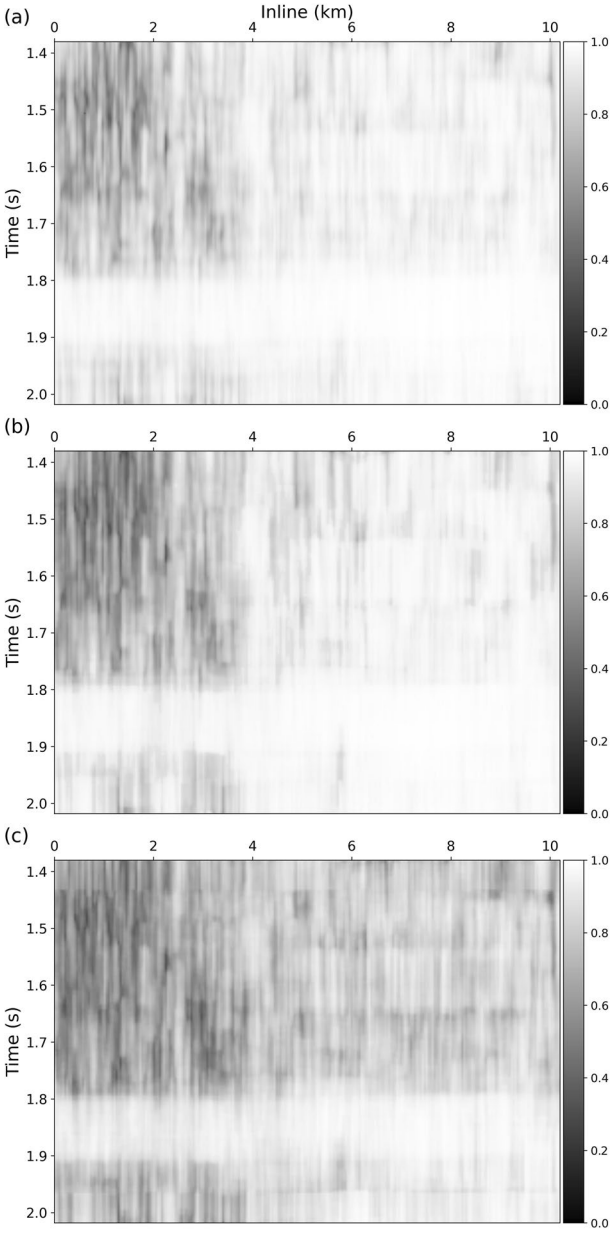

Fig. 19 Comparison of C3 coherence for a crossline profile (crossline = 2.6 km) of real field data. **a** Real field seismic data, **b** results from the proposed SeisDPS-3D method, and **c** results from the FISTA algorithm

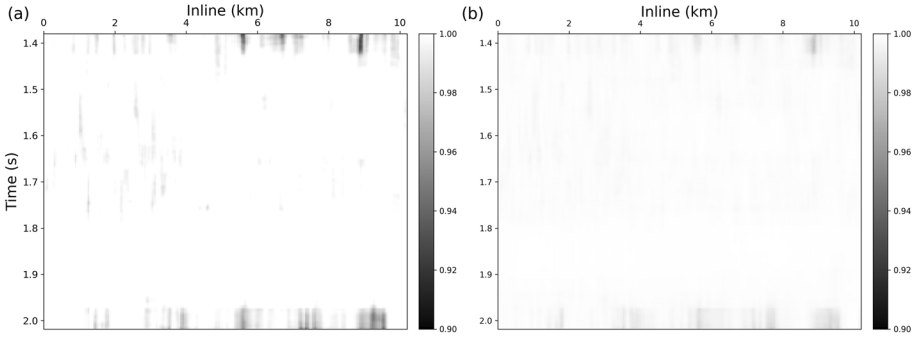

Fig. 20 Local similarity (crossline = 2.6 km) between the re-convolved results and the real field 3D data. **a** Results from the proposed SeisDPS-3D method, and **b** results from the FISTA algorithm

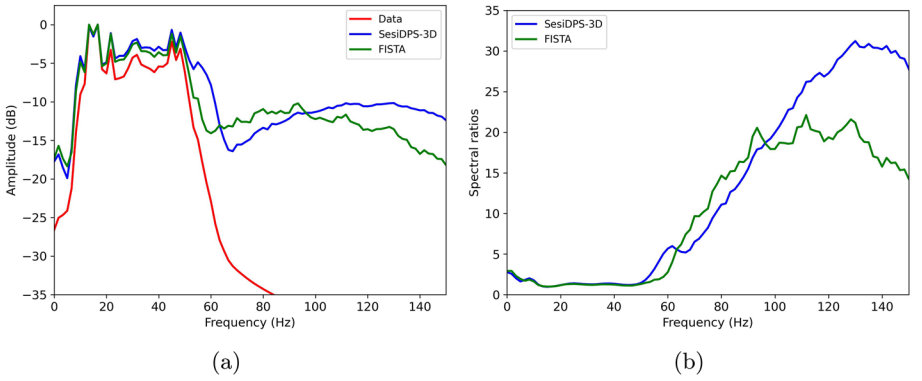

Fig. 21 Comparative spectral analysis of the real field data and the corresponding deconvolution results. **a** Multi-trace normalized average amplitude spectrum, where the red, blue, and green lines represent the real field 3D data, the results from the SeisDPS-3D method, and the results from the FISTA algorithm, respectively. **b** Spectral ratios, where the blue and green lines represent the SeisDPS-3D method and the FISTA algorithm, respectively

verifies the effectiveness of adding regularization. As in Fig. 22a, it can also be concluded that η has an optimal value. Moreover, λ also has an optimal value. Finally, the impact of λ and ξ is analyzed when $\eta = 0.5$ (Fig. 22c). Figure 22c clearly shows that ξ has the most significant impact on the results, and λ has two values with similar performance.

Regarding the additional hyperparameter K required for 3D data processing, we performed a thorough sensitivity analysis to determine its optimal value. Since the primary profile in our study is the inline profile, the K value primarily affects the crossline profile. To maintain conciseness and clarity, Fig. 23 presents representative results that illustrate the impact of this parameter. As shown in Fig. 23, the reconstruction performance exhibits a non-monotonic trend as K increases. At $K = 3$ (Fig. 23b), the constraints are insufficient, resulting in significant lateral discontinuities. Increasing the value to $K = 5$ (Fig. 23c) yields the most favorable results, where the spatial continuity of reflection coefficients is significantly enhanced and closely matches the original synthetic data, particularly in the highlighted yellow and green regions. However, a further increase to $K = 7$ (Fig. 23d) leads to a decline in quality and the reappearance of discontinuities. Although

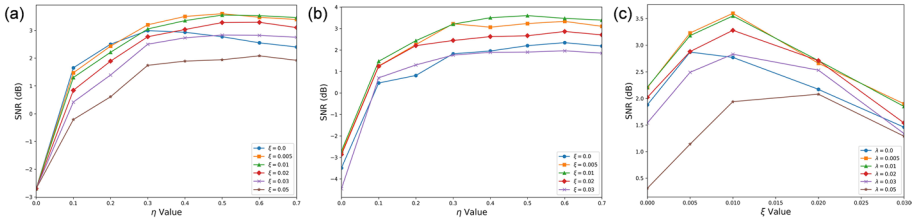

Fig. 22 The impact of different hyperparameters on result accuracy. **a** Results with different hyperparameters when ξ= 0.01, **b** results with different hyperparameters when $\lambda = 0.005$, **c** Results with different hyperparameters when $\eta = 0.5$

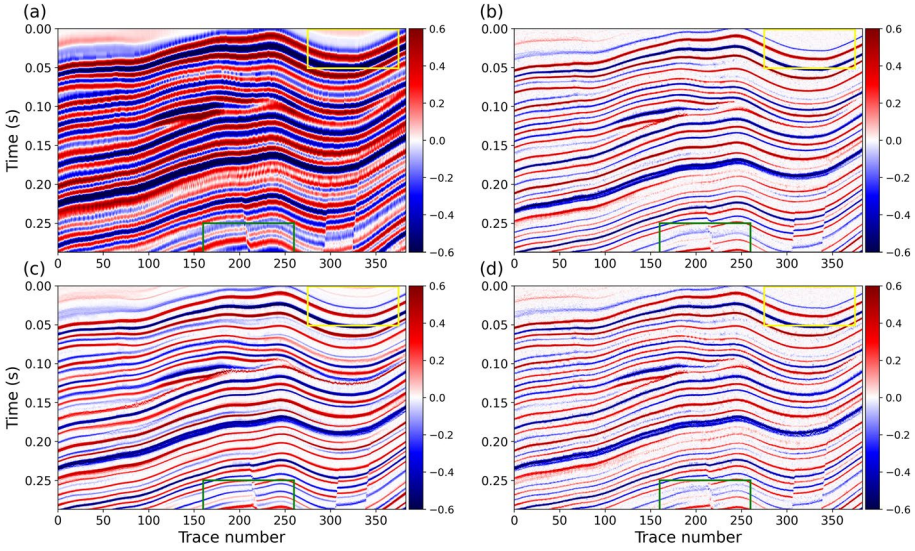

Fig. 23 Crossline results (crossline = 150) of synthetic seismic data. **a** Original synthetic seismic data, **b**, **c**, and **d** present the results from the proposed SeisDPS-3D method, with the hyperparameters K set to 3, 5, and 7, respectively

setting $K = 7$ offers higher computational efficiency, as shown in Table 3, the associated compromise in reconstruction fidelity (i.e., the recurrence of discontinuities) is undesirable. Consequently, to prioritize high-fidelity deconvolution while maintaining reasonable computational cost, we set the hyperparameter K to 5.

Overall, all four parameters have an optimal solution. In practice, they are generally set to $\eta = 0.5$, $\lambda = 0.01$, ξ= 0.01, and $K = 5$ and then fine-tuned during the testing phase. For example, if a sparser result is desired, λ is increased to obtain the optimal results ultimately.

4.2 Computational Efficiency

We evaluated the computational efficiency of the proposed method with the experiment conducted on an NVIDIA GeForce RTX 3090 GPU. A single sampling is performed on a slice of 3D data with dimensions of 256×320, and the FISTA algorithm is executed on

Table 3 The runtime of different hyperparameter K of SeisDPS-3D ($256 \times 256 \times 320$)

K	2	3	4	5	6	7
Inference time (h)	7.12	6.91	6.77	6.58	6.34	6.18

Table 4 Comparison of inference time and memory usage on 2D synthetic data (256×320)

Method	SeisDPS	FISTA (CPU)	FISTA (GPU)
Inference time (s)	81.28	69.64	2.93
GPU memory usage (G)	3.09	NA	0.28

Table 5 Comparison of inference time and memory usage on 3D synthetic data ($256 \times 256 \times 320$)

Method	SeisDPS-3D	SeisDPS	FISTA (CPU)	FISTA (GPU)
Inference time (h)	6.58	5.77	6.18	0.21
GPU memory usage (G)	3.25	3.09	NA	0.28

the CPU and the same GPU. The results are shown in Table 4. Similarly, a single sampling is performed on all 3D data (dimensions $256 \times 256 \times 320$), and the results are shown in Table 5, where SeisDPS refers to using a 2D method to process 3D data (slice by slice). Comparing the memory usage of SeisDPS-3D and SeisDPS, after improving the SeisDPS method, memory usage increased by only 0.16 GB. The increase in inference time is due to the additional unconditional sampling correction. However, because of the inherently high computational cost of diffusion models, the runtime for both methods significantly exceeds that of the FISTA algorithm, whether using the GPU or CPU version of FISTA. Our diffusion-based methods represent a strategic trade-off where superior robustness and reconstruction fidelity are achieved at the expense of increased computational overhead.

4.3 Uncertainty Quantification

To analyze the uncertainty of the proposed method, we performed 1, 000 deconvolution tests on the noisy synthetic data with an SNR of 18 dB, as shown in Fig. 4a, each initialized with a different random seed. The corresponding mean and variance maps of the results are presented in Fig. 24. The mean map represents the finalized deconvolution result. The variance map is on the order of 10^{-3}, indicating a relatively small variance for the normalized data. The areas of high uncertainty primarily occur where the reflection coefficients exhibit geological structure, while in other areas, the uncertainty is almost negligible. In addition, the uncertainty is somewhat higher in areas with strong reflections. This occurs because the relative changes of the reflection coefficients are recovered primarily, which can lead to relatively large fluctuations in the strong reflection areas. Nevertheless, the overall uncertainty remains low, confirming the stability of the deconvolution results.

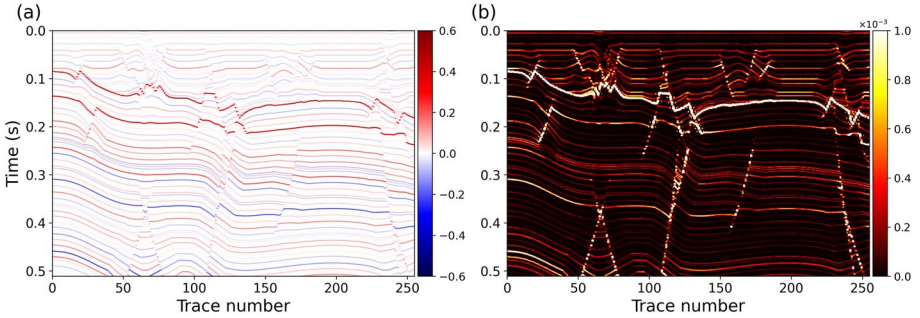

Fig. 24 Uncertainty quantification of the results from 1000 tests: **a** mean map, **b** variance map

4.4 Future Work of Nonstationary Deconvolution

Considering that existing work primarily focuses on the stationary wavelet assumption, and in view of the structure of our subsequent work, we demonstrate the applicability of our method to nonstationary wavelets without introducing a new major research direction. We first generate two nonstationary wavelets. Both nonstationary wavelets linearly attenuate from 40 Hz down to 10 Hz and 25 Hz, respectively. These wavelets are then convolved with synthetic reflection coefficients and add noise to obtain the synthetic nonstationary noisy seismic data, as shown in Fig. 25a and c, respectively. We subsequently estimated the statistical seismic wavelet for each case and performed deconvolution using the SeisDPS method. The resulting deconvolved results are displayed in Fig. 25b and d. It is evident from these figures that when the wavelet attenuation is more severe, the deconvolution results at greater depths are noticeably poorer. Conversely, performance improves as the wavelet attenuation decreases.

The deconvolution problem for nonstationary seismic data can be conceptually decomposed into multiple stationary seismic deconvolution subproblems (Grossman et al. 2002). In each subproblem, the seismic wavelet is assumed to be stationary, allowing our proposed method to be applied, and thus holds potential for extension to nonstationary wavelet deconvolution.

5 Conclusion

This study introduces SeisDPS, a novel method for reconstructing 2D seismic reflection coefficients using diffusion models. SeisDPS performs deconvolution across the entire 2D profile, addressing the limitations of traditional trace-by-trace techniques and significantly enhancing lateral continuity. We extend this approach to three dimensions through SeisDPS-3D, which leverages a pre-trained 2D diffusion model to achieve volumetric deconvolution with only a 5% increase in memory usage, offering a computationally efficient solution for 3D seismic data processing.

The 3D algorithm operates by applying SeisDPS along one direction (e.g., inline profiles) while employing unconditional sampling along the orthogonal direction (e.g., crosslines) and vice versa. This strategy uses the implicit prior encoded by the 2D diffusion model as a 3D constraint, thereby improving lateral continuity in the deconvolved volume. Importantly, the method is trained exclusively on synthetic reflection coefficients, yet it demonstrates strong

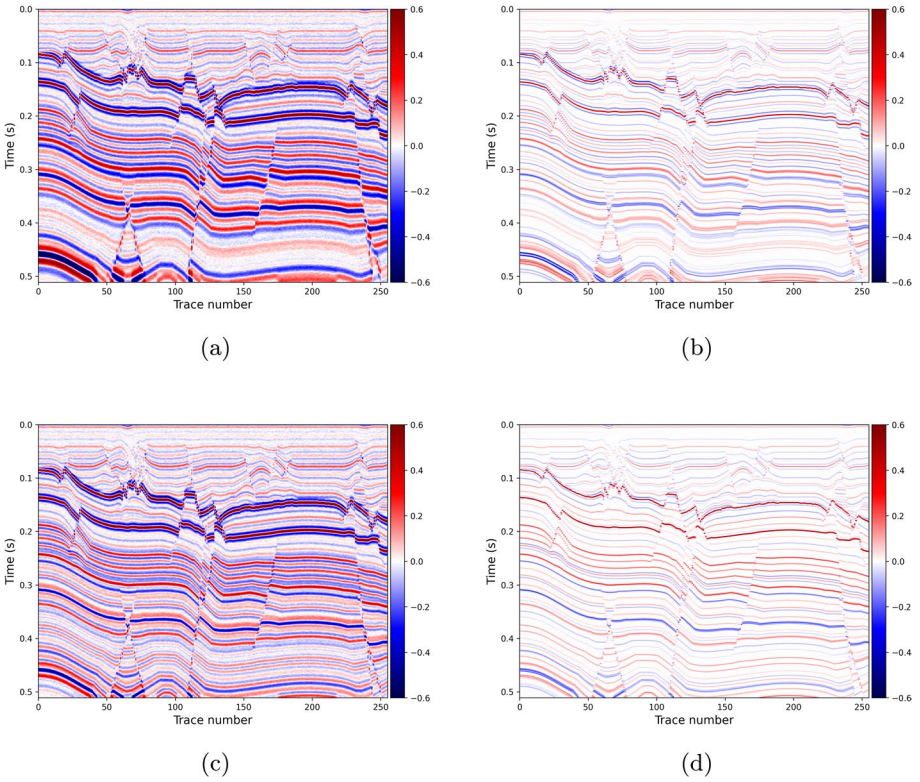

Fig. 25 Deconvolution results on synthetic data with nonstationary wavelets. **a** and **c** show synthetic data of seismic wavelet attenuated from 40 Hz to 10 Hz and 25 Hz, respectively. **b** and **d** show the corresponding deconvolution results for (**a**) and (**c**), respectively

generalization across synthetic and real datasets, including 2D and 3D field data. Fine-tuning of hyperparameters reveals robust performance, with notable improvements in noise resilience and spatial coherence compared to traditional deconvolution techniques.

Our implementation targets stationary wavelet deconvolution; future extensions will address nonstationary scenarios, expanding the method's applicability.

Moreover, with the continued development of generative models and improved learning of priors, the framework proposed in this work can be further developed to achieve superior results.

Finally, although deconvolution is a mature and foundational technique in seismic data processing, accurate reflectivity estimation remains a critical challenge for seismic interpreters and seismic data processing practitioners. Machine learning offers compelling opportunities to incorporate prior information in novel, implicitly defined ways that can enhance seismic resolution. It remains uncertain whether these methodologies will achieve widespread adoption in industrial practice. Nevertheless, we contend that innovative explorations–such as those presented here–are indispensable for propelling seismic data processing into new and promising directions.

Appendix A 2D Noise-Free Data Experiment

To illustrate the efficacy of the proposed 2D method, we performed tests on noise-free data. The SeisDPS method has three primary hyperparameters. This section sets the step size η to 0.3, and the hyperparameters λ and ξ to 0.005. In the FISTA algorithm, the regularization parameter $\mu = 0.05$, with two conditions for stopping the iteration: either the maximum iteration K_{iter} of 500 is reached, or the convergence tolerance falls below 10^{-5}.

See Figs. 26, 27 and 28.

The deconvolution results of synthetic, noise-free seismic data are shown in Fig. 28. Figure 28b demonstrates that the SeisDPS method performs well overall, producing results that closely resemble the true reflection coefficients. Furthermore, zooming in on the green-boxed regions in Fig. 28a–c, as shown in Fig. 28d–f, demonstrates that the SeisDPS method outperforms the FISTA algorithm in capturing finer details, especially as shown in the yellow box. Additionally, the SeisDPS method exhibits better spatial continuity because our method performs deconvolution on 2D profiles, whereas the FISTA algorithm operates trace-by-trace without considering inherent spatial continuity.

Figure 27 illustrates the multi-trace average amplitude spectra for both the true and deconvolved results. Below 125 Hz, the SeisDPS method yields more accurate results than the sparse-spike (FISTA) method. There is only a slight deviation at higher frequencies, similar to what is observed with the FISTA algorithm. Since seismic data typically exhibits relatively low-frequency components, we can conclude that the SeisDPS method effectively recovers the original reflection coefficients. To further validate its performance, we compare the results using data from trace No.180, as shown in Fig. 26. The green box in the figure highlights the excessive amplitude produced by the FISTA algorithm, potentially due to local overfitting when the algorithm balances data fidelity and solution sparsity.

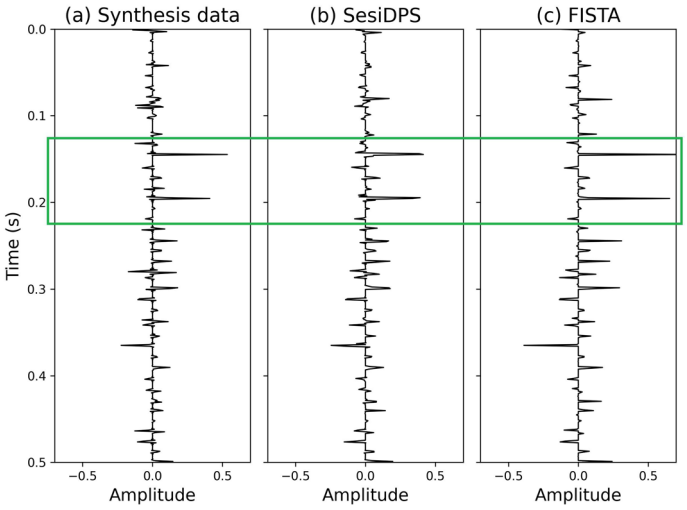

Fig. 26 Trace No.180. **a** True reflection coefficients, **b** the SeisDPS method, **c** the FISTA algorithm. The green box highlights the excessive amplitude generated by FISTA, possibly related to local overestimation arising from the algorithm's trade-off between data fidelity and solution sparsity

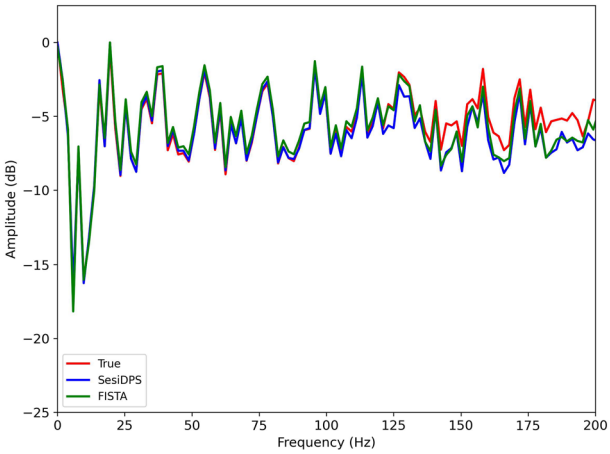

Fig. 27 Multi-trace normalized average amplitude spectrum

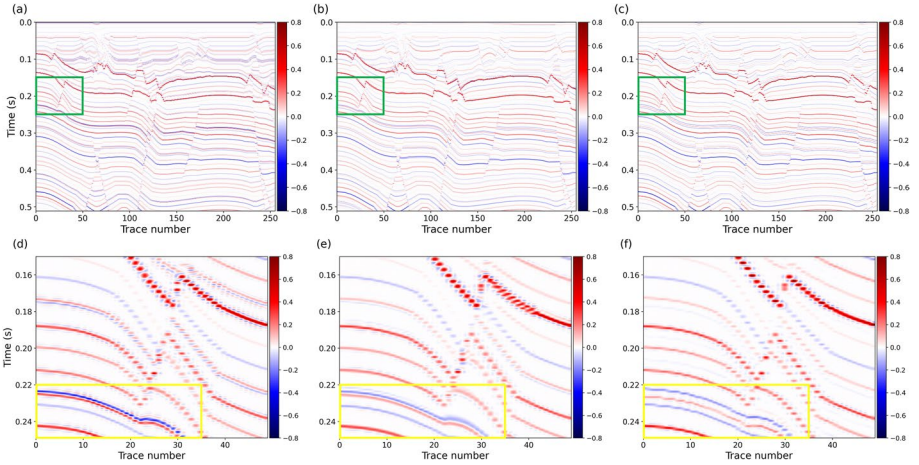

Fig. 28 Deconvolution of 2D synthetic data. **a** True reflection coefficients. **b** The results are obtained using the proposed method (SeisDPS). **c** The results from the FISTA algorithm. **d**–**f** present zoomed-in views of the regions outlined by the green boxes in (**a**–**c**), respectively. As indicated within the yellow boxes in Fig. (**d**–**f**), the thin-layer SeisDPS method in (**e**) closely approximates the synthetic data, outperforming FISTA in this regard

Appendix B Different Wavelets and Frequencies Experiment on 2D Data

Since our method requires an initial estimation of the wavelet, and in previous experiments, only the Ricker wavelet was tested, we now conduct tests on 2D data to further demonstrate our method's robustness with different seismic wavelets. We use the same clean 2D reflection coefficients as in the previous experiments, convolve them with different seismic

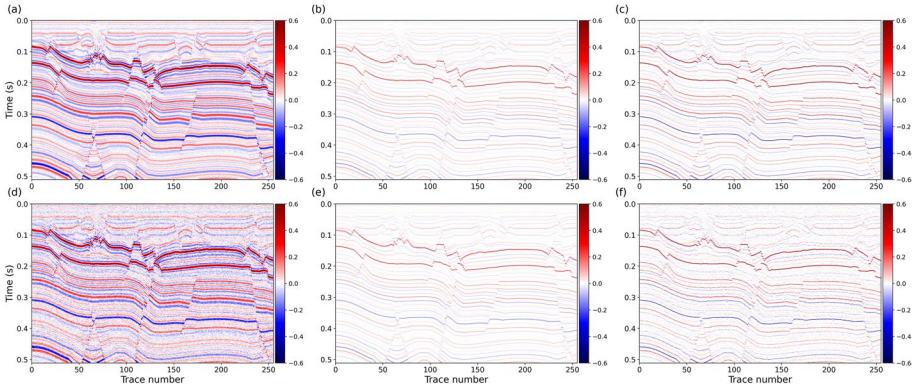

Fig. 29 Deconvolution results under different SNR using convolution with the **Klauder wavelet**. **a** and **d** show the synthetic data at SNRs values of 18 dB and 9 dB, respectively; **b** and **e** present the corresponding SeisDPS results; **c** and **f** show the results obtained using FISTA

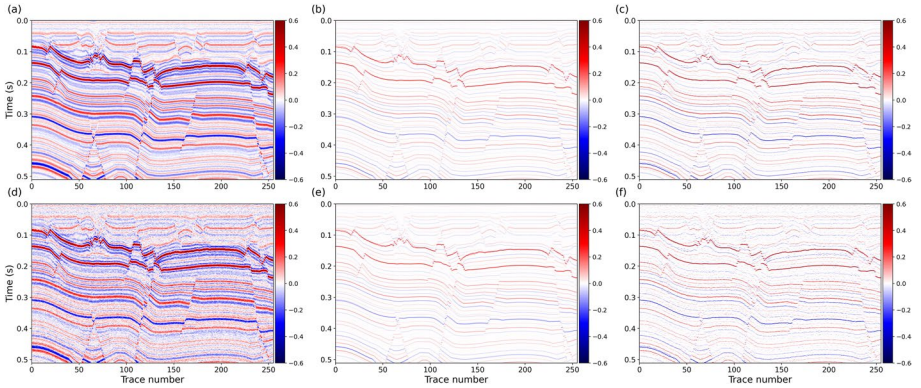

Fig. 30 Deconvolution results under different SNR using convolution with the **Ormsby wavelet**. **a** and **d** show the synthetic data at SNRs values of 18 dB and 9 dB, respectively; **b** and **e** present the corresponding SeisDPS results; **c** and **f** show the results obtained using FISTA

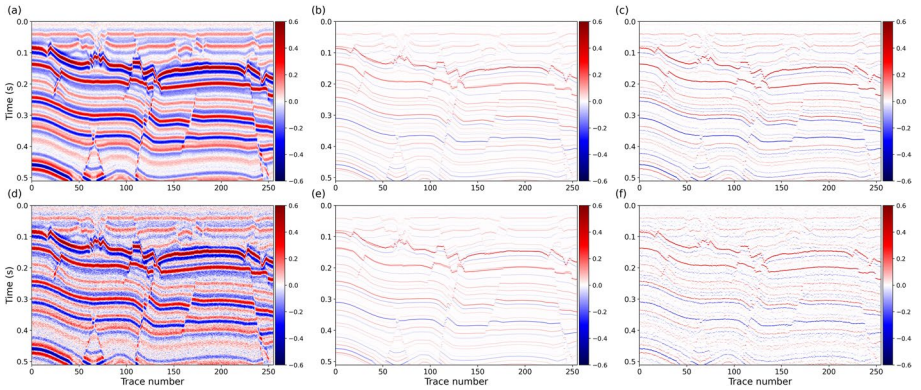

Fig. 31 Deconvolution results under convolution with a **15 Hz Ricker wavelet** for different SNR. **a** and **d** show the synthetic data at SNRs values of 18 dB and 9 dB, respectively; **b** and **e** present the corresponding SeisDPS results; **c** and **f** show the results obtained using FISTA

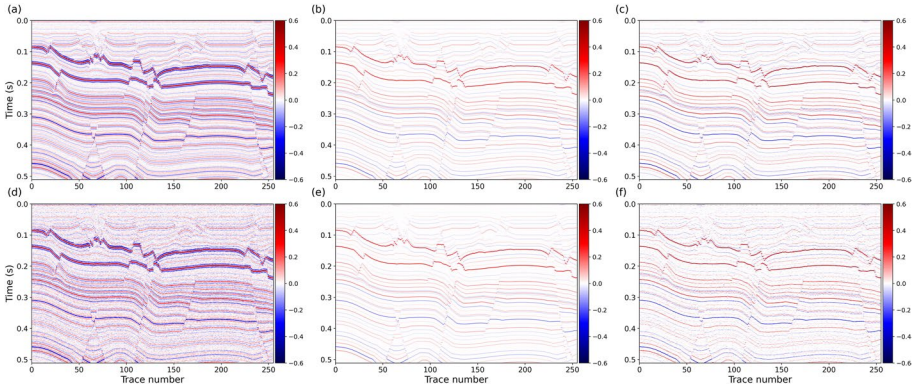

Fig. 32 Deconvolution results under convolution with a **45 Hz Ricker wavelet** for different SNR. **a** and **d** show the synthetic data at SNRs values of 18 dB and 9 dB, respectively; **b** and **e** present the corresponding SeisDPS results; **c** and **f** shows the results obtained using FISTA

wavelets, and add noise with different SNRs to obtain synthetic noisy seismic data. The hyperparameters used remain consistent with the previous treatment of noisy 2D data, ensuring that only the seismic wavelet differs in this test. This experiment uses two different seismic wavelets: the Klauder wavelet and the Ormsby wavelet. See Figs. 29, 30, 31 and 32.

Figure 29 and 30 shows the results obtained from seismic data synthesized with different wavelets at different noise levels. Our method maintains satisfactory spatial continuity and reconstructs relatively accurate reflection coefficients at various noise levels. These results are consistent with those obtained using the Ricker wavelet. The experiment demonstrates that our method is highly robust to various wavelets and noise levels.

To further evaluate the robustness of our method to seismic wavelet frequencies, we convolve reflection coefficients with Ricker wavelets of different frequencies to generate synthetic seismic data. This experiment also incorporates different noise levels, while the hyperparameters remain unchanged to ensure that only the wavelet frequency varies. Figures 31 and 32 display the results using 15 Hz and 45 Hz Ricker wavelets, respectively. The results indicate that wavelet frequency is a key factor influencing the accuracy of the results. As the wavelet frequency increases, the reconstructed reflection coefficients contain more details, which is consistent with the trend observed in the FISTA algorithm. This is because higher frequencies provide better time resolution.

Our method applies to different seismic wavelets and has no specific frequency requirements. It is highly robust to noise at various wavelets and frequencies and can effectively exploit the spatial structure of the data to reconstruct reflection coefficients with better spatial continuity.

Acknowledgements This work was supported in part by the Key Program of the National Natural Science Foundation of China (No. 42530802), in part by the National Key Research and Development Program of China under grant no. 2021YFA0716904, in part by the National Natural Science Foundation of China under grant no. 42374135, in part by the Fundamental Research Funds for the Central Universities under grant no. xzy012023073 and in part by the Innovation Capability Support Program of Shaanxi under the Grant No. 2021TD-08.

Data Availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Alfarraj M, AlRegib G (2019) Semisupervised sequence modeling for elastic impedance inversion. Interpretation 7(3):SE237–SE249

Anderson BD (1982) Reverse-time diffusion equation models. Stoch Process Appl 12(3):313–326

Araya-Polo M, Jennings J, Adler A, Dahlke T (2018) Deep-learning tomography. Lead Edge 37(1):58–66

Backus G, Gilbert F (1968) The resolving power of gross earth data. Geophys J Int 16(2):169–205

Beck A, Teboulle M (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J Imag Sci 2(1):183–202

Berkhout A (1977) Least-squares inverse filtering and wavelet deconvolution. Geophysics 42(7):1369–1383

Burg JP (1972) The relationship between maximum entropy spectra and maximum likelihood spectra. Geophysics 37(2):375–376

Chai X, Tang G, Lin K, Yan Z, Gu H, Peng R et al (2021) Deep learning for multitrace sparse-spike deconvolution. Geophysics 86(3):V207–V218

Chen H, Gao J, Zhang W, Yang P (2021) Seismic acoustic impedance inversion via optimization-inspired semisupervised deep learning. IEEE Trans Geosci Remote Sens 60:1–11

Chen H, Sacchi MD, Haghshenas Lari H, Gao J, Jiang X (2023) Nonstationary seismic reflectivity inversion based on prior-engaged semisupervised deep learning method. Geophysics 88(1):WA115–WA128

Choi J, Kim S, Jeong Y, Gwon Y, Yoon S (2021) Ilvr: Conditioning method for denoising diffusion probabilistic models. arXiv preprint arXiv:2108.02938

Chopra S, Castagna J, Portniaguine O (2006) Seismic resolution and thin-bed reflectivity inversion. CSEG Rec 31(1):19–25

Chung H, Kim J, Mccann MT, Klasky ML, Ye JC (2022) Diffusion posterior sampling for general noisy inverse problems. arXiv preprint arXiv:2209.14687

Chung H, Ryu D, McCann MT, Klasky ML, Ye JC (2023) Solving 3d inverse problems using pre-trained 2d diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition p. 22542–22551

Claerbout JF (1971) Toward a unified theory of reflector mapping. Geophysics 36(3):467–481

Claerbout JF (1976) Fundamentals of geophysical data processing, vol 274. Citeseer, Princeton

Claerbout JF, Muir F (1973) Robust modeling with erratic data. Geophysics 38(5):826–844

Daras G, Chung H, Lai CH, Mitsufuji Y, Ye JC, Milanfar P, et al (2024) A survey on diffusion models for inverse problems. arXiv preprint arXiv:2410.00083

Debeye HWJ, van Riel P (1990) L-Norm Deconvolution. Geophys Prospect 38(4):381–403

Efron B (2011) Tweedie's formula and selection bias. J Am Stat Assoc 106(496):1602–1614

Gao Z, Hu S, Li C, Chen H, Jiang X, Pan Z et al (2021) A deep-learning-based generalized convolutional model for seismic data and its application in seismic deconvolution. IEEE Trans Geosci Remote Sens 60:1–17

Gholami A (2016) A fast automatic multichannel blind seismic inversion for high-resolution impedance recovery. Geophysics 81(5):V357–V364

Gholami A, Sacchi MD (2012) A fast and automatic sparse deconvolution in the presence of outliers. IEEE Trans Geosci Remote Sens 50(10):4105–4116

Gholamy A, Kreinovich V (2014) Why Ricker wavelets are successful in processing seismic data: Towards a theoretical explanation. In: 2014 IEEE Symposium on Computational Intelligence for Engineering Solutions (CIES) IEEE; p. 11–16

Grossman JP, Margrave GF, Lamoureux MP (2002) Adaptive, nonuniform Gabor frames from partitions of unity. CREWES Res Rep 14:1–10

Hamid H, Pidlisecky A (2015) Multitrace impedance inversion with lateral constraints. Geophysics 80(6):M101–M111

Hennenfent G, van den Berg E, Friedlander MP, Herrmann FJ (2008) New insights into one-norm solvers from the Pareto curve. Geophysics 73(4):A23–A26

Ho J, Jain A, Abbeel P (2020) Denoising diffusion probabilistic models. Adv Neural Inf Process Syst 33:6840–6851

Hosken J (1988) Ricker wavelets in their various guises. First Break. 6(1)
Kazemi N, Sacchi MD (2014) Sparse multichannel blind deconvolution. Geophysics 79(5):V143–V152
Kingma DP, Ba J (2014) Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980
Lee S, Chung H, Park M, Park J, Ryu WS, Ye JC (2023) Improving 3D imaging with pre-trained perpendicular 2D diffusion models. In: Proceedings of the IEEE/CVF International Conference on Computer Vision; p. 10710–10720
Li ZX, Li ZC, Lu WK (2016) Multichannel predictive deconvolution based on the fast iterative shrinkage-thresholding algorithm. Geophysics 81(1):V17–V30
Liu Y, Li Y, Peng J, Wang M (2025) Semi-Supervised Learning for AVO Inversion with Strong Spatial Feature Constraints. arXiv preprint arXiv:2501.15473
Merrifield TP, Griffith DP, Zamanian SA, Gesbert S, Sen S, De La Torre GJ et al (2022) Synthetic seismic data for training deep learning networks. Interpretation 10(3):SE31–SE39
Peacock KL, Treitel S (1969) Predictive deconvolution: Theory and practice. Geophysics 34(2):155–169
Porsani M, Ursin B (2007) Direct multichannel predictive deconvolution. Geophysics 72:H11
Rahman S, Elsheikh AH, Jaya MS (2024) A Deep Semi-Supervised Learning Approach for Seismic Reflectivity Inversion. IEEE Transactions on Geoscience and Remote Sensing
Ricker N (1953) The form and laws of propagation of seismic wavelets. Geophysics 18(1):10–40
Robinson EA (1957) Predictive decomposition of seismic traces. Geophysics 22(4):767–778
Sacchi MD (1997) Reweighting strategies in seismic deconvolution. Geophys J Int 129(3):651–656
Sacchi MD, Velis DR, Cominguez AH (1994) Minimum entropy deconvolution with frequency-domain constraints. Geophysics 59(6):938–945
Sohl-Dickstein J, Weiss E, Maheswaranathan N, Ganguli S (2015) Deep unsupervised learning using non-equilibrium thermodynamics. In: International conference on machine learning pmlr. p. 2256–2265
Song Y, Ermon S (2019) Generative modeling by estimating gradients of the data distribution. Advances in neural information processing systems. 32
Song Y, Sohl-Dickstein J, Kingma DP, Kumar A, Ermon S, Poole B (2020) Score-based generative modeling through stochastic differential equations. arXiv preprint arXiv:2011.13456
Taylor HL, Banks SC, McCoy JF (1979) Deconvolution with the ℓ_1 norm. Geophysics 44(1):39–52. https://doi.org/10.1190/1.1440921
Torres K, Sacchi MD (2023) Deep decomposition learning for reflectivity inversion. Geophys Prospect 71(6):963–982
Ulrych TJ (1971) Application of homomorphic deconvolution to seismology. Geophysics 36(4):650–660
Ulrych TJ, Velis DR, Sacchi MD (1995) Wavelet estimation revisited. Lead Edge 14(11):1139–1143
Vincent P (2011) A connection between score matching and denoising autoencoders. Neural Comput 23(7):1661–1674
Wang YQ, Wang Q, Lu WK, Ge Q, Yan XF (2022) Seismic impedance inversion based on cycle-consistent generative adversarial network. Pet Sci 19(1):147–161
Wiggins RA (1978) Minimum entropy deconvolution. Geoexploration 16(1–2):21–35
Wu X, Yan S, Bi Z, Zhang S, Si H (2021) Deep learning for multidimensional seismic impedance inversion. Geophysics 86(5):R735–R745
Wu Y, Wang W, Wang Y, McMechan GA (2025) Deep learning-based self-supervised multiparameter inversion. Geophysical Journal International. 243(2):ggaf332
Yu B, Ozdemir S, Dong Y, Shao W, Pan T, Shi K, et al (2025) Robust whole-body PET image denoising using 3D diffusion models: evaluation across various scanners, tracers, and dose levels. European Journal of Nuclear Medicine and Molecular Imaging. p. 1–14
Yu H, Chen W, Wang X, Liu D (2025) Unsupervised Diffusion Model for Seismic Deconvolution. IEEE Geoscience and Remote Sensing Letters. p. 1–1. https://doi.org/10.1109/LGRS.2025.3549055
Zhang J, Li J, Chen X, Li Y, Huang G, Chen Y (2021) Robust deep learning seismic inversion with a priori initial model constraint. Geophys J Int 225(3):2001–2019
Zhang R, Castagna J (2011) Seismic sparse-layer reflectivity inversion using basis pursuit decomposition. Geophysics 76(6):R147–R158
Zhang R, Sen MK, Srinivasan S (2013) Multi-trace basis pursuit inversion with spatial regularization. J Geophys Eng 10(3):035012

Authors and Affiliations

Dawei Liu[1,2] · Hongzhi Yu[1] · Mauricio D. Sacchi[2]

✉ Dawei Liu
409791715@qq.com

[1] School of Information and Communication Engineering, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, China

[2] Department of Physics, University of Alberta, Edmonton, AB T6G 2E1, Canada