Scientific Research Publishing

Applied Mathematics, 2026, 17(11), 720-743
https://www.scirp.org/journal/am
ISSN Online: 2152-7393
ISSN Print: 2152-7385

Reconstruction of Missing Well-Log DTC Curves Using Low-Rank Matrix Recovery

Haitao Li, Yizhu Fang*, Yanru Chen, Yu Zou, Shu Zheng

Planning Department, PetroChina Southwest Oil and Gas Field Company, Chengdu, China
Email: *fangyizhu_CNPC@outlook.com

How to cite this paper: Li, H.T., Fang, Y.Z., Chen, Y.R., Zou, Y. and Zheng, S. (2026) Reconstruction of Missing Well-Log DTC Curves Using Low-Rank Matrix Recovery. *Applied Mathematics*, **17**, 720-743.
https://doi.org/10.4236/am.2026.1711040

Received: October 11, 2026
Accepted: November 15, 2026
Published: November 18, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/

Open Access

Abstract

To address partial missing values in compressional sonic travel time (DTC) curves, this study develops a missing-log reconstruction method based on low-rank matrix recovery. First, multiple well-log variables, including gamma ray, bulk density, neutron porosity, resistivity, total porosity, water saturation, and DTC, are integrated into a standardized multi-log observation matrix. Considering the correlations among different well-log variables and the vertical continuity of logging responses along depth, a depth-wise smoothness constraint is introduced to improve the geological consistency of the reconstructed DTC curves. Artificial masking experiments are conducted on complete well-log datasets to evaluate the reconstruction performance of the proposed method. Isolated missing points and continuous missing intervals are designed separately to simulate different missing-data patterns in practical logging applications. The proposed method is compared with baseline methods, including linear interpolation and pure low-rank recovery. The experimental results show that the proposed method achieves better reconstruction accuracy for missing DTC curves, while effectively preserving the overall variation trend and depth continuity of the curves. These results indicate that combining low-rank matrix recovery with a depth-continuity prior provides an effective and geologically reasonable solution for missing well-log reconstruction.

Keywords

Well-Log Data, Missing Data Reconstruction, Low-Rank Matrix Recovery, Compressional Sonic Travel Time, DTC, Depth Continuity

1. Introduction

In recent years, the reconstruction of missing well-log data has been investigated

using a variety of data-driven approaches. Traditional interpolation-based methods are still widely used in practical applications because of their simple implementation and low computational cost [1]. However, these methods mainly rely on local continuity along the depth direction and reconstruct missing values only from neighboring samples of the same log. As a result, they are unable to fully exploit the correlations among different logging curves. Their performance may deteriorate when the missing interval is long, when abrupt lithological changes occur, or when the log response is controlled by nonlinear relationships among multiple geological and petrophysical factors. In such cases, interpolation may produce overly smooth curves and fail to recover the true local variation of the missing log.

Recent studies further show that missing well-log imputation remains an active research topic. For example, benchmark comparisons of machine-learning methods indicate that reconstruction performance may vary with the selected model, logging variables, and missing-data characteristics [2]. In addition, collaborative-filtering-based imputation has been introduced to exploit similarity information among logging samples, providing another data-driven perspective for missing-value reconstruction [3]. These studies suggest that missing well-log reconstruction is not only a simple interpolation problem, but also a data-driven estimation problem that depends on both inter-log relationships and the structure of missing patterns.

To overcome these limitations, machine learning methods have been introduced to predict missing logging curves by learning nonlinear relationships from available well-log data. Early studies used artificial neural networks and related machine-learning techniques to estimate petrophysical properties or generate synthetic well logs from existing logging measurements [4]-[7]. More recently, deep-learning-based methods have also been applied to missing well-log prediction and have shown promising performance in real reservoir cases [8] [9]. Compared with interpolation-based methods, machine learning approaches are better able to capture nonlinear mapping relationships among multiple logging curves. However, these methods usually treat samples at different depths as relatively independent observations, and the continuity constraint along the depth direction within the same well is often insufficiently considered. Therefore, in continuous missing intervals or in formations with complex geological variations, the predicted results may exhibit discontinuous fluctuations along the depth direction, which can affect the geological plausibility and morphological consistency of the reconstructed curve. Deep learning models, including convolutional neural networks, recurrent neural networks, spatio-temporal neural networks, self-attention-based architectures, and graph neural networks, have also been applied to well-log prediction and reconstruction tasks [10]-[13]. Although these methods have strong representation capabilities and can model sequential dependencies to some extent, their performance usually depends on large amounts of training data and careful model tuning.

To tackle these drawbacks of existing methods, this study puts forward a miss-

ing well-log reconstruction method that combines low-rank matrix recovery with depth-wise smoothness correction. The proposed method integrates multiple well-log curves into a unified observation matrix and exploits the inherent low-rank structure of multi-log data to estimate the missing parts of the target log. Besides, given that well-log responses in the same well usually change continuously as depth increases, we also add adepth-wise smoothness correction mechanism to make the reconstruction results more vertically consistent and geologically reasonable [14]. In short, the proposed method combines the correlation between different well-log curves and the vertical continuity within a single well, thereby improving the reliability and accuracy of missing well-log reconstruction.

In this study, the DTC curves are selected as the main reconstruction target, while GR (Gamma Ray), RHOB (Bulk Density), NPHI (Neutron Porosity), RT (Resistivity), PHIT (Total Porosity), and water saturation (SW) are used as auxiliary logging variables. Artificial masking experiments are conducted on complete well-log datasets to evaluate the reconstruction performance [15]. The proposed method is assessed using mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), coefficient of determination (R^2), and correlation coefficient (Corr). In addition, comparative experiments against linear interpolation and pure low-rank recovery without smoothness correction are carried out to examine both the overall effectiveness of the low-rank framework and the contribution of the proposed depth smoothness correction. The results demonstrate that the proposed framework can achieve accurate, stable, and geologically plausible reconstruction for incomplete well-log data.

2. Preliminaries

2.1. Multi-Log Data Representation

In well-log interpretation, measurements obtained from multiple logging curves are typically integrated into a unified depth-by-variable matrix to provide support for reconstruction analysis. It is assumed that the dataset comprises n depth samples and p logging variables, where the selected variables include GR, RHOB, NPHI, DTC, RT, PHIT, and SW. These measurements are organized into an observation matrix $\boldsymbol{X} \in \mathbb{R}^{n\times p}$, in which each row corresponds to a depth sample and each column corresponds to a single logging curve [16]. The primary objective of this study is to reconstruct the missing values in the DTC log with the assistance of the auxiliary variables.

Because different logging variables have different physical units and numerical ranges, the data matrix is preprocessed before reconstruction. Specifically, the selected variables are converted to numeric format, the RT variable is logarithmically transformed, and all variables are standardized. After preprocessing, the standardized matrix is denoted by $\boldsymbol{Z}$. To describe missing-data locations, a bi-

nary observation mask matrix $\boldsymbol{\Omega}$ is introduced, where $\Omega_{ij}=1$ indicates that the entry is observed and $\Omega_{ij}=0$ indicates that it is missing.

2.2. Low-Rank Modeling Background

Based on the multi-log matrix representation introduced in Section 2.1, low-rank modeling provides a natural way to characterize the shared structure among different logging variables. In subsurface formations, logging responses are controlled by common geological and petrophysical factors, such as lithology, porosity, fluid content, and compaction. As a result, different logs are often strongly correlated rather than independent. This inter-variable dependency implies that the standardized multi-log matrix Z contains redundant information and can therefore be approximated by a low-rank structure.

Formally, the singular value decomposition (SVD) of $\boldsymbol{Z}$ can be written as

$$\boldsymbol{Z}=\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\top},$$

where $\boldsymbol{U}\in\mathbb{R}^{n\times n}$ and $\boldsymbol{V}\in\mathbb{R}^{p\times p}$ are orthogonal matrices, and $\boldsymbol{\Sigma}\in\mathbb{R}^{n\times p}$ is a diagonal matrix whose diagonal entries are the singular values of $\boldsymbol{Z}$ arranged in non-increasing order.

If the singular values decay rapidly, $\boldsymbol{Z}$ can be well approximated by a rank-r truncated SVD:

$$\boldsymbol{Z}\approx\boldsymbol{Z}_r=\boldsymbol{U}_r\boldsymbol{\Sigma}_r\boldsymbol{V}_r^{\top},$$

where $r\ll\min(n,p)$, $\boldsymbol{U}_r\in\mathbb{R}^{n\times r}$, $\boldsymbol{\Sigma}_r\in\mathbb{R}^{r\times r}$, and $\boldsymbol{V}_r\in\mathbb{R}^{p\times r}$. According to the Eckart-Young theorem, $\boldsymbol{Z}_r$ is the best rank-r approximation to $\boldsymbol{Z}$ under the Frobenius norm.

For the purpose of subsequent optimization, the truncated SVD can be equivalently expressed in a bilinear factorization form:

$$Z_r=AB^{\top},$$

where

$$A=U_r\Sigma_r^{1/2},$$
$$B=V_r\Sigma_r^{1/2}.$$

Here, $A\in\mathbb{R}^{n\times r}$ and $B\in\mathbb{R}^{p\times r}$. Under this representation, the i-th row of A characterizes the latent features of the i-th depth sample, whereas the j-th row of B describes how the j-th logging variable is associated with the latent factors.

This factorized form provides an effective way to represent the shared correlation structure among different well-log curves. Such a low-rank property is particularly advantageous for recovering missing entries in the data matrix, because it allows the reconstruction model to exploit the underlying dependencies among variables rather than treating each log independently [17] [18]. Low-rank matrix recovery has also been applied to oil-well logging data processing, showing its potential for extracting coherent logging responses and suppressing irregular per-

turbations in logging curves [19].

2.3. Depth Continuity Prior

In addition to the mutual correlation among different logging curves, a single well-log curve usually exhibits continuity along the depth direction within the same borehole. Under normal geological conditions, the measured response changes gradually with depth rather than fluctuating abruptly. Sharp variations generally occur near formation boundaries or within intervals with obvious lithological changes. Therefore, depthwise continuity can be regarded as an important geological prior for missing-log reconstruction [20] [21]. Previous geophysical studies have also shown that introducing smoothness constraints based on well-log priors can improve both the numerical stability and the geological plausibility of inversion results [22]. Related studies on multiple well-seismic ties further indicate that depth-consistent alignment and joint use of well information are important for improving the consistency of subsurface interpretation [23].

Although low-rank modeling can capture the global correlations among different logging variables, relying only on inter-log correlation may still lead to unreasonable local fluctuations in the reconstructed curve, especially within continuous missing intervals. In other words, global correlation information alone is not sufficient to guarantee that the reconstructed DTC curve is smooth and geologically consistent along depth. It is therefore necessary to further introduce a depth-continuity constraint to preserve the local morphology of the logging curve.

For a depth-ordered logging sequence, depth continuity can be expressed by penalizing the first-order differences between adjacent depth samples. This penalty discourages abrupt changes between neighboring reconstructed values and encourages the recovered curve to vary smoothly along depth. Let $\boldsymbol{v}$ denote the current reconstructed estimate of a missing logging sequence, and let $\boldsymbol{x}$ denote the corrected sequence after imposing the depth-continuity constraint. The correction can be formulated as an optimization problem in which $\boldsymbol{x}$ is required to remain close to $\boldsymbol{v}$ while satisfying a smoothness requirement.

This correction can be interpreted using the proximal-operator framework. For a proper convex function $g(\boldsymbol{x})$, the proximal operator of g at $\boldsymbol{v}$ is defined as

$$\operatorname{prox}_g(\boldsymbol{v}) = \arg\min_{\boldsymbol{x}} \left\{ \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{v}\|_2^2 + g(\boldsymbol{x}) \right\}.$$

In this definition, $\boldsymbol{v}$ is the input vector before correction, $\boldsymbol{x}$ is the optimized output vector, and $g(\boldsymbol{x})$ represents the structural constraint imposed on the solution. In the present study, $g(\boldsymbol{x})$ is associated with the depthwise smoothness of the reconstructed log. The first term keeps the corrected sequence close to the original reconstruction, whereas the second term enforces the desired smoothness along depth. Thus, the proximal formulation provides a convenient optimization form for the subsequent depth-continuity correction.

3. Methodology

3.1. Low-Rank Matrix Recovery Model

Using the standardized multi-log matrix $\boldsymbol{Z}$ and the observation mask $\boldsymbol{\Omega}$ defined in Section 2.1, the missing DTC reconstruction problem can be formulated as a matrix recovery task. Since different logging variables are controlled by a limited number of common geological and petrophysical factors, the complete multi-log matrix is expected to have an approximately low-rank structure. A natural formulation is therefore to seek a recovered matrix $\boldsymbol{M}$ with the lowest possible rank while keeping the observed entries consistent with the available measurements:

$$\min_{\boldsymbol{M}} \frac{1}{2}\left\|\boldsymbol{\Omega}\odot\left(\boldsymbol{Z}-\boldsymbol{M}\right)\right\|_F^2+\lambda\,\mathrm{rank}\left(\boldsymbol{M}\right),$$

where $\odot$ denotes the Hadamard product, $\left\|\cdot\right\|_F$ is the Frobenius norm, and $\lambda>0$ controls the strength of the low-rank constraint.

However, directly minimizing the rank function is computationally difficult because $\mathrm{rank}\left(\boldsymbol{M}\right)$ is non-convex and discontinuous. In matrix recovery problems, a common strategy is to replace the rank function with the nuclear norm, which is defined as the sum of the singular values of the matrix:

$$\left\|\boldsymbol{M}\right\|_*=\sum_k \sigma_k\left(\boldsymbol{M}\right),$$

where $\sigma_k\left(\boldsymbol{M}\right)$ denotes the k-th singular value of $\boldsymbol{M}$. The nuclear norm is widely used as a convex surrogate for the rank function and encourages the recovered matrix to have a low-rank structure. Singular-value-thresholding-based methods provide a representative computational framework for nuclear-norm-regularized matrix completion [24]. This leads to the following nuclear-norm regularized recovery model:

$$\min_{\boldsymbol{M}}\left\|\frac{1}{2}\boldsymbol{\Omega}\odot\left(\boldsymbol{Z}-\boldsymbol{M}\right)\right\|_F^2+\lambda\left\|\boldsymbol{M}\right\|_*.$$

For computational implementation, the low-rank matrix $\boldsymbol{M}$ is further represented by the product of two low-dimensional factor matrices:

$$\boldsymbol{M}=\boldsymbol{A}\boldsymbol{B}^{\top},$$

where $\boldsymbol{A}\in\mathbb{R}^{n\times r}$ and $\boldsymbol{B}\in\mathbb{R}^{p\times r}$ are latent factor matrices, and $r\ll\min\left(n,p\right)$ is the prescribed rank. Under this factorized representation, the recovery model used in this study is written as

$$\min_{\boldsymbol{A},\boldsymbol{B}}\left\|\boldsymbol{\Omega}\odot\left(\boldsymbol{Z}-\boldsymbol{A}\boldsymbol{B}^{\top}\right)\right\|_F^2+\lambda\left(\left\|\boldsymbol{A}\right\|_F^2+\left\|\boldsymbol{B}\right\|_F^2\right). \tag{1}$$

In Equation (1), the reconstruction error is evaluated only on the observed entries, so that the available measurements are preserved during model fitting. The regularization term constrains the magnitude of the latent factors and improves the numerical stability of the solution [25]. Compared with directly estimating all entries of $\boldsymbol{M}$, which involves np unknowns, the factorized form only requires

$r(n+p)$ unknowns. When r is much smaller than $\min(n,p)$, this representation substantially reduces the number of parameters and makes the recovery problem more tractable.

After optimization, the reconstructed standardized matrix is obtained as

$$\boldsymbol{M} = \boldsymbol{A}\boldsymbol{B}^{\top}.$$

The entries of $\boldsymbol{M}$ corresponding to the missing positions in the DTC column are then used as the preliminary reconstruction results. These values are subsequently refined by the depth smoothness correction introduced in the next subsection.

3.2. Depth Smoothness Correction

The low-rank recovery model effectively captures the global correlations among logging variables, but it does not explicitly impose vertical continuity along the depth direction. In real subsurface settings, most well-log responses vary gradually with depth, except near formation boundaries or intervals with sharp lithological changes [26]. Therefore, when low-rank reconstruction is used alone, the recovered curves may still exhibit unrealistic local oscillations, especially within masked intervals. To incorporate this geological prior into the reconstruction process, a depth-wise smoothness correction is introduced.

For each well w, suppose that the reconstructed target log contains n_w depth samples ordered by depth. Let $\boldsymbol{x}_w \in \mathbb{R}^{n_w}$ denote the corrected target-log segment for well w, and let $\boldsymbol{m}_w \in \mathbb{R}^{n_w}$ denote the current reconstructed segment extracted from the low-rank estimate $\boldsymbol{M} = \boldsymbol{A}\boldsymbol{B}^{\top}$. Define the first-order difference operator

$$\boldsymbol{D}_w = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \in \mathbb{R}^{(n_w-1)\times n_w}.$$

Then, $\boldsymbol{D}_w\boldsymbol{x}_w$ represents the first-order differences between adjacent depth samples. Therefore, the depth-wise smoothness of the reconstructed DTC sequence can be quantified by the sum of squared adjacent differences:

$$\left\|\boldsymbol{D}_w\boldsymbol{x}_w\right\|_2^2 = \sum_{i=1}^{n_w-1}\left(x_{w,i+1} - x_{w,i}\right)^2.$$

In the alternating optimization procedure, the following expression is treated as the smoothing subproblem within the overall objective. When updating the smoothing variable x_w, the low-rank factors $\boldsymbol{A}$ and $\boldsymbol{B}$, and hence the current estimate m_w, are temporarily held fixed. Thus, x_w is the only unknown optimized in this step; the smoothing subproblem is not independent of the overall objective.

To incorporate depthwise continuity while preserving consistency with the current low-rank reconstruction, a well-wise quadratic optimization subproblem is formulated as follows:

$$x_w = \arg\min_x \left\{ \frac{1}{2}\|x - m_w\|_2^2 + \frac{\alpha}{2}\|D_w x\|_2^2 \right\}. \quad (2)$$

This problem can be interpreted as a proximal step associated with the smoothness regularizer. Specifically,

$$x_w = \text{prox}_{\frac{\alpha}{2}\|D_w \cdot\|_2^2}(m_w),$$

where prox denotes the proximal operator. Since Equation (2) is a strictly convex quadratic problem, its solution is given in closed form by

$$x_w = \left(I + \alpha D_w^\top D_w\right)^{-1} m_w.$$

Although the complete reconstruction process can be conceptually associated with the following regularized objective, the optimization in this study is implemented in a decoupled manner. Specifically, low-rank recovery is first performed through ALS by alternately updating $\boldsymbol{A}$ and $\boldsymbol{B}$ using the observed entries. With the resulting low-rank estimate held fixed, the depth-wise smoothness term is then solved as the well-wise proximal subproblem above. The smoothed DTC estimates are used to form the final reconstruction and are not fed back into a subsequent ALS factor update. Accordingly, the overall reconstruction model can be written as

$$\min_{A,B,\{x_w\}} \left\| \Omega \odot \left(Z - AB^\top \right) \right\|_F^2 + \lambda \left(\|A\|_F^2 + \|B\|_F^2 \right) + \alpha \sum_w \|D_w x_w\|_2^2. \quad (3)$$

In practice, Equation (3) is not solved as a fully coupled joint update. The implementation first completes the ALS low-rank fitting using only the observed entries, constructs the preliminary estimate, and then solves the smoothing subproblem with the low-rank factors fixed. The resulting smoothed values are not reintroduced into the ALS factor updates. Therefore, the depth-wise correction is a subproblem-level realization of the smoothness regularization term rather than an additional ALS update.

3.3. Optimization Procedure

The reconstruction model presented here is solved via an alternating least squares (ALS) procedure, a choice motivated by the bilinear structure of the low-rank factorization problem in $\boldsymbol{A}$ and $\boldsymbol{B}$: a joint optimization of both factor matrices proves challenging, whereas fixing one of them reduces the subproblem to a regularized least-squares formulation that admits a closed-form solution [27] [28]. The overall optimization therefore proceeds by alternately updating $\boldsymbol{A}$ and $\boldsymbol{B}$.

Let $\boldsymbol{Z}$ stand for the standardized complete data matrix, and let $\boldsymbol{\Omega}$ represent the observation mask matrix. An entry Ω_{ij} set to 1 indicates that the corresponding value in $\boldsymbol{Z}$ is treated as observed throughout the optimization. An entry set to 0 signals that the value is regarded as missing and therefore becomes a target for reconstruction. In the masking experiment, the true values at the masked positions do remain present within $\boldsymbol{Z}$. Through the action of the mask $\boldsymbol{\Omega}$, how-

ever, they are kept apart from the ALS fitting process and play no role in guiding the updates.

The optimization is initialized by assigning random values to the factor matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. The two factor matrices are then updated iteratively using the ALS scheme, which is commonly adopted in low-rank matrix completion problems [29]. In each iteration, one factor matrix is fixed while the other is updated by solving a set of regularized least-squares subproblems. The detailed update procedure is described as follows.

Step 1. Update of $\boldsymbol{A}$

When B is fixed, the update of A can be performed row by row. For the i-th sample, only the observed entries indicated by the i-th row of Ω are used. Let

$$\mathcal{O}_i = \{ j \mid \Omega_{ij} = 1 \}.$$

Accordingly, the corresponding submatrix of B is denoted by B_i, and the observed values in the i-th row of Z are denoted by Z_i. The update of the i-th row of A, denoted by A_i, is obtained by solving the following regularized least-squares problem:

$$\min_{A_i} \frac{1}{2} \left\| Z_i - A_i B_i^{\top} \right\|_2^2 + \frac{\lambda}{2} \left\| A_i \right\|_2^2.$$

Taking the derivative with respect to A_i yields

$$\frac{\partial}{\partial A_i} \left(\frac{1}{2} \left\| Z_i - A_i B_i^{\top} \right\|_2^2 + \frac{\lambda}{2} \left\| A_i \right\|_2^2 \right) = -\left(Z_i - A_i B_i^{\top} \right) B_i + \lambda A_i.$$

Setting the derivative to zero gives the corresponding normal equation. Therefore, the closed-form update of A_i is

$$A_i = Z_i B_i \left(B_i^{\top} B_i + \lambda I \right)^{-1}.$$

This step estimates the latent representation of the i-th sample based only on its available observations.

Step 2. Update of $\boldsymbol{B}$

Similarly, when A is fixed, the update of B can be performed row by row. For the j-th logging variable, only the observed entries indicated by the j-th column of Ω are used. Let

$$\mathcal{O}_j = \{ i \mid \Omega_{ij} = 1 \}.$$

Then, the corresponding submatrix of A is denoted by A_j, and the observed values in the j-th column of Z are denoted by Z_j. The update of the j-th row of B, denoted by B_j, is obtained by solving the following regularized least-squares problem:

$$\min_{B_j} \frac{1}{2} \left\| Z_j - A_j B_j^{\top} \right\|_2^2 + \frac{\lambda}{2} \left\| B_j \right\|_2^2.$$

Taking the derivative with respect to B_j yields

$$\frac{\partial}{\partial B_j}\left(\frac{1}{2}\left\|Z_j - A_j B_j^{\top}\right\|_2^2 + \frac{\lambda}{2}\left\|B_j\right\|_2^2\right) = -\left(Z_j - A_j B_j^{\top}\right)^{\top} A_j + \lambda B_j.$$

Setting the derivative to zero gives the corresponding normal equation. Therefore, the closed-form update of B_j is

$$B_j = Z_j^{\top} A_j \left(A_j^{\top} A_j + \lambda I\right)^{-1}.$$

This step estimates the latent representation of the j-th logging variable under the current sample-factor matrix.

Step 3. Construction of the Low-Rank Estimate

After both factor matrices have been updated, the low-rank estimate of the standardized data matrix is obtained as

$$\boldsymbol{M} = \boldsymbol{A}\boldsymbol{B}^{\top}.$$

The matrix $\boldsymbol{M}$ provides model-based estimates for all entries of the standardized data matrix. At this stage, however, it is only regarded as an intermediate estimate. The final reconstruction is obtained after the subsequent depth-wise smoothing and data-consistency enforcement steps.

Step 4. Depth Smoothness Correction

For each well, the DTC sequence extracted from the preliminary estimate $\boldsymbol{M}^{(0)}$ is refined using the depth-wise smoothing strategy. This step reduces unrealistic local fluctuations and improves the continuity of the reconstructed DTC curve along depth. The resulting smoothed DTC values are used for final data completion only; they are not used to update $\boldsymbol{A}$ or $\boldsymbol{B}$ in a later ALS iteration. Let $\tilde{\boldsymbol{M}}$ denote the matrix after this smoothing correction.

Step 5. Restoration of Observed Entries

After the depth-wise smoothing step, data consistency is enforced by combining the original observed entries with the smoothed reconstructed values at the masked positions. Let $\tilde{\boldsymbol{M}}$ denote the matrix after low-rank reconstruction and depth-wise smoothing, and let $\hat{\boldsymbol{M}}$ denote the final standardized reconstruction. Then,

$$\hat{M}_{ij} = \begin{cases} Z_{ij}, & \Omega_{ij} = 1, \\ \tilde{M}_{ij}, & \Omega_{ij} = 0. \end{cases}$$

This step does not indicate that the observed entries are reconstructed. Rather, it ensures that the known measurements remain fixed in the final output, whereas only the masked entries are filled by the reconstruction model.

Step 6. Iteration and Inverse Transformation

The ALS factor updates are repeated until the preset maximum number of iterations is reached. After the final low-rank estimate has been obtained, the depth-wise smoothing and data-consistency steps are applied once to produce the final standardized reconstruction $\hat{\boldsymbol{M}}$. The final standardized reconstruction is then converted back to the original measurement scale using the inverse standardization parameters obtained during preprocessing. For the RT variable, an additional inverse transformation is required because a logarithmic transformation was ap-

plied before standardization. Therefore, the corresponding values are exponentiated after inverse standardization to recover the original resistivity scale. Finally, the reconstructed DTC values are extracted from the DTC column for subsequent evaluation and visualization.

3.4. Algorithm Framework

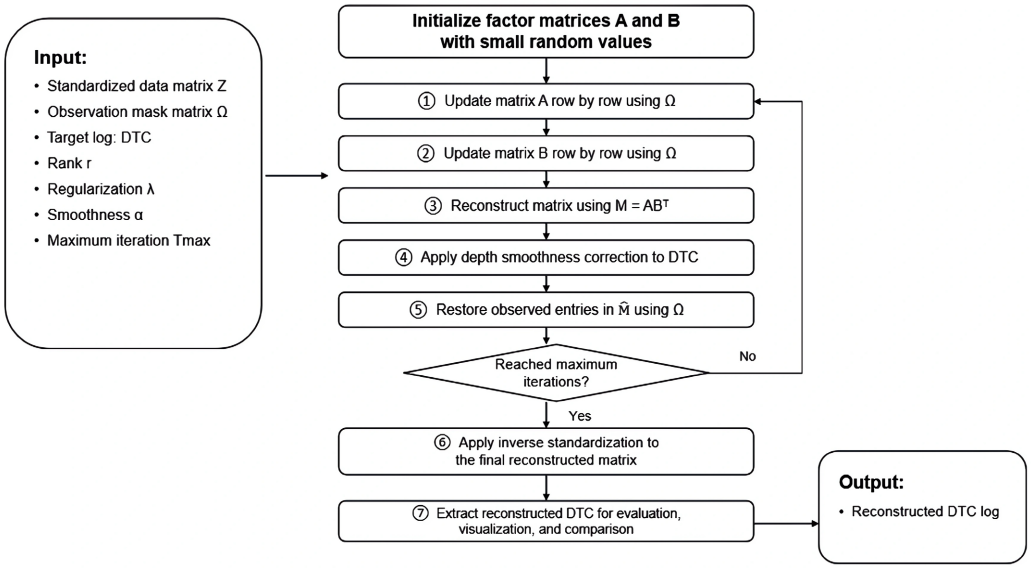


Figure 1. Framework of matrix recovery algorithm.

Algorithm 1 DTC reconstruction framework based on low-rank matrix factorization and depth-wise smoothness correction

Input: standardized multi-log matrix $\boldsymbol{Z}$; observation mask matrix $\boldsymbol{\Omega}$; well index information and depth order; rank r; regularization coefficient λ; smoothness coefficient α; maximum number of iterations T

Output: reconstructed DTC log in the original measurement scale

1: Initialize the factor matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ with small random values

2: **for** $t = 1$ to T do

3: Update the sample-factor matrix $\boldsymbol{A}$ row by row using only the observed entries indicated by $\boldsymbol{\Omega}$

4: Update the variable-factor matrix $\boldsymbol{B}$ row by row using only the observed entries indicated by $\boldsymbol{\Omega}$

5: Construct the preliminary low-rank estimate: $\boldsymbol{M} = \boldsymbol{A}\boldsymbol{B}^{\top}$

6: The preliminary estimate $\boldsymbol{M}$ is used only for the ALS iteration; smoothing is not fed back into the factor updates

7: **end**

8: After the ALS iterations, extract the DTC sequence of each well from $\boldsymbol{M}$ and apply depth-wise smoothness correction based on the first-order difference operator, yielding $\tilde{\boldsymbol{M}}$

9: Enforce data consistency by restoring the observed entries:

$$\hat{\boldsymbol{M}}_{ij} = \begin{cases} \boldsymbol{Z}_{ij}, & \Omega_{ij} = 1, \\ \tilde{\boldsymbol{M}}_{ij}, & \Omega_{ij} = 0. \end{cases}$$

10: Apply inverse standardization to $\hat{\boldsymbol{M}}$ to obtain the reconstructed matrix $\hat{\boldsymbol{X}}$ in the original measurement scale

11: Extract the reconstructed DTC values from $\hat{\boldsymbol{X}}$ for quantitative evaluation, visualization, and comparison with baseline methods

To clearly describe the implementation procedure of the proposed reconstruction method, the main computational steps are summarized in Algorithm 1. Algorithm 1 presents the iterative low-rank factor update, depth-wise smoothness correction, data-consistency restoration, and inverse transformation steps in a pseudo-code form. In addition, Figure 1 provides a graphical overview of the reconstruction workflow, illustrating the relationship among the input data, iterative matrix recovery process, and final reconstructed DTC output.

For a more intuitive understanding of the proposed framework, the graphical workflow of the reconstruction procedure is shown in Figure 1.

4. Experimental Setup

4.1. Dataset and Preprocessing

The dataset used in this study was supplied as a complete multi-log spreadsheet containing ten wells labeled W01 - W10. Each well contains 600 samples over a depth range of 2500.0 - 2799.5 m, with a uniform depth-grid spacing of 0.5 m, yielding 6000 samples in total. Records were included in the analysis on the basis of complete records for DTC and all six auxiliary variables. Before artificial masking, all selected wells therefore contain complete DTC curves, and the original DTC values serve as reference values in the artificial-masking experiments. During reconstruction, data from all wells are organized into a unified multi-log matrix for joint multi-well reconstruction. The main information of the dataset is summarized in Table 1.

Table 1. Summary of the dataset used in this study.

Item	Value
Number of wells	10
Total number of valid samples	6000
Samples per well	600
Depth range	2500.0 - 2799.5 m
Synthetic depth-grid spacing	red0.5 m
Target log	DTC
Auxiliary logs	GR, RHOB, NPHI, RT, PHIT, SW
Reconstruction mode	Joint multi-well reconstruction
Evaluation mode	Well-wise masked-sample evaluation

PHIT and SW were imported directly from the supplied dataset rather than calculated within the reconstruction code. However, the supplied materials do not document the original data source, geological setting, initial well-selection procedure, or the methods used to obtain PHIT and SW. Consequently, whether these auxiliary variables were originally derived from DTC cannot be established. If either variable depends directly or indirectly on DTC, retaining its original values

at masked DTC locations may introduce target-information leakage. The present results should therefore be interpreted as reconstruction performance conditional on the supplied auxiliary variables, rather than as evidence of performance using independently obtained inputs.

Before reconstruction, preprocessing is performed to improve the numerical stability of the optimization. Since the RT curve usually exhibits a highly skewed distribution and may contain values with large magnitudes, a logarithmic transformation is first applied to RT:

$$X_{i,\mathrm{RT}}^{\log} = \log\left(X_{i,\mathrm{RT}}\right),$$

where $X_{i,\mathrm{RT}}$ denotes the original RT value at the i-th depth sample, and $X_{i,\mathrm{RT}}^{\log}$ denotes the corresponding log-transformed value. This transformation reduces the influence of extreme resistivity values.

For each masking experiment, the masking pattern is generated before standardization. Let $\Omega_{ij} = 1$ indicate that entry (i, j) is available to the model and $\Omega_{ij} = 0$ indicate that it is masked. The mean and standard deviation for each logging variable are calculated using only its unmasked observations. Let $\boldsymbol{X}$ denote the preprocessed data matrix before standardization, and let $\boldsymbol{Z}$ denote the standardized matrix. The standardization is given by

$$Z_{ij} = \frac{X_{ij} - \mu_j^{(\Omega)}}{\sigma_j^{(\Omega)}},$$

where $\mu_j^{(\Omega)}$ and $\sigma_j^{(\Omega)}$ are computed from the entries for which $\Omega_{ij} = 1$ in variable j. In particular, the true DTC values at masked positions are excluded when calculating the DTC normalization parameters. The same observed-only parameters are then used to standardize the data for model fitting and to transform predictions back to the original scale. This transformation places different logging variables on comparable numerical scales and prevents variables with larger magnitudes from dominating the estimation of the low-rank factors.

4.2. Experimental Design and Baseline Methods

To evaluate the reconstruction performance under controlled conditions, artificial masking is applied to the complete DTC curves. The mask is created before calculating the standardization parameters. DTC means and standard deviations are calculated using only unmasked DTC observations; the original values at masked positions are retained solely for evaluation and are excluded both from normalization-statistic estimation and from model fitting through the observation mask matrix $\boldsymbol{\Omega}$. Auxiliary variables remain available as specified by the experimental design.

Masking strategy. The masking operation is performed only on the target DTC log, while the auxiliary logs remain complete. To better approximate practical missing-data patterns, both isolated missing points and continuous missing intervals are included. The total masking ratio is 7.63% of all DTC samples. Of the 458 masked DTC samples, 295 are isolated points (64.41%) and 163 belong to contin-

uous intervals (35.59%). For interval-length analysis, adjacent masked positions are grouped into a single continuous interval. This grouping yields 11 short intervals of 2 - 10 samples and four longer intervals of 35 - 50 samples, as shown in Figure 7(b). Thus, the 2 - 10-sample group consists of adjacent masked positions that form short runs, whereas the four deliberately generated long intervals form the 35 - 50-sample group. All evaluation metrics are calculated only at the masked DTC positions, so that the results directly reflect the model's ability to recover missing values.

Baseline methods. Two baseline methods are used for comparison. The first is linear interpolation, which reconstructs missing values within each well based only on local depth-wise continuity. The second is pure low-rank recovery, which uses the same matrix factorization framework as the proposed method but does not include the depth-wise smoothness correction. These two baselines are used to examine the effect of multi-log low-rank modeling and the additional contribution of depth-wise smoothness correction, respectively.

Parameter settings. The main parameters used in the reconstruction experiments are listed in Table 2. The rank of the low-rank factorization was set to $r = 3$, which determines the dimension of the latent factor space. The regularization coefficient was fixed at $\lambda = 0.05$ to constrain the magnitude of the factor matrices and improve numerical stability. The depth-wise smoothing step size was set to $\alpha = 0.05$ to control the strength of the smoothness correction. The maximum number of ALS iterations was set to $T_{\max} = 5000$. To ensure reproducibility, the random seed for initializing the factor matrices was fixed at 0.

The choice of $r = 3$ provides a compact latent representation of the seven-variable multi-log matrix while retaining the dominant cross-log correlation structure; larger ranks introduce additional degrees of freedom and may reduce the regularizing effect of the low-rank model. The value $\lambda = 0.05$ was selected to provide moderate shrinkage of the factor matrices and to prevent numerically unstable factor magnitudes. The value $\alpha = 0.05$ was selected to suppress implausible point-to-point oscillations while preserving the main depth-wise variation of the DTC curve. These parameter values were fixed before the final evaluation and were kept constant across all wells and masking patterns. They were not tuned using the true DTC values at the masked positions; those values were used only for the final evaluation.

Table 2. Parameter settings used in the reconstruction experiments.

Parameter	Symbol	Value
Rank of low-rank factorization	r	3
Regularization coefficient	λ	0.05
Smoothness step size	α	0.05
Maximum number of ALS iterations	$T_{\max}$	5000
Random seed	–	0

5. Results and Discussion

This section evaluates the proposed method from both quantitative and visual perspectives. The analysis includes overall reconstruction metrics, single-well curve comparisons, reconstructed-versus-true scatter plots, error distributions, well-wise error statistics, and comparisons with baseline methods. These results jointly assess the accuracy, stability, and geological plausibility of the reconstructed DTC curves.

5.1. Overall Quantitative Performance

The reconstruction performance of the proposed method is evaluated first on the masked positions of the DTC log and because the observed entries are restored to their original values during reconstruction all evaluation metrics are computed exclusively on the masked samples. The reported results thus reflect the method's genuine missing-data recovery capability without being influenced by the observed portion of the data that remains unchanged.

The quantitative evaluation of DTC reconstruction yielded the following results. The proposed method achieved an MAE of 1.810, an RMSE of 2.263, and a MAPE of 2.429%. The R^2 value reached 0.848, and the correlation coefficient stood at 0.923. Taken together, these figures suggest that the reconstructed DTC values lie reasonably close to the true measurements. The MAE and RMSE remain modest, pointing to limited average errors and to the fact that larger deviations are kept in check. The low MAPE further implies that the relative reconstruction error, when viewed against the actual DTC magnitudes, is fairly small. The R^2 and correlation values, both relatively high.

The results indicate that the proposed approach combining low-rank reconstruction with depth smoothness correction is capable of recovering missing DTC values in well-log data with reasonable effectiveness.

5.2. Single-Well Curve Reconstruction Analysis

A representative well is selected for closer inspection of the local reconstruction behavior, and Figure 2 displays the juxtaposition of the true DTC curve, the masked version, and the reconstructed output.

Figure 2 displays how the reconstructed DTC curve behaves in relation to the original log. Along the depth axis, the reconstruction follows the overall variation trend of the true curve with reasonable fidelity. Within the intervals that were masked, it manages to reconnect the adjacent observed segments in a manner that is relatively smooth. This outcome suggests that the method is capable of drawing upon the multi-log information available to it while still preserving a degree of continuity in the DTC response.

The enlarged view further illustrates the reconstruction performance around the masked positions. In these local intervals, the reconstructed values are generally consistent with the corresponding true DTC values, indicating that the pro-

posed method can recover the local variation pattern of the DTC curve with satisfactory accuracy. Moreover, no obvious discontinuities or unrealistic oscillations are observed in the reconstructed curve, suggesting that the depth-wise smoothness correction helps preserve local continuity across the missing intervals.

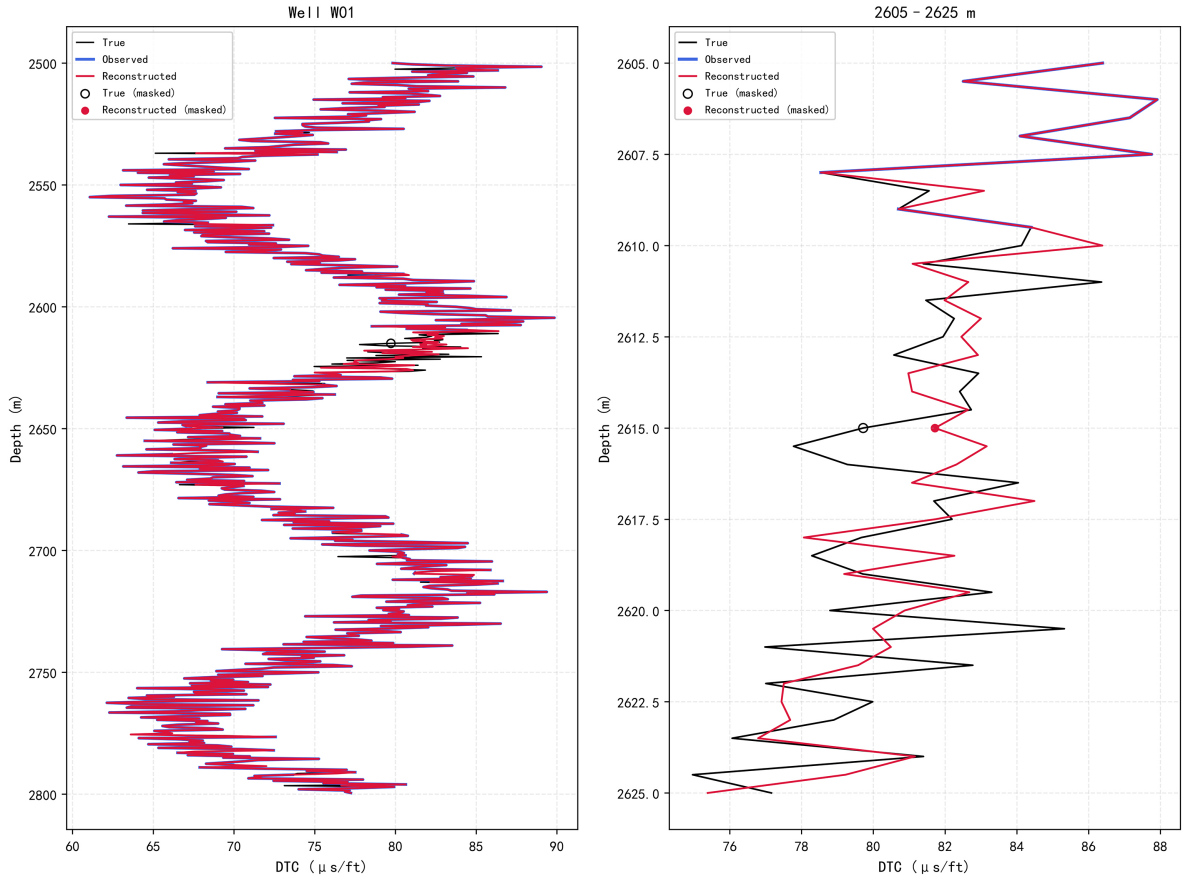


Figure 2. Single-well DTC reconstruction results: overall view and local zoom-in.

Overall, the single-well reconstruction result provides an intuitive visualization of the proposed method's performance. The reconstructed DTC curve not only matches the true curve closely in value, but also maintains the expected continuity along the depth direction.

5.3. Scatter Plot Analysis of Reconstructed and True Values

A scatter plot constructed at the masked positions, shown in **Figure 3**, further examines the agreement between the reconstructed DTC values and the true measurements.

Figure 3 shows that the bulk of the scatter points lie clustered around the diagonal line. Such an alignment points to a generally favorable agreement between the reconstructed values and the true measurements, and it is consistent with the relatively high R^2 and Corr figures that were reported earlier in Section 5.1.

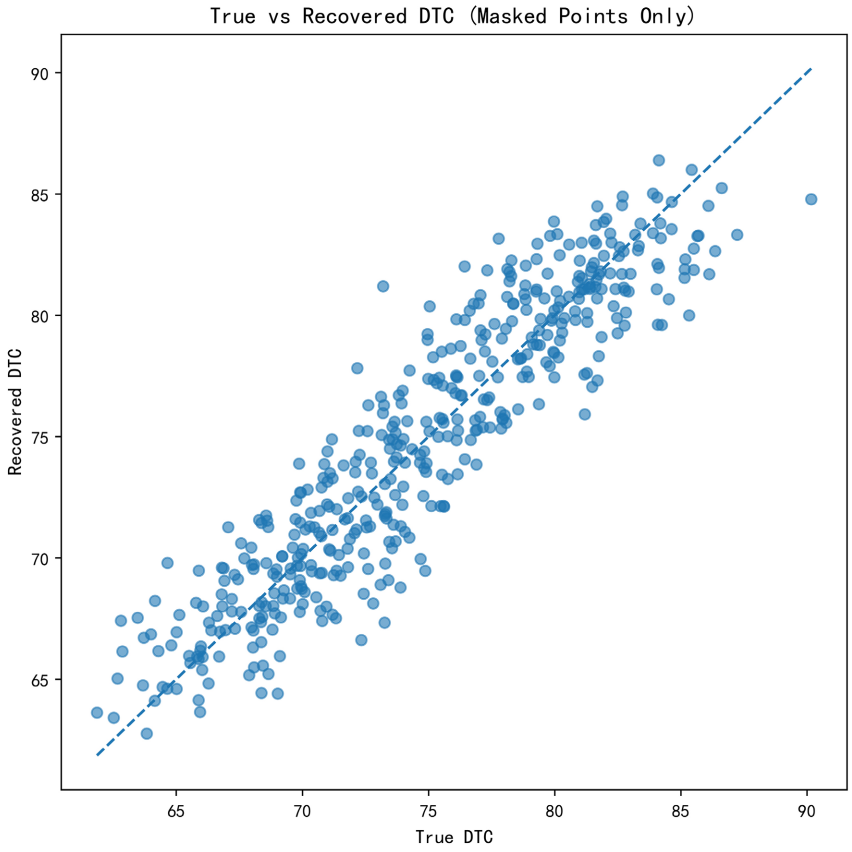


Figure 3. Scatter plot of reconstructed versus true DTC values at masked positions.

5.4. Error Distribution Analysis

To further analyze the distribution characteristics of the reconstruction errors, an error histogram is presented in **Figure 4**.

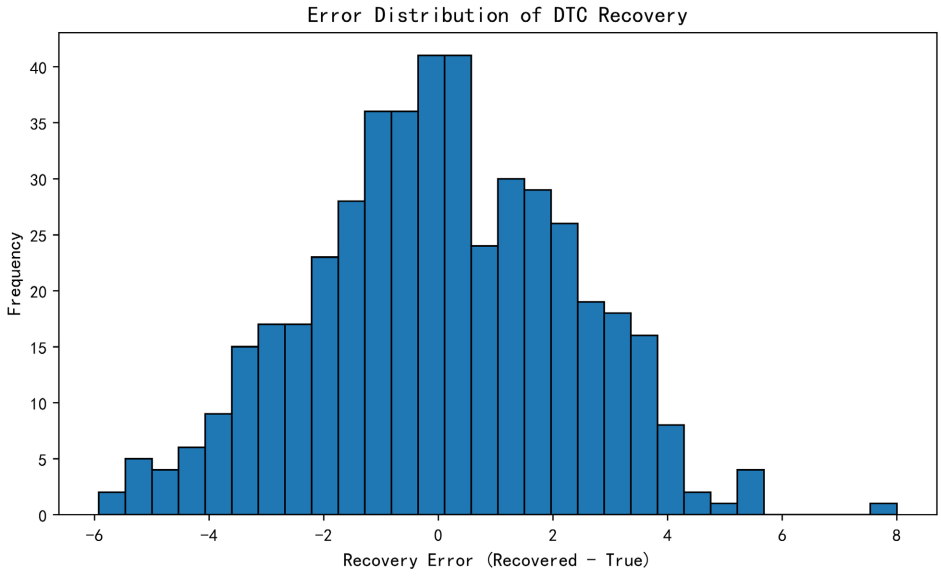


Figure 4. Error distribution of DTC recovery.

As shown in Figure 4, most reconstruction errors are concentrated around zero within a relatively narrow range. This indicates that, for the majority of masked samples, the reconstructed DTC values are close to the corresponding true values, reflecting stable reconstruction behavior across the missing positions.

5.5. Wellwise Stability Analysis

To evaluate the cross-well stability of the proposed method, the MAE of DTC reconstruction is calculated separately for each well. The well-wise MAE results are shown in Figure 5.

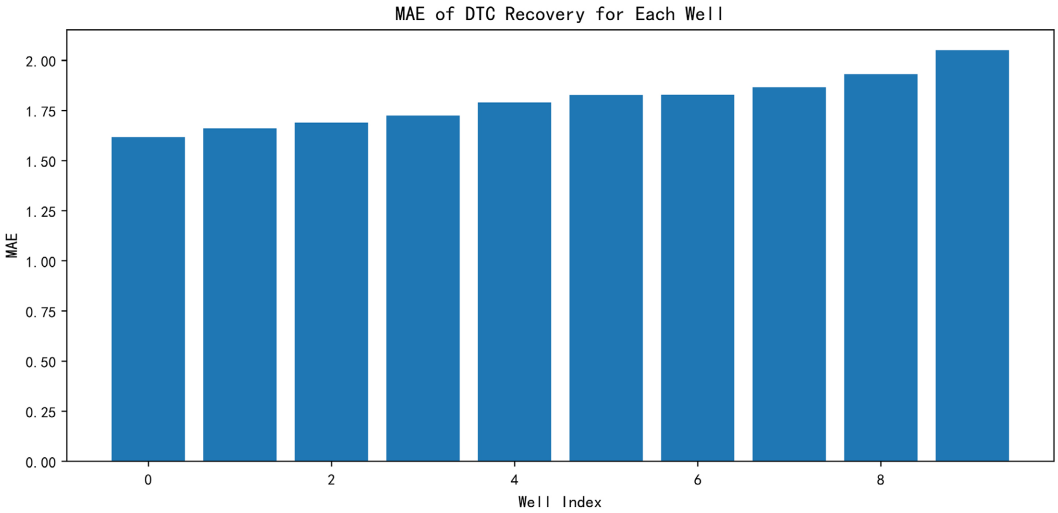


Figure 5. Well-wise MAE values of DTC reconstruction across all test wells.

As shown in Figure 5, the MAE values of different wells remain within a relatively limited range. This indicates that the proposed method does not rely on the characteristics of a single well, but maintains stable reconstruction performance across multiple wells. Such cross-well stability is important for well-log reconstruction, because logging responses may vary with geological conditions and reservoir properties.

Although several wells exhibit slightly higher MAE values, the overall inter-well variation remains limited. These differences may be related to stronger local fluctuations in the DTC curves, more complex lithological changes, or the distribution of masked positions within individual wells. Nevertheless, the well-wise results show that the proposed method maintains satisfactory robustness across different wells.

Overall, the well-wise MAE analysis demonstrates the cross-well applicability of the proposed reconstruction framework.

5.6. Comparison with Baseline Methods

Comparative experiments are conducted with two baseline methods to further evaluate the effectiveness of the proposed method: linear interpolation and pure

low-rank recovery without depth-wise smoothness correction [30]. The quantitative comparison is summarized in Table 3, and the corresponding visual comparison is shown in Figure 6.

As shown in Table 3, pure low-rank recovery achieves better reconstruction performance than linear interpolation. This result indicates that the low-rank factorization framework can effectively exploit the global correlation structure among multiple logging variables. Unlike linear interpolation, which relies only on local depth-wise continuity within a single log, low-rank recovery incorporates auxiliary information from other logging curves and therefore provides a more effective basis for recovering missing DTC values.

The proposed method further improves the reconstruction results compared with pure low-rank recovery. Although the numerical improvements in some metrics are moderate, consistent gains are obtained in MAE, RMSE, MAPE, R^2, and Corr. This suggests that the depth-wise smoothness correction provides additional information beyond the global low-rank structure. By suppressing unrealistic local fluctuations and enhancing vertical continuity, the correction makes the reconstructed DTC curve more consistent with the expected geological behavior.

Table 3. Comparison of reconstruction performance among three methods for DTC curve.

Method	Curve	MAE	RMSE	MAPE (%)	R^2	Corr
Low-Rank + Smoothing	DTC	1.810	2.263	2.429	0.848	0.923
Pure Low-Rank	DTC	1.833	2.283	2.460	0.845	0.922
Linear Interpolation	DTC	2.517	3.144	3.395	0.707	0.847

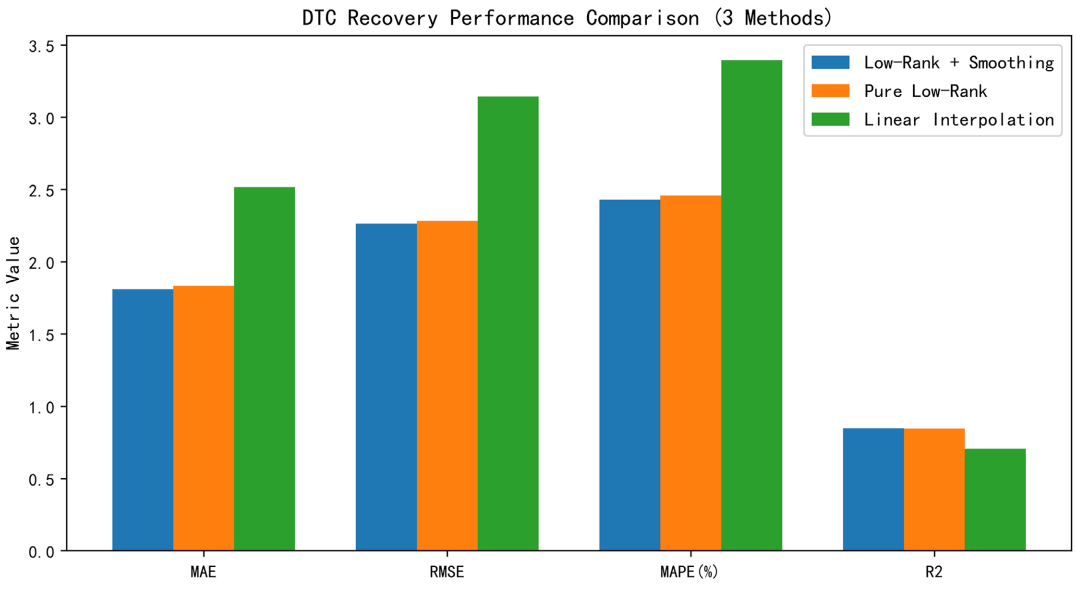


Figure 6. Comparison of the proposed method with baseline methods for DTC reconstruction.

The visual comparison in Figure 6 is consistent with the quantitative results in Table 3. Among the three methods, the proposed method achieves the lowest

MAE, RMSE, and MAPE, together with the highest R^2 and Corr. These results show that the combination of global low-rank modeling and local depth-wise continuity correction leads to more accurate and geologically plausible DTC reconstruction.

5.7. Performance Analysis under Different Missing Patterns

To further examine the applicability of the proposed method under different missing-data conditions, the masked DTC samples were grouped according to the missing pattern and the length of continuous missing intervals. The corresponding RMSE comparison is presented in Figure 7.

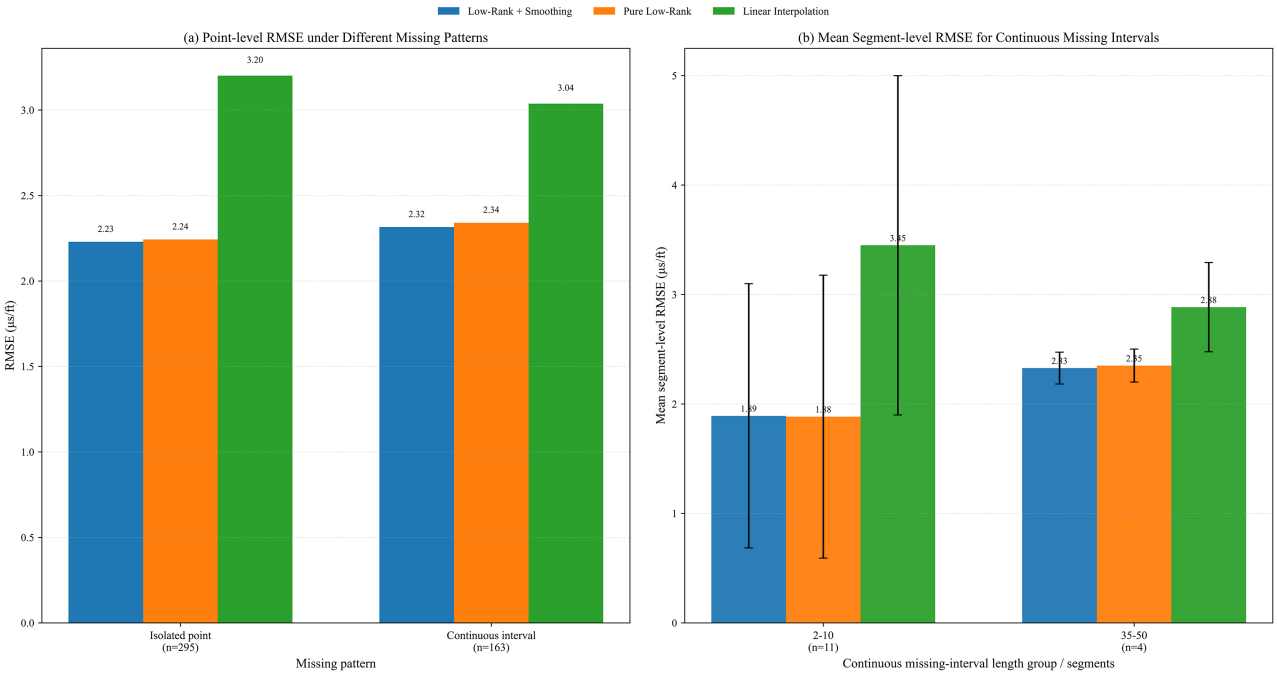


Figure 7. DTC reconstruction RMSE for different missing patterns and interval lengths.

As shown in Figure 7(a), the proposed Low-Rank + Smoothing method achieves the lowest point-level RMSE under both isolated missing points and continuous missing intervals. Compared with linear interpolation, the low-rank-based methods show a clear advantage, indicating that the correlations among multiple logging curves provide more effective information for DTC reconstruction than relying only on local depth-wise interpolation.

Although the RMSE difference between Low-Rank + Smoothing and pure low-rank recovery is relatively small, the proposed method consistently yields slightly lower errors under both missing patterns. This result indicates that the low-rank structure plays the dominant role in capturing the global correlations among different logging curves and is the main source of numerical accuracy improvement. Meanwhile, the depth-wise smoothing correction further improves the local stability of the reconstructed DTC curve. For continuous missing intervals, this cor-

rection is meaningful because the reconstructed values should not only be close to the true measurements, but also remain consistent with the expected depth-wise continuity of well-log responses.

Figure 7(b) further evaluates the continuous missing intervals from a segment-level perspective. The error bars represent the standard deviation of segment-level RMSE values within each length group. The 2 - 10-sample group contains 11 runs formed by adjacent masked positions, while the 35 - 50-sample group contains the four deliberately generated long intervals. The relatively large error bar of linear interpolation in the 2 - 10 group indicates that the reconstruction difficulty varies considerably among short continuous missing intervals, suggesting that some short intervals may be located in locally fluctuating DTC sections. In the 35 - 50 group, the lower RMSE of linear interpolation does not imply that longer gaps are easier to reconstruct. Instead, it may be related to the smoother local morphology of the selected long intervals and the limited number of segments in this group.

5.8. Discussion

The results indicate that combining low-rank recovery with depth-wise smoothness correction is effective for reconstructing missing DTC values. The improvement over linear interpolation suggests that DTC reconstruction should not rely only on the local continuity of a single curve. By organizing multiple logs into a unified matrix, the low-rank model can exploit the correlation structure among GR, RHOB, NPHI, RT, PHIT, SW, and DTC, thereby providing additional information for estimating the masked DTC entries.

The comparison with pure low-rank recovery further highlights the contribution of the depth-wise smoothness correction. Although low-rank modeling captures the global relationship among logging variables, it does not explicitly constrain the local morphology of the reconstructed curve along depth. The smoothing correction introduces the depth-continuity prior into the reconstruction process, which helps suppress unrealistic local fluctuations and improves the continuity of reconstructed DTC curves, especially within continuous missing intervals.

The visual results are consistent with the quantitative evaluation. The single-well curve comparison shows that the reconstructed DTC curve follows the main trend of the true curve and reconnects masked intervals smoothly. The scatter plot and error distribution show that most reconstructed values remain close to the corresponding true values, while the well-wise MAE results indicate stable reconstruction performance across different wells. These results support the numerical accuracy and geological plausibility of the proposed framework.

Some reconstruction errors are still observed, which may be related to strong local lithological variations, abrupt DTC changes, or the location and length of masked intervals. In addition, the present study focuses on DTC reconstruction under an artificial masking scheme. The applicability of the method to other logging curves, more irregular missing patterns, measurement noise, and systematic

parameter selection still requires further investigation.

6. Conclusions

This study proposed a DTC reconstruction method for incomplete well-log data by combining low-rank matrix recovery with depth-wise smoothness correction. Multiple logging variables were organized into a unified observation matrix, allowing the reconstruction model to exploit the global correlation structure among different logs. On this basis, a depth-wise smoothness correction was introduced to preserve the vertical continuity of the reconstructed DTC curve.

Artificial masking experiments were conducted on complete well-log datasets to evaluate the reconstruction performance. The results were assessed using MAE, RMSE, MAPE, R^2, and the correlation coefficient. On the masked DTC samples, the proposed method achieved satisfactory reconstruction accuracy and outperformed linear interpolation and pure low-rank recovery. These results indicate that the combination of multi-log low-rank modeling and depth-continuity correction can improve the reconstruction of missing DTC values.

The masking scheme included both isolated missing points and continuous missing intervals, which allowed the proposed method to be tested under different missing-data patterns. The results show that the method can recover pointwise missing values while also maintaining reasonable curve continuity across continuous missing intervals. This suggests that the proposed framework has potential applicability to incomplete well-log datasets encountered in practical logging and reservoir evaluation.

This study is limited to DTC reconstruction under artificial masking conditions. Future work may extend the proposed framework to other types of logging curves and evaluate its performance under more irregular field missing patterns, including longer and more complex depth gaps. In addition, systematic parameter selection and adaptive smoothness control deserve further investigation.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Mukherjee, B., Sain, K. and Wu, X. (2024) Missing Log Prediction Using Machine Learning Perspectives: A Case Study from Upper Assam Basin. *Earth Science Informatics*, **17**, 3071-3093. https://doi.org/10.1007/s12145-024-01323-5

[2] Gama, P.H.T., Faria, J., Sena, J., Neves, F., Riffel, V.R., Perez, L., *et al.* (2024) Imputation in Well Log Data: A Benchmark for Machine Learning Methods. *Computers & Geosciences*, **196**, Article 105789. https://doi.org/10.1016/j.cageo.2024.105789

[3] Kim, M.J. and Cho, Y. (2024) Imputation of Missing Values in Well Log Data Using K-Nearest Neighbor Collaborative Filtering. *Computers & Geosciences*, **193**, Article 105712. https://doi.org/10.1016/j.cageo.2024.105712

[4] Fung, C.C., Wong, K.W., Eren, H., Charlebois, R. and Crocker, H. (1996) Modular Artificial Neural Network for Prediction of Petrophysical Properties from Well Log

Data. *Proceedings of IEEE Instrumentation and Measurement Technology Conference*, **46**, 1295-1299.

[5] Rolon, L., Mohaghegh, S.D., Ameri, S., Gaskari, R. and McDaniel, B. (2009) Using Artificial Neural Networks to Generate Synthetic Well Logs. *Journal of Natural Gas Science and Engineering*, **1**, 118-133. https://doi.org/10.1016/j.jngse.2009.08.003

[6] Akinnikawe, O., Lyne, S. and Roberts, J. (2018) Synthetic Well Log Generation Using Machine Learning Techniques. *Proceedings of the 6th Unconventional Resources Technology Conference*, Houston, 23-25 July 2018.

[7] Zhang, D., Chen, Y. and Meng, J. (2018) Synthetic Well Logs Generation via Recurrent Neural Networks. *Petroleum Exploration and Development*, **45**, 629-639. https://doi.org/10.1016/s1876-3804(18)30068-5

[8] Pham, N., Wu, X. and Zabihi Naeini, E. (2020) Missing Well Log Prediction Using Convolutional Long Short-Term Memory Network. *Geophysics*, **85**, WA159-WA171. https://doi.org/10.1190/geo2019-0282.1

[9] Wang, J., Cao, J., Fu, J. and Xu, H. (2022) Missing Well Logs Prediction Using Deep Learning Integrated Neural Network with the Self-Attention Mechanism. *Energy*, **261**, Article 125270. https://doi.org/10.1016/j.energy.2022.125270

[10] Kanfar, R., Shaikh, O., Yousefzadeh, M. and Mukerji, T. (2020) Real-Time Well Log Prediction from Drilling Data Using Deep Learning. *International Petroleum Technology Conference*, Dhahran, 13-15 January 2020. https://doi.org/10.2523/iptc-19693-ms

[11] Wang, J., Cao, J., You, J., Cheng, M. and Zhou, P. (2021) A Method for Well Log Data Generation Based on a Spatio-Temporal Neural Network. *Journal of Geophysics and Engineering*, **18**, 700-711. https://doi.org/10.1093/jge/gxab046

[12] Lin, L., Wei, H., Wu, T., Zhang, P., Zhong, Z. and Li, C. (2023) Missing Well-Log Reconstruction Using a Sequence Self-Attention Deep-Learning Framework. *Geophysics*, **88**, D391-D410. https://doi.org/10.1190/geo2022-0757.1

[13] Wang, H., Liu, X., Yang, Y. and Zhang, Y. (2024) FlexLogNet: A Flexible Deep Learning-Based Well-Log Completion Method with Heterogeneous Graph Neural Networks. *Computers & Geosciences*, **187**, Article 105611.

[14] Cheng, C., Gao, Y., Chen, Y., Jiao, S., Jiang, Y., Yi, J., *et al.* (2022) Reconstruction Method of Old Well Logging Curves Based on BI-LSTM Model—Taking Feixianguan Formation in East Sichuan as an Example. *Coatings*, **12**, Article 113. https://doi.org/10.3390/coatings12020113

[15] Feng, R., Grana, D. and Balling, N. (2021) Imputation of Missing Well Log Data by Random Forest and Its Uncertainty Analysis. *Computers & Geosciences*, **152**, Article 104763. https://doi.org/10.1016/j.cageo.2021.104763

[16] Tatsipie, N.R.K. and Sheng, J.J. (2021) Generating Pseudo Well Logs for a Part of the Upper Bakken Using Recurrent Neural Networks. *Journal of Petroleum Science and Engineering*, **200**, Article 108253. https://doi.org/10.1016/j.petrol.2020.108253

[17] Gunasekar, S., Yamada, M., Yin, D. and Chang, Y. (2015) Consistent Collective Matrix Completion under Joint Low Rank Structure. *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*, **38**, 306-314.

[18] Wang, Z., Lai, M., Lu, Z., Fan, W., Davulcu, H. and Ye, J. (2015) Orthogonal Rank-One Matrix Pursuit for Low Rank Matrix Completion. *SIAM Journal on Scientific Computing*, **37**, A488-A514. https://doi.org/10.1137/130934271

[19] Wang, Y., Xia, K., Niu, W. and Ahamd, A. (2017) A Denoising Method by Low-Rank Matrix Recovery and Its Application in Oil Well Logging. *Geophysical Prospecting*

for Petroleum, **56**, 644-650.

[20] Acharya, S. and Fabian, K. (2023) Dynamic Depth Alignment of Well-Bore Measurements Using Machine Learning. 84*th EAGE Annual Conference & Exhibition*, **2023**, 1-5. https://doi.org/10.3997/2214-4609.2023101047

[21] Westeng, K., Aursand, P., Viset, F. and Van Crombrugge, Y. (2026) Dynamic Depth Alignment of Well Logs: A Continuous Optimization Framework for Enhanced Petrophysical and Rock Physics Interpretation. *Petrophysics—The SPWLA Journal of Formation Evaluation and Reservoir Description*, **67**, 54-67. https://doi.org/10.30632/pjv67n1-2026a4

[22] Pilkington, M. and Todoeschuck, J.P. (1991) Naturally Smooth Inversions with a Priori Information from Well Logs. *Geophysics*, **56**, 1811-1818. https://doi.org/10.1190/1.1442993

[23] Wu, X. and Caumon, G. (2017) Simultaneous Multiple Well-Seismic Ties Using Flattened Synthetic and Real Seismograms. *Geophysics*, **82**, IM13-IM20. https://doi.org/10.1190/geo2016-0295.1

[24] Cai, J., Candès, E.J. and Shen, Z. (2010) A Singular Value Thresholding Algorithm for Matrix Completion. *SIAM Journal on Optimization*, **20**, 1956-1982. https://doi.org/10.1137/080738970

[25] Shi, L. and Zou, C. (2020) Noisy Low-Rank Matrix Completion under General Bases. *Stat*, **9**, e303. https://doi.org/10.1002/sta4.303

[26] Yang, L., Wang, S., Chen, X., Chen, W., Saad, O.M. and Chen, Y. (2023) Deep-Learning Missing Well-Log Prediction via Long Short-Term Memory Network with Attention-Period Mechanism. *Geophysics*, **88**, D31-D48. https://doi.org/10.1190/geo2020-0749.1

[27] Keshavan, R.H., Montanari, A. and Oh, S. (2010) Matrix Completion from a Few Entries. *IEEE Transactions on Information Theory*, **56**, 2980-2998. https://doi.org/10.1109/tit.2010.2046205

[28] Zhao, L., Babu, P. and Palomar, D.P. (2016) Efficient Algorithms on Robust Low-Rank Matrix Completion against Outliers. *IEEE Transactions on Signal Processing*, **64**, 4767-4780. https://doi.org/10.1109/tsp.2016.2572049

[29] Hastie, T., Mazumder, R., Lee, J.D. and Zadeh, R. (2015) Matrix Completion and Low-Rank SVD via Fast Alternating Least Squares. *Journal of Machine Learning Research*, **16**, 3367-3402.

[30] Wang, G., Teng, H., Qiao, L., Yu, H., Cui, Y. and Xiao, K. (2024) Well Logging Reconstruction Based on a Temporal Convolutional Network and Bidirectional Gated Recurrent Unit Network with Attention Mechanism Optimized by Improved Sand Cat Swarm Optimization. *Energies*, **17**, Article 2710. https://doi.org/10.3390/en17112710