ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

Data-driven high-resolution gas-bearing prediction in tight sandstones: A case study from block L, Eastern Ordos Basin

Lixin Tian[†]*, Shuai Sun[†], Qixin Li, Jingxue Shi

Cnooc Research Institute Ltd., Chaoyang District, Beijing, China

(This article belongs to the *Special Issue: Geophysical Inversion and Intelligent Prediction Technologies for Complex Hydrocarbon Reservoirs*)

Abstract

The Upper Paleozoic Shihezi Formation in Block L of the eastern Ordos Basin harbors extensive tight sandstone gas reservoirs. However, these reservoirs exhibit strong heterogeneity, thin sand bodies, and overlapping elastic properties between gas- and water-bearing layers, which significantly limit the effectiveness of conventional pre-stack inversion methods in delineating thin sand bodies and predicting gas saturation. To address these challenges, we propose an integrated high-resolution gas prediction technique combining geostatistical inversion with deep learning. First, within a Bayesian sequential inversion framework, we jointly inverted well-log data, seismic data, and geological constraints to obtain high-resolution elastic parameters, substantially improving the identification of thin sand bodies (<5 m). Second, we employed a long short-term memory network to extract temporal features from inverted elastic parameter sequences and establish a non-linear mapping between gas/water-sensitive attributes and water saturation; this step incorporates horizon constraints and an attribute optimization strategy to enhance prediction accuracy. Field applications demonstrated that our method achieved superior performance compared to conventional approaches, with an 85% consistency rate between predicted gas saturation and drilling results. The integration of geostatistical inversion and deep learning provides a robust workflow for characterizing thin, heterogeneous tight gas reservoirs, offering significant potential for optimizing exploration and development strategies in the Ordos Basin.

Keywords: Ordos Basin; Tight sandstone gas; Geostatistical inversion; Deep learning; Long short-term memory network; Gas-bearing prediction

[†]These authors contributed equally to this work.

***Corresponding author:**
Lixin Tian
(ssun30143@gmail.com)

Citation: Tian L, Sun S, Li Q, Shi J. Data-driven high-resolution gas-bearing prediction in tight sandstones: A case study from block L, Eastern Ordos Basin. *J Seismic Explor*. doi: 10.36922/JSE025320053

Received: August 6, 2025

Revised: September 1, 2025

Accepted: September 3, 2025

Published online: September 30, 2025

1. Introduction

The Permian Shihezi Formation in Block L of the northeastern Ordos Basin harbors large-scale tight sandstone gas reservoirs with proven geological reserves exceeding 10 billion cubic meters, making it a critical gas-producing interval in the basin.[1-4] These reservoirs are deposited in a fluvial-deltaic environment influenced by seasonal flooding, characterized by thin-bedded (2–12 m thick, with 70% of layers <5 m) and lenticular sand bodies exhibiting strong lateral heterogeneity and frequent vertical interbedding

with mudstones. The lithology primarily comprises quartz sandstones underlain by coal-bearing source rocks of the Shanxi-Taiyuan Formations, forming typical tight gas reservoirs through a "source-reservoir pressure differential" driving mechanism.[5] However, reservoir prediction in this area faces three major challenges: (i) impedance contrast limitations: post-stack impedance inversion is hindered by the minimal acoustic impedance contrast between sandstone and mudstone, restricting effective spatial prediction of sand bodies. (ii) Resolution constraints: conventional pre-stack inversion is limited by the seismic data's dominant frequency (30 Hz in the target zone), resulting in a theoretical resolution limit ($\lambda/4 \approx 38$ m) that far exceeds the average sandstone thickness (<15 m). (iii) Fluid discrimination difficulty: gas-bearing and water-bearing layers exhibit substantial overlap in P-impedance versus Vp/Vs crossplots, rendering rock physics template methods ineffective for gas saturation prediction. Geostatistical inversion, which integrates geological priors with stochastic simulation, has emerged as a key solution for thin-bed reservoir characterization.[6] This approach has been successfully validated in continental thin sandstones[7] and coal bed methane reservoirs.[8]

Extensive research has focused on seismic gas-bearing prediction.[9-14] Since gas saturation has minimal influence on seismic waveforms (often obscured by noise), conventional methods typically derive elastic parameters through seismic inversion before identifying gas-bearing zones. For instance, Zong *et al.*[15] developed fluid-sensitive factors via direct P- and S-wave inversion, while Zong and Yin[16] constructed sensitivity factors using amplitude versus offset (AVO) linear equations to estimate Young's modulus and Poisson's ratio, thereby reducing cumulative errors from traditional elastic parameter inversion. Although pre-stack inversion-derived elastic parameters can effectively identify reservoirs,[17,18] they remain limited by thin-bed tuning effects.

For the Shihezi Formation's thin tight sandstones—where gas-water overlap is severe—conventional fluid identification methods fail due to insufficient seismic resolution and ineffective elastic sensitivity factors. Recent advances in deep learning have introduced data-driven approaches for gas prediction.[19,20] Early work by Hampson *et al.*[21] demonstrated successful porosity prediction using probabilistic neural networks to extract seismic attributes from waveform data. Similarly, Zhong *et al.*[22] showed that connected neural network (CNN)-based permeability prediction models outperform traditional genetic algorithms, while Das and Mukerji[23] achieved direct porosity and clay content inversion from post-stack data using CNN-trained synthetic models. Notably, Chen *et al.*[24] found that recurrent neural networks (RNNs) significantly outperform support vector machines and random forests in well-log time-series modeling, offering new opportunities for time-sensitive reservoir parameter prediction.

To address these challenges, this study proposes a data-driven high-resolution gas-bearing prediction framework for tight sandstones, combining: (1) pre-stack geostatistical inversion to integrate well-log, seismic and geological data for high-resolution elastic parameter estimation and (2) a long short-term memory (LSTM) network to establish a nonlinear mapping between time-series elastic parameters and water saturation (Sw), constrained by geological horizons and optimized attribute selection. Field applications demonstrate that our method significantly improves thin-bed identification and gas-prediction accuracy, providing a robust technical solution for tight gas exploration in the Ordos Basin with broader applicability to similar reservoirs.

2. Methodology

2.1. Technical workflow

The direct inversion of Sw from seismic data remains challenging due to its significantly weaker sensitivity to seismic waveform characteristics compared to elastic parameters.[25] In contrast, elastic parameters not only predominantly control seismic wavefield dynamics[26] but also exhibit more quantifiable physical relationships with Sw through established rock physics models. To address these challenges, we developed an integrated workflow combining geostatistical inversion with deep learning (Figure 1), which consists of three key components: (i) high-resolution geostatistical inversion: conducting pre-stack geostatistical inversion using well logs, geological structural frameworks, and 3D seismic data to overcome the bandwidth limitations of conventional seismic inversion and obtain high-resolution elastic parameters for

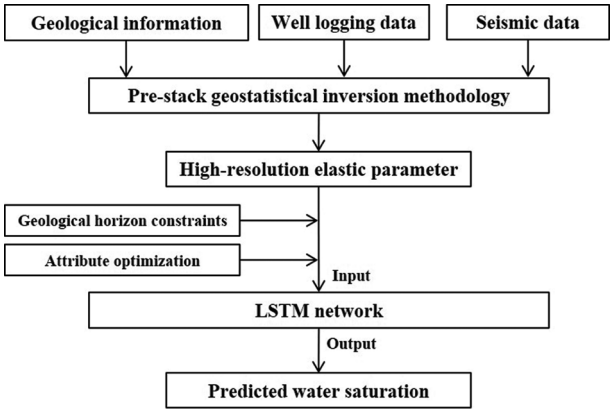

Figure 1. The workflow of data-driven high-resolution gas-bearing property prediction
Abbreviation: LSTM: Long short-term memory.

thin sand body identification; (ii) LSTM-based saturation modeling: leveraging the unique sequential modeling capability of LSTM networks while incorporating horizon constraints for segmented refinement learning, where attribute optimization techniques are employed to select elastic attribute combinations most sensitive to Sw, thereby establishing a sequential mapping between high-resolution elastic data and Sw to address the overlap issue of elastic parameters in gas-water layers; (iii) Model training and application: training the network using inverted traces adjacent to wells, with the fully trained LSTM model ultimately being applied to the tight sandstone reservoirs of the Shihezi Formation.

2.2. Geostatistical inversion

Unlike conventional deterministic inversion, geostatistical inversion statistically integrates prior information from well logs and geological data with seismic observations d_{obs} and estimates the posterior distribution of model parameters, m, through Bayesian inversion.[27] The expectation of the posterior probability solution is given by:

$$\tilde{\mathrm{m}} = \mathrm{m}_{prior} + \mathrm{C}_m \mathrm{G}^T (\mathrm{G}\mathrm{C}_m \mathrm{G}^T + \mathrm{C}_d)^{-1} \left(\mathrm{d}_{obs} - \mathrm{G}\mathrm{m}_{prior}\right) \quad \text{(I)}$$

The posterior covariance is expressed as:

$$\tilde{\mathrm{C}}_m = \mathrm{C}_m - \mathrm{C}_m \mathrm{G}^T (\mathrm{G}\mathrm{C}_m \mathrm{G}^T + \mathrm{C}_d)^{-1} \mathrm{G}\mathrm{C}_m \quad \text{(II)}$$

where $\mathrm{m}_{prior} = \left[\ln V_p, \ln V_s, \ln \rho\right]^T$ is a $3n_m$ column vector composed of smoothed background models for P-wave velocity, S-wave velocity, and density. C_m denotes the $3n_m \times 3n_m$ prior model covariance matrix; C_d is the seismic covariance matrix, estimated through well synthetic seismograms and field seismic data adjacent to wells.

Building upon the sequential simulation concept,[18,28] the sequential inversion framework classifies observed data into two distinct categories: Type A and Type B data. Type A data are direct measurements of model parameters, including well log data and previously simulated grid points. Type B data are indirectly acquired measurement data, specifically referring to pre-stack seismic angle gathers in this context. By jointly incorporating both data types, the forward equation can be reformulated as:

$$\begin{bmatrix} \mathrm{d}_{obsA} \\ \mathrm{d}_{obsB} \end{bmatrix} = \begin{bmatrix} \mathrm{G}_A & 0 \\ 0 & \mathrm{G}_B \end{bmatrix} \begin{bmatrix} \mathrm{m}_A \\ \mathrm{m}_B \end{bmatrix} + \begin{bmatrix} \mathrm{e}_A \\ \mathrm{e}_B \end{bmatrix} \quad \text{(III)}$$

Where d_{obs}, G_A, m_A, e_A represent the observed data, forward operator, model parameters, and error terms for Type A data, respectively; d_{obsB}, G_A, m_A, e_A denote the observed data, forward operator, model parameters, and error terms for type B data, respectively; G_A is simply an identity matrix. Performing Bayesian inversion on the joint data in **Equation III** yields the posterior expectation constrained by both well log data (type A) and pre-stack seismic data (type B):

$$\tilde{\mathrm{m}}_{A+B} = \mathrm{m}_{priorA+B} + \mathrm{C}_{mA+B} \mathrm{G}_{A+B}^T (\mathrm{G}_{A+B} \mathrm{C}_{mA+B} \mathrm{G}_{A+B}^T + \mathrm{C}_{dA+B})^{-1} (\mathrm{d}_{obsA} - \mathrm{G}\mathrm{m}_{priorA+B}) \quad \text{(IV)}$$

and posterior covariance:

$$\tilde{\mathrm{C}}_{mA+B} = \mathrm{C}_{mA+B} - \mathrm{C}_{mA+B} \mathrm{G}_{A+B}^T (\mathrm{G}_{A+B} \mathrm{C}_{mA+B} \mathrm{G}_{A+B}^T + \mathrm{C}_{dA+B})^{-1} \mathrm{G}_{A+B} \mathrm{C}_{mA+B} \quad \text{(V)}$$

where the covariance matrices C_{mA+B} (model) and C_{dA+B} (data) are formally expressed as:

$$\mathrm{C}_{mA+B} = \begin{bmatrix} \mathrm{C}_{mAA} & \mathrm{C}_{mAB} \\ \mathrm{C}_{mAB}^T & \mathrm{C}_{mBB} \end{bmatrix}, \mathrm{C}_{dA+B} = \begin{bmatrix} 0 & 0 \\ 0 & \mathrm{C}_{dBB} \end{bmatrix} \quad \text{(VI)}$$

The model covariance matrices for Type A and Type B data are denoted as C_{mAA} and C_{mBB}, respectively, while C_{mAB} represents their cross-covariance matrix. The data covariance matrix for type B observations is specified as C_{dBB}. The Bayesian sequential stochastic inversion framework treats well log data and previously simulated points as hard constraints. Under the joint constraints of geostatistical information and seismic data, these data participate in computing subsequent grid points. Consequently, well-derived information propagates throughout the stochastic simulation path, endowing the inverted elastic parameters with high-resolution characteristics.

2.3. LSTM network

LSTM networks,[29] a specialized variant of RNNs,[30] were specifically designed to address the vanishing gradient problem in traditional RNNs while preserving long-range temporal dependencies. The basic LSTM unit, illustrated in Figure 2, consists of three core components: the input gate, forget gate, and output gate. These gates selectively regulate information flow, enabling the modeling of long-term temporal dependencies. Specifically, this mechanism allows the network to dynamically store or discard temporal features, significantly enhancing its ability to process long sequential data. The input-output relationship of an LSTM unit can be described by the following equations:

$$f_t = \sigma(W_{xf} X_t + W_{yf} Y_{t-1} + W_{cf} \circ C_{t-1} + b_f) \quad \text{(VII-a)}$$

$$i_t = \sigma(W_{xi} X_t + W_{yi} Y_{t-1} + W_{ci} \circ C_{t-1} + b_i) \quad \text{(VII-b)}$$

$$C_t = f_t \circ C_{t-1} + i_t \circ \tanh(W_{xc} X_t + W_{yc} Y_{t-1} + b_c) \quad \text{(VII-c)}$$

$$O_t = \sigma(W_{xo}X_t + W_{yo}Y_{t-1} + W_{co} \circ C_t + b_o) \quad \text{(VII-d)}$$

$$Y_t = O_t \tanh(C_t) \quad \text{(VII-e)}$$

where ∘ denotes the Hadamard product; *represents convolution; W and b correspond to the weight matrices and bias vectors of the LSTM network, respectively; X_t signifies the input data; Y_t denotes the output parameters; C_t represents the cell state (memory unit); and i_t, f_t, and O_t indicate the input gate, forget gate, and output gate, respectively.

Figure 3 illustrates the deep learning architecture for Sw prediction based on an LSTM network. The core feature of this architecture lies in its use of elastic parameter time series as network inputs, as opposed to traditional single-point input patterns. Specifically, at each time step t, a multidimensional elastic vector X_t = [EP_1(t), EP_2(t), EP_3(t)…] is fed into the network. This design offers dual advantages: first, the temporal characteristics of the data are explicitly modeled through the LSTM gating mechanism; second, the joint input of elastic parameters at the same time step (such as P-impedance and Vp/Vs) captures the petrophysical correlations between parameters. In terms of technical implementation, the length in the time dimension is selected based on the wavelength of the seismic wavelet, while the choice of elastic data is determined through attribute optimization results. Regarding the network output mechanism, the LSTM outputs the hidden state Y_t at each time step, which serves as input to a fully connected

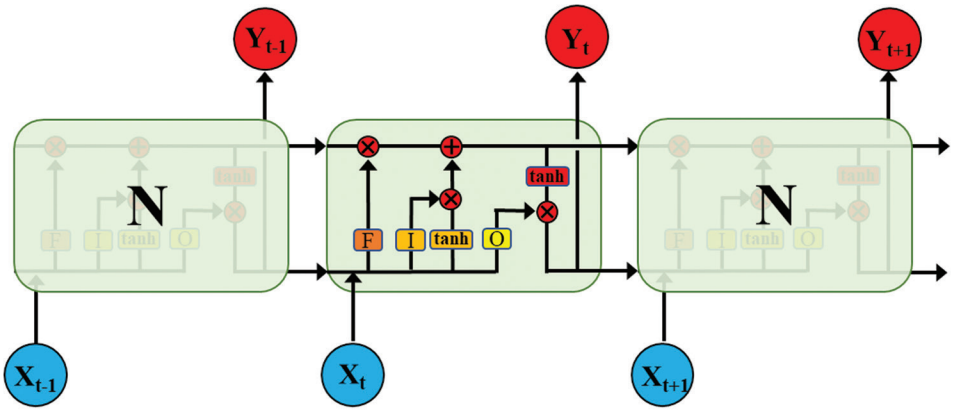

Figure 2. The architecture of the LSTM network. Symbol "N" represents the unit of LSTM; the terms X_t, $t = 1, 2, ..., n$ and Y_t, $t = 1, 2, ..., n$ represent input and output sequence data, respectively.
Abbreviation: LSTM: Long short-term memory.

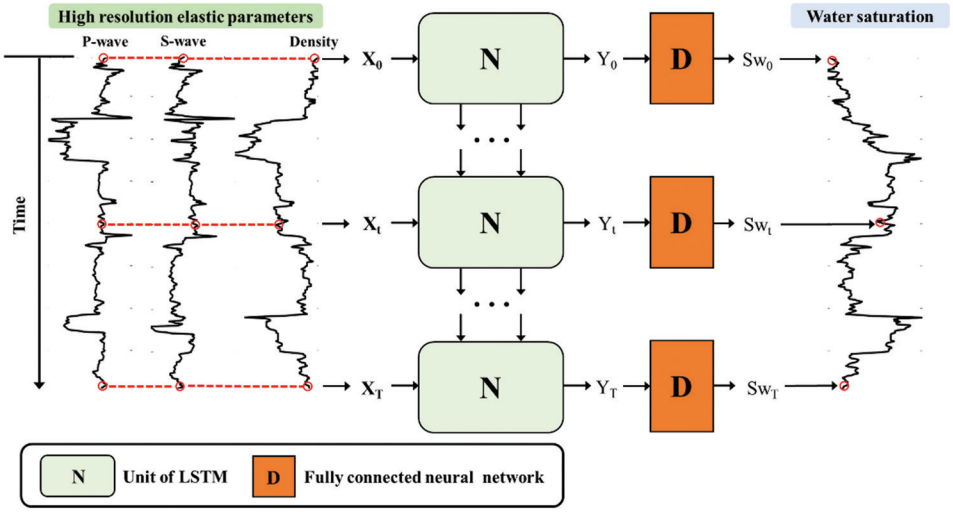

Figure 3. Deep learning-based high-resolution gas-bearing property prediction architecture. Symbol "N" represents the unit of LSTM, with detailed structure shown in Figure 2. Symbol "D" represents the FCNN, which is used in the conversion of time-series to Sw.
Abbreviation: FCNN: Fully connected neural network; LSTM: Long short-term memory; Sw: Water saturation.

neural network (FCNN). The predicted Sw value can be expressed as:

$$Sw_t = \sigma_P(W_\rho Y_t + b_\rho) \quad \text{(VIII)}$$

where the weight matrix W_p and bias vector b_p constitute the trainable parameters of the FCNN, and Y_t represents the hidden state output of the LSTM unit at time step t. The network employs ReLU activation functions[31] for nonlinear transformation: $\sigma_p(x) = max\ (0,x)$. Notably, this architecture adopts a time-step-shared weight parameter mechanism, meaning the FCNN's weights and biases remain constant across different time steps. This design achieves parameter efficiency and overfitting suppression. The complete set of learnable parameters includes:

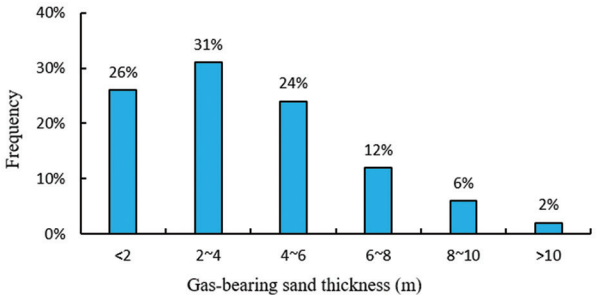

Figure 4. Statistical distribution of gas-bearing sandstone thickness in the Shihezi Formation, Ordos Basin

$$\Theta = \{W_{xf}, W_{yf}, W_{cf}, b_f, W_{xi}, W_{yi}, W_{ci}, b_i, W_{xc}, W_{yc}, b_c, W_{xo}, W_{yo}, W_{co}, b_o, W_\rho, b_\rho\} \quad \text{(IX)}$$

3. Case study

3.1. Target formation overview

The target interval of the Shihezi formation in the study area is predominantly composed of tight sandstone intercalated with mudstone. Statistical analysis reveals that individual sand bodies have an average thickness below 15 m, with gas-bearing sand bodies thinner than 6 m constituting over 80% of the total reservoir units (Figure 4). Petrophysical characterization demonstrates that these tight sandstones are distinguished by remarkably low Vp/Vs ratios (Vp/Vs < 1.8). Figure 5 presents a comparative analysis between Vp/Vs curves derived from well log interpretation (red solid line) and conventional pre-stack inversion (blue solid line), with lithological interpretation indicating sandstone intervals in yellow and gas-bearing sandstones in red. Due to resolution constraints inherent in conventional inversion methodologies, only thicker sand bodies can be confidently identified, while thinner sand bodies exhibit poor resolution. This resolution limitation directly compromises accurate reservoir assessment. Furthermore, petrophysical cross-plot analysis indicates substantial overlap between gas-bearing and water-bearing layers within the sandstone reservoirs

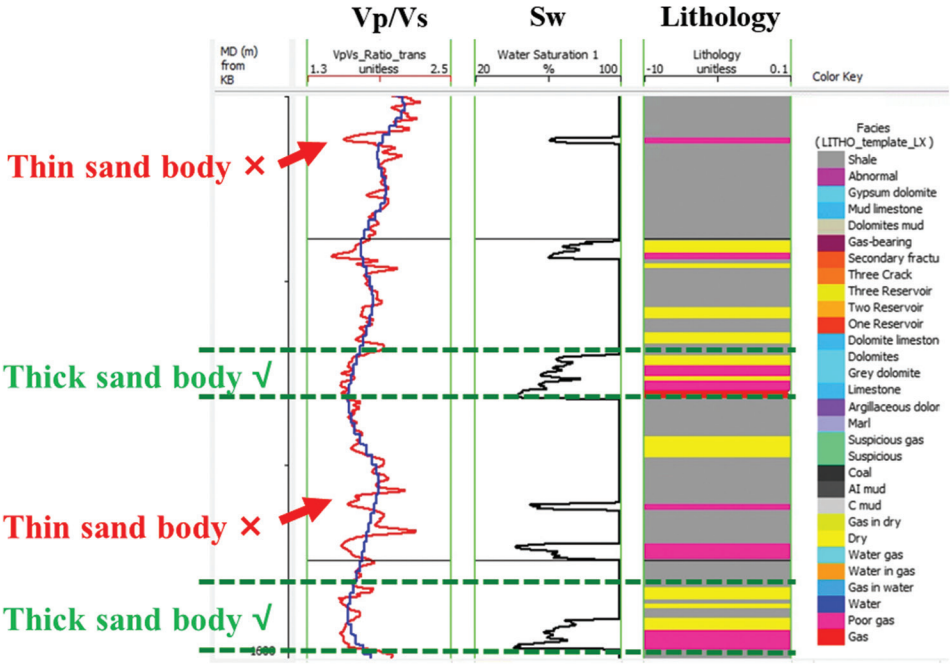

Figure 5. Lithology identification using conventional pre-stack inversion Vp/Vs. The red and blue solid lines represent the well logging and conventional pre-stack inverted Vp/Vs curves, respectively. In the lithology interpretation, yellow and red colors indicate sandstone and gas-bearing sandstone, respectively.

(Figure 6). This overlap significantly challenges effective fluid discrimination when employing conventional elastic parameter cross-plotting techniques.

3.2. High-resolution pre-stack geostatistical inversion

3.2.1. Vertical range determination from well log statistics

The variogram, serving as a fundamental geostatistical tool, provides quantitative characterization of reservoir parameter spatial variability.[32] The range parameter plays a particularly crucial role in defining reservoir thickness and lateral continuity patterns. Analysis of 94 well logs from the Shihezi Formation (Figure 7) and geological data demonstrates that thin sand layers are widely developed in the study area, with approximately 80% of gas-bearing sand bodies having thicknesses <6 m. Based on this finding, the vertical range was determined to be 1 ms. Simultaneously, statistical results reveal significant lateral variations in tight sandstones, with 85% of sand body widths distributed within the 300–1500 m range. Consequently, the lateral range was set to 800 m. It must be emphasized that determining the lateral range requires comprehensive consideration of both the depositional characteristics of the target formation and reservoir prediction results to ensure the rationality of parameter settings.

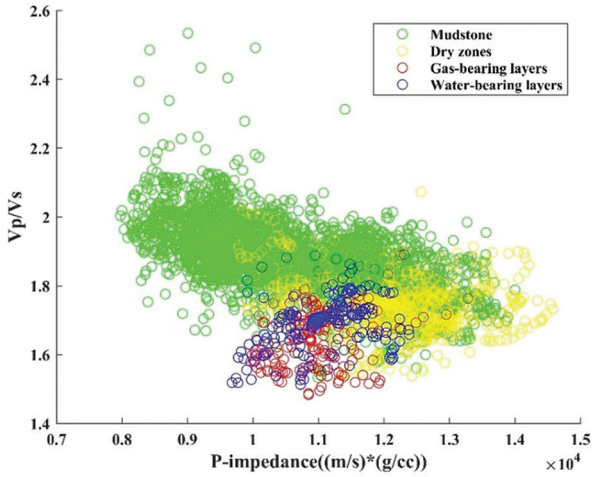

Figure 6. Crossplot analysis of P-impedance versus Vp/Vs in Shihezi formation, where green, yellow, blue, and red circles represent shale, dry layer, water-bearing layer, and gas-bearing layer, respectively. Significant overlap is observed between gas-bearing and water-bearing sandstones.

3.2.2. Application results of elastic parameter inversion

The pre-stack geostatistical inversion method integrates four key data types: (i) geological grid models, (ii) variogram parameters, (iii) well-log data, and (iv) pre-stack seismic data to construct a Bayesian inversion framework for estimating posterior probability distributions. We employed sequential Gaussian simulation to generate multiple realizations (n = 100) of elastic parameters from the tight sandstone reservoirs in the Shihezi Formation, Block L. The inversion used pre-stack seismic gathers with 20 m CDP spacing, 1 ms time sampling, and 5°–35° incidence angles. Figure 8 compares three datasets near well A: (i) measured well-log data (black curves), (ii) conventional pre-stack inversion results (blue curves), and (iii) P50 geostatistical inversion results (red curves) for Vp, Vs, density, and Vp/Vs ratio. Applying the Vp/Vs < 1.8 sandstone discrimination criterion, the geostatistical approach resolved thin gas-bearing sand layers (2–5 m thickness) that conventional inversion failed to detect. Figure 9 displays cross-well sections comparing: (i) conventional versus (ii) Geostatistical inversion results, annotated with lithology interpretations from Wells B–D (yellow: sandstone; red: gas-bearing sand). The geostatistical Vp/Vs results show superior vertical resolution (arrow indicators). Collectively, these results demonstrate that pre-stack geostatistical inversion technology can significantly improve resolution and enable detailed characterization of thin sand bodies.

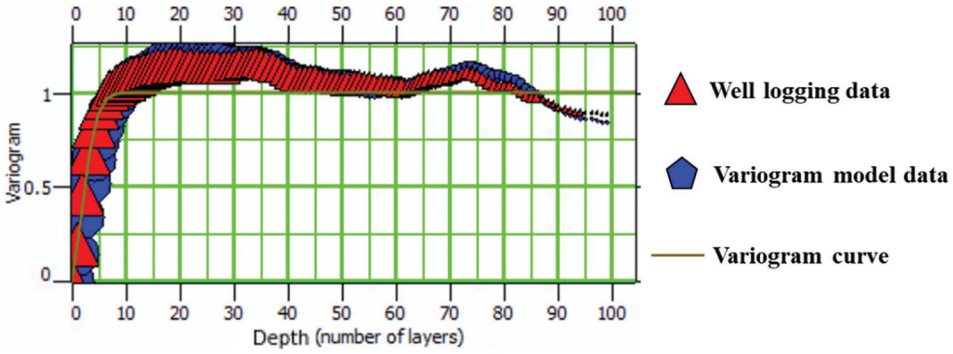

Figure 7. Variogram analysis of well logging parameters in Shihezi formation

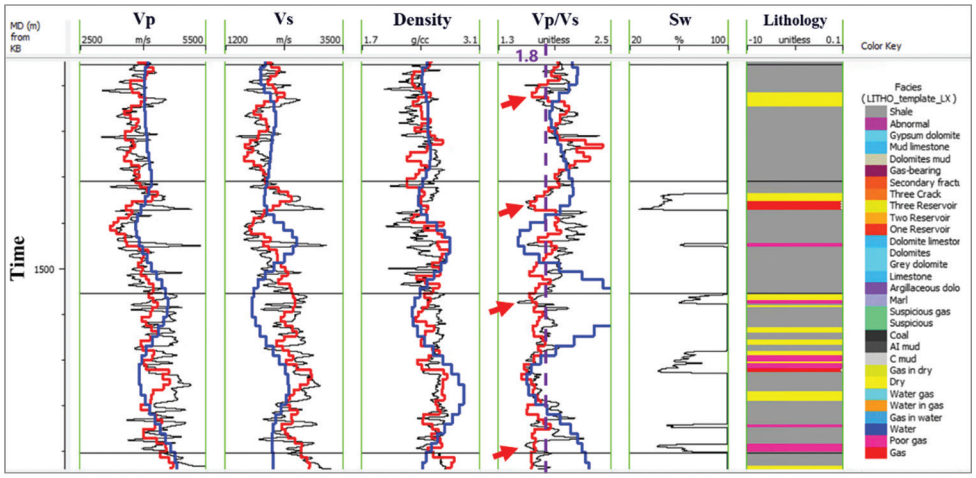

Figure 8. Comparison of inversion results near Well A seismic gather. The black, blue, and red solid lines represent the well log data, conventional pre-stack inversion results, and geostatistical pre-stack inversion results, respectively.

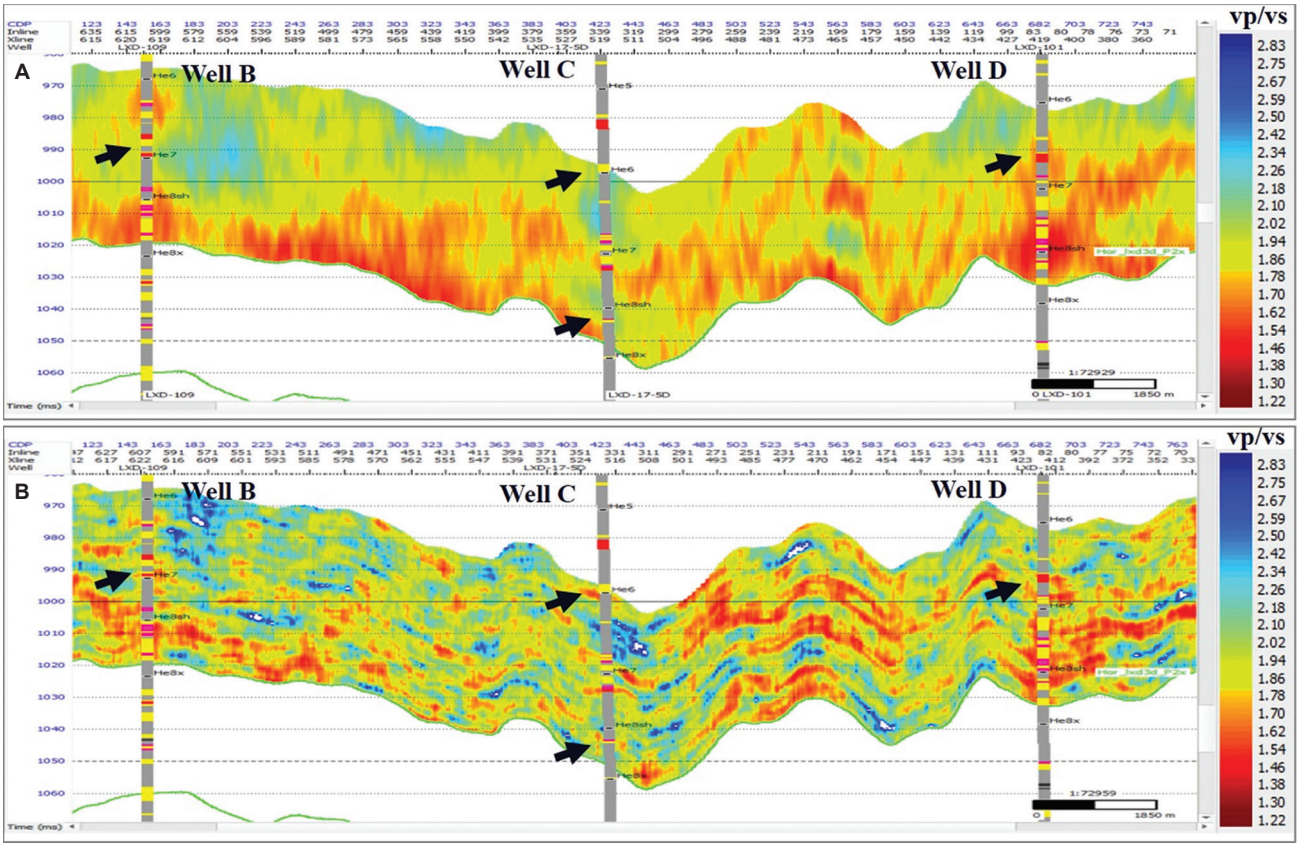

Figure 9. Comparative analysis of well-tie profiles based on elastic parameter inversion. (A) Conventional pre-stack inversion. (B) High-resolution geostatistical pre-stack inversion.

3.3. LSTM-based Sw prediction

3.3.1. Geological horizon constraints implementation

In this section, we employ a geological horizon-constrained approach to enhance the predictive performance of the LSTM network. The specific implementation procedure is as follows: first, we utilize the horizon information of the Shihezi formation to segment both well logging and seismic data. During data processing, we adhere to the "intra-horizon cross-validation" principle, meaning that data

from the same horizon can be mutually used as training and testing sets, while data from different horizons are strictly isolated. For the experiment, Well A was selected as the blind test well, with data from Wells B, C, and D used to train the LSTM network. Figure 10 presents a comparative analysis of the results before and after applying constraints: Figure 10A displays the unconstrained data from Well A, including three Vp, Vs, and density; Figure 10B shows the horizon-constrained data, where eight distinct colors represent eight different depositional periods; Figure 10C compares the lithology prediction results: the left side presents well log interpretation results (yellow indicating sandstone and brown indicating mudstone), where the unconstrained LSTM predictions show significant misjudgment in sand-rich intervals with lower resolution for thin sand layers; whereas after applying geological constraints, the accuracy of sandstone-mudstone identification improves markedly. This improvement stems from the following mechanism: the macroscopic trends of well logs reflect variations in depositional environments across different geological periods. By dividing well logs into contemporaneous depositional segments through geological constraints, the differences in data distribution within each segment more authentically reflect lithological variations, thereby enabling the LSTM network to more accurately learn reservoir characteristics.

3.3.2. Attribute optimization and model training

An LSTM neural network model was constructed for gas-bearing prediction using high-resolution elastic inversion results as training data. The fundamental elastic parameters, including Vp, Vs, and density, were mathematically processed to compute multiple gas-sensitive indicators such as Vp/Vs ratio, Poisson's ratio, acoustic impedance, and bulk modulus. Additional sensitive attributes were subsequently generated through mathematical transformations. During network training, an attribute selection method based on loss function gradient descent was employed, ultimately identifying the 15 most contributive key attributes (Figure 11), which were then used as the final input sequences for the LSTM network. The specific architectural parameters of the LSTM network are presented in Table 1. The input vector X_t adopts a 150 ms time-series length, with this parameter setting matching both the seismic wavelet length and LSTM timesteps. The network structure comprises 32 hidden units, with the output layer Y_t having dimensions of 150 × 32. After transformation through the fully connected layer, the final output is a predicted sequence of Sw with dimensions of 150 × 1. Model validation results (Figure 12) demonstrate that on the training set, the average correlation coefficient of predictions across multiple wells reached 0.86, with a mean absolute error of Sw at 3.9%. In blind well validation, the average correlation coefficient was 0.76, with Sw mean absolute error approximately 8%, confirming the neural network model's excellent training effectiveness and prediction accuracy. Furthermore, statistical results of drilling confirmation rates (Figure 12C) show that all eight validation wells achieved match rates exceeding 75%, with Wells 1 and 5 reaching 85%, thereby further verifying the reliability of this method in practical applications.

3.3.3. Application results of gas-bearing prediction

The trained LSTM model was applied to predict tight sandstone reservoirs in the Shihezi Formation of the Ordos Basin. Figure 13 displays a gas saturation prediction profile intersecting wells, visualized using a gradient color scale from blue to red (representing gas saturation ranging from

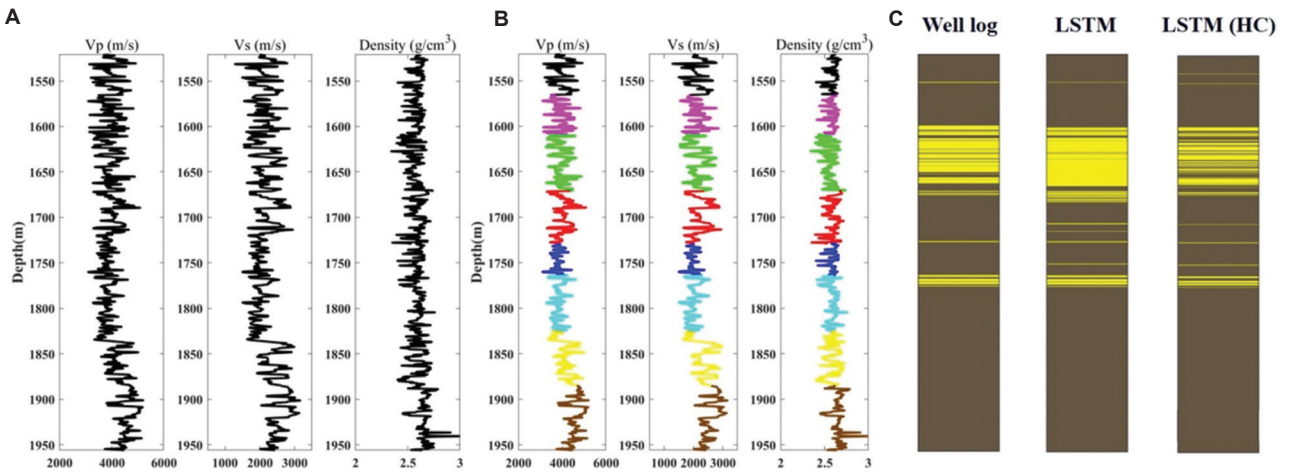

Figure 10. Comparison of lithology prediction results in Well A with/without horizon constraints, yellow and brown represent sandstone and mudstone, respectively. (A) Before applying horizon constraints. (B) Applying horizon constraints. (C) Comparison of lithology prediction results.

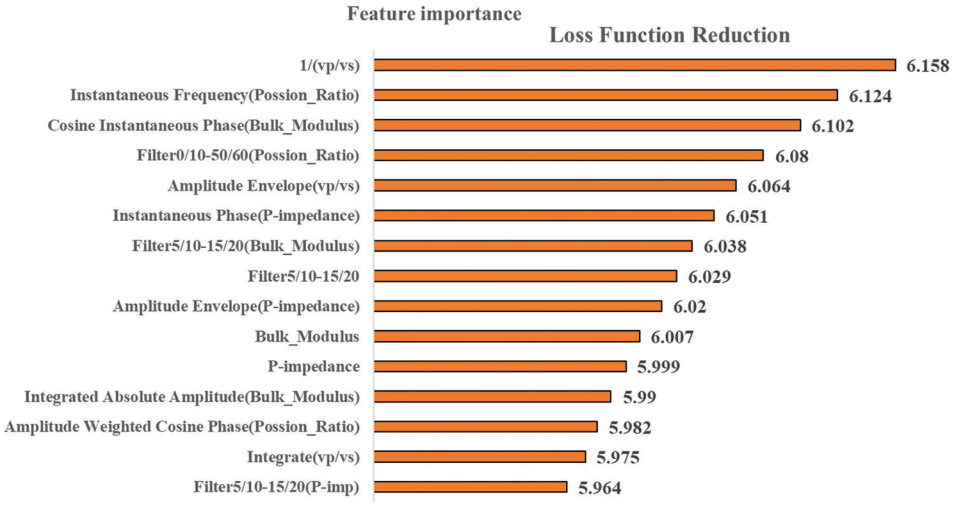

Figure 11. Fifteen optimized elastic attributes as LSTM network inputs
Abbreviation: LSTM: Long short-term memory.

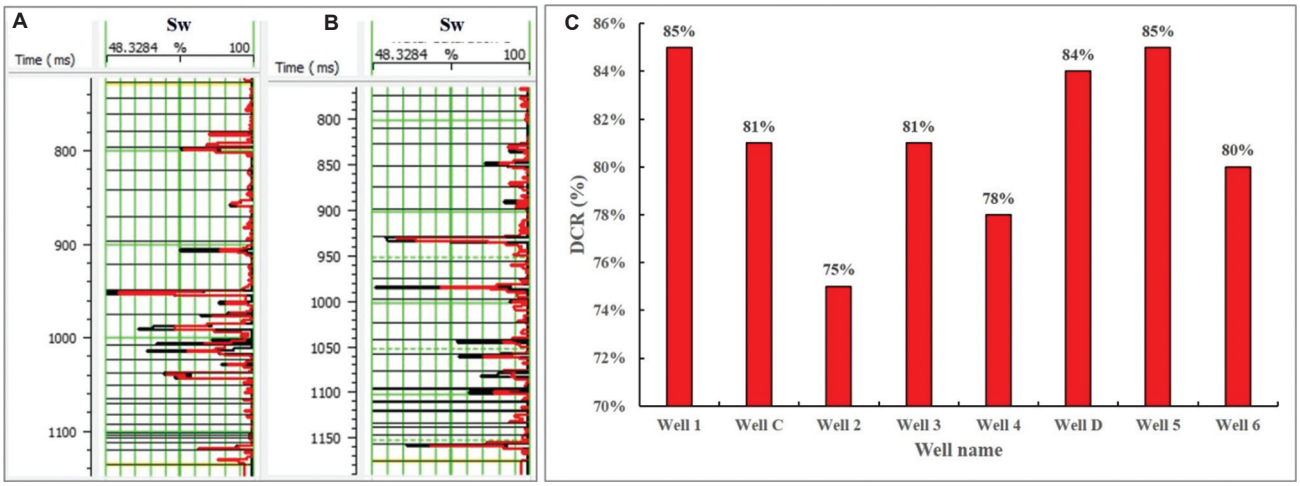

Figure 12. Training and validation results of LSTM neural network for water saturation prediction. (A) Training results of Well A. (B) Blind well testing of Well C. (C) Multi-well cross-validation DCR statistics.
Abbreviations: DSR: Drilling confirmation rate; LSTM: Long short-term memory.

0% to 60%). The prediction results demonstrate excellent agreement with well log interpretations, validating the reliability and applicability of the LSTM network in quantitative gas saturation prediction. Further analysis of horizon slice results for key well groups in the He8 Member (Figure 14) reveals that the lateral distribution characteristics of gas saturation closely align with both sand body distribution and hydrocarbon indications from well logs. The application results indicate that this method can effectively enhance the characterization accuracy of tight sandstone gas reservoirs, providing a novel technical approach for the exploration and development of similar hydrocarbon reservoirs.

4. Discussion

The study has several limitations that warrant further improvement in future work:

(i) Geostatistical inversion sensitivity: The inversion results are highly sensitive to variogram parameters and prior models. While enhancing resolution, this approach may introduce modeling artifacts. Therefore, caution must be exercised when applying geostatistical methods for resolution improvement.

(ii) Physics-aware LSTM development: Current data-driven LSTM networks lack explicit rock physics constraints. Future research should integrate rock

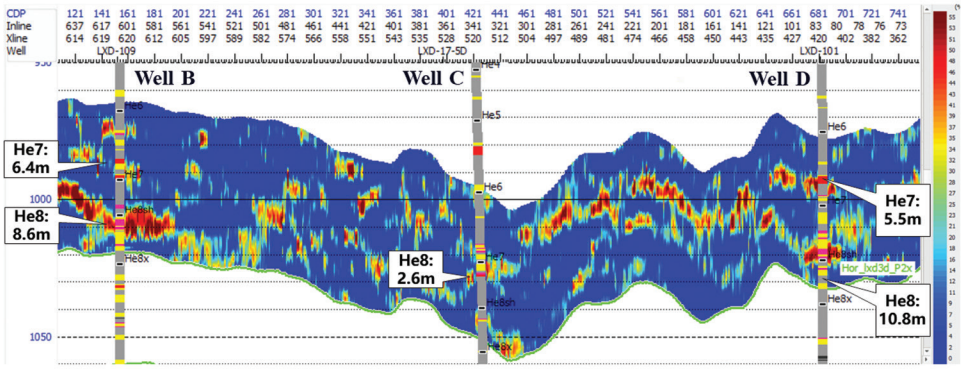

Figure 13. Well-tie profile of gas saturation prediction using LSTM neural network
Abbreviation: LSTM: Long short-term memory.

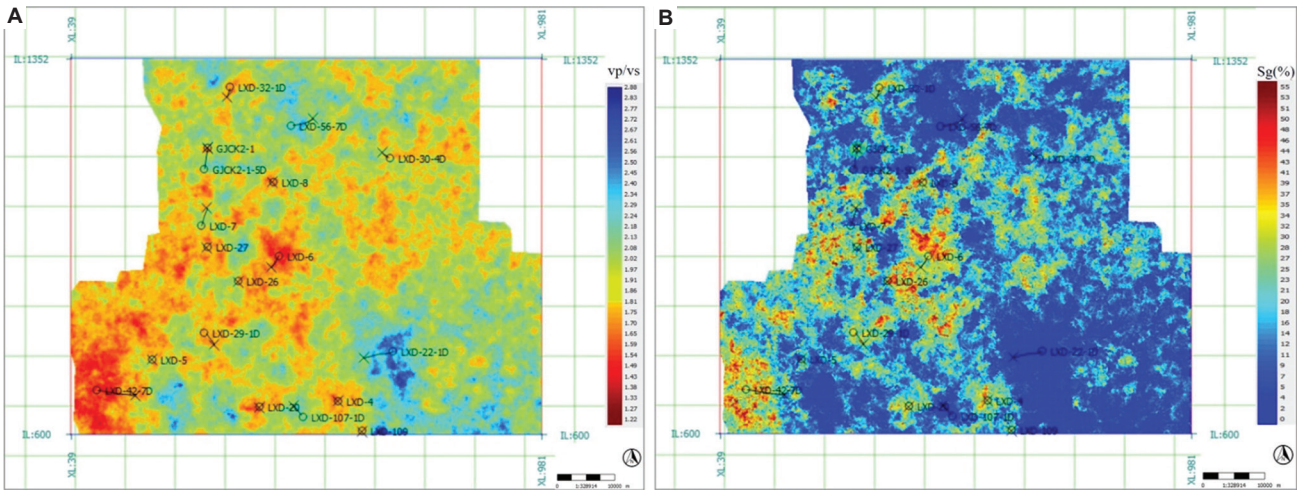

Figure 14. Inversion horizon slice of He8 member reservoir parameters. (A) Tight sand distribution. (B) Gas saturation prediction map.

Table 1. Architecture parameters of LSTM network for water saturation prediction

Parameter	Value
Dimension of the input vector X_t, t=1,2,…, n	150×15
Dimension of the LSTM output vector Y_t, t=1,2,…, n	150×32
Dimension of the FCNN output vector S_w, t=1,2,…, n	150
Number of time steps	150
Hidden units	32
Layers of LSTM	1
Layers of FCNN	1
Dimension of $W_{xf}, W_{xi}, W_{xc}, W_{xo}, W_{cf}, W_{ci}, W_{co}$	150×32
Dimension of $W_{yf}, W_{yi}, W_{yc}, W_{yo}, W_\rho$	32×1
Dimension of b_f, b_i, b_c, b_o	32
Dimension of b_ρ	1,

Abbreviations: FCNN: Fully connected neural network; LSTM: Long short-term memory.

physics models to develop physics-constrained LSTM approaches for gas-bearing property prediction.

In summary, geostatistical inversion relies heavily on prior geological knowledge, while deep learning techniques require extensive training datasets. Given the mature geological understanding and data availability during hydrocarbon development phases, this methodology is particularly suitable for late-field development stages, providing critical support for well placement optimization.

5. Conclusion

This study focuses on tight sandstone reservoirs in block L of the Ordos Basin and develops a data-driven high-resolution gas-bearing prediction technology for tight sandstones. The proposed method effectively predicts thin sand bodies and their gas-bearing potential, providing critical support for well placement optimization. The key findings are summarized as follows: first, the implementation of geostatistical methods to construct detailed 3D grid models

and statistical variograms significantly enhances seismic inversion resolution, enabling the successful identification of thin sand bodies with thicknesses below 5 m in the Ordos Basin. Second, the application of geological horizon constraints, which systematically organizes training data according to sedimentary cycles, effectively mitigates interference between data from different depositional periods and substantially improves prediction accuracy. Finally, the prediction model based on LSTM networks fully considers the time-varying characteristics of elastic parameter sequences, overcoming the limitations of traditional deep neural networks that rely on single-point predictions. When combined with attribute optimization techniques, this approach significantly improves the accuracy of seismic reservoir prediction and provides reliable guidance for well placement and optimization in tight gas blocks.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare no potential conflict of interest.

Author contributions

Conceptualization: Lixin Tian
Formal analysis: Jingxue Shi
Investigation: Qixin Li
Methodology: Shuai Sun
Writing–original draft: Shuai Sun
Writing–review & editing: Shuai Sun

Availability of data

Data are available from the corresponding author upon reasonable request.

References

1. Lan C, Zhang J, Tao W, Zhang Y, Yang M, Wang J. Sedimentary characteristics and evolution of the upper carboniferous Taiyuan formation, Shenmu Gasfield, Northeastern Ordos Basin. *Acta Geol Sin*. 2011;85(4):533-542.

2. Lan C, Zhang Y, Zhang J, Yang M, Wang J. Reservoir characteristics and controlling factors of Taiyuan Formation in Shenmu Gas Field. *J Xi'an Petrol Univ (Natural Science Edition)*. 2010;25(1):7-11.

   doi: 10.11743/ogg20130105

3. Zhu G, Li B, Li Z, Du J, Liu Y, Wu L. Practices and development trend of unconventional natural gas exploration in eastern margin of Ordos Basin: Taking Linxing-Shenfu gas field as an example. *China Offshore Oil Gas*. 2022;34(4):16-29.

   doi: 10.11935/j.issn.1673-1506.2022.04.002

4. Zhu Y, Zhao Z, Zhang D, *et al*. Accumulation conditions and accumulation laws of tight gas in Shenfu area, northeast of Ordos Basin. *China Offshore Oil Gas*. 2022;34(4):55-64.

   doi: 10.11935/j.issn.1673-1506.2022.04.005

5. Zou C, Zhang G, Yang Z, *et al.* Unconventional petroleum geology. *Petrol Explor Dev*. 2012;40(4):2153.

   doi: 10.1016/S1876-3804(13)60053-1

6. Haas A, Dubrule O. Geostatistical inversion-a sequential method of stochastic reservoir modelling constrained by seismic data. *First Break*. 1994;12(11):561-569.

   doi: 10.3997/1365-2397.1994034

7. Huang H, Zhang R, Wei S. Research on application of seismic nonlinear random inversion to reservoir prediction in the thin sandstone of continental deposits. *Acta Petrol Sin*. 2009;30(3):386-390.

   doi: 10.7623/syxb2009011

8. Liu Z, Zhang L, Huo L, Cao M, Ding Q, Gao G. Thin coalbed methane reservoir identification by geostatistics inversion. *Oil Geophys Prospect*. 2012;47(S1):30-34.

9. Archie GE. The electrical resistivity log as an aid in determining some reservoir characteristics. *Trans AIME*. 1942;146:54-62.

   doi: 10.2118/942054-G

10. Barone A, Sen MK. An improved classification method that combines feature selection with nonlinear Bayesian classification and regression: A case study on pore-fluid prediction. In: *87th Annual International Meeting, SEG, Expanded Abstracts*; 2007. p. 3062-3066.

    doi: 10.1190/segam2017-17790222.1

11. Xu J, Xu L, Qin Y. Two effective methods for calculating water saturations in shale-gas reservoirs. *Geophysics*. 2017;82(3):D187-D197.

    doi: 10.1190/geo2016-0462.1

12. Hu L, Chen M, Jin H. Gas prediction in tight sandstone reservoirs based on a seismic dispersion attribute derived from frequency-dependent AVO inversion. *Processes*. 2025;13:2210.

    doi: 10.3390/pr13072210

13. Tao X, Cao J, Zhao L, Li H, Ren Y, Jian P. Gas-bearing prediction in tight sandstone reservoirs based on multinetwork integration. *Interpretation*. 2024;12(2):T177-T185.

    doi: 10.1190/INT-2023-0091.1

14. Sun S, Nie J, Qu Z, *et al*. Oil saturation estimation and uncertainty evaluation by modeling-data-driven Gaussian

mixture conditional generative adversarial networks. In: *First International Meeting for Applied Geoscience & Energy Expanded Abstracts*; 2021. p. 1691-1695.

doi: 10.1190/segam2021-3577905.1

15. Zong Z, Yin X, Wu G. Fluid identification method based on compressional and shear modulus direct inversion. *Chin J Geophys*. 2012;55(1):284-292.

doi: 10.6038/j.issn.0001-5733.2012.01.028

16. Zong Z, Yin X. Direct inversion of Young's and Poisson impedances for fluid discrimination. *Geofluids*. 2016;16:1006-1016.

doi: 10.1111/gfl.12202

17. Jia L, Li L, Wang Q, Ma J, Wang H, Wang D. Fluid identification factor inversion based on generalized elastic impedance. *Geophys Prospect Petrol*. 2018;57(2):302-311.

doi: 10.3969/j.issn.1000-1441.2018.02.016

18. Yin X, He W, Huan X. Bayesian sequential gaussian simulation methodology. *J China Univ Petrol*. 2005;29(5):5.

19. Bergen KJ, Johnson PA, De Hoop MV, Beroza GC. Machine learning for data-driven discovery in solid Earth geoscience. *Science*. 2019;363(6433):eaau0323.

doi: 10.1126/science.aau0323

20. Feng R, Mejer Hansen T, Grana D, Balling N. An unsupervised deep-learning method for porosity estimation based on post-stack seismic data. *Geophysics*. 2020;85(6):M97-M105.

doi: 10.1190/geo2020-0121.1

21. Hampson DP, Schuelke JS, Quirein JA. Use of multiattribute transforms to predict log properties from seismic data. *Geophysics*. 2001;66(1):220-236.

doi: 10.1190/1.1444899

22. Zhong Z, Carr TR, Wu X, Wang G. Application of a convolutional neural network in permeability prediction: A case study in the Jacksonburg-Stringtown oil field, West Virginia, USA. *Geophysics*. 2019;84(6):B363-B373.

doi: 10.1190/geo2018-0588.1

23. Das V, Mukerji T. Petrophysical properties prediction from prestack seismic data using convolutional neural networks. *Geophysics*. 2020;85(5):N41-N55.

doi: 10.1190/geo2019-0650.1

24. Chen W, Yang L, Zha B, Zhang M, Chen Y. Deep learning reservoir porosity prediction based on multilayer long short-term memory network. *Geophysics*. 2020;85(4):WA213-WA225.

doi: 10.1190/geo2019-0261.1

25. Smith JR, Johnson AB, Zhang L. Seismic sensitivity to water saturation in tight gas sandstones: A rock physics perspective. *Geophysics*. 2018;83(2):MR89-MR101.

26. Zong Z, Yin X, Wu G. Elastic parameter-driven seismic characterization of gas-bearing reservoirs using hybrid inversion. *Geophysics*. 2020;85(3):WA1-WA12.

27. Buland A, Omre H. Bayesian linearized AVO inversion. *Geophysics*. 2003;68(1):185-198.

doi: 10.1190/1.1543206

28. Journal AG, Alabert F. Non-Gaussian data expansion in the Earth Sciences. *Terra Nova*. 1989;1:123-134.

doi: 10.1111/j.1365-3121.1989.tb00344.x

29. Tai KS, Socher R, Manning CD. Improved semantic representations from tree-structured long short-term memory networks. In: *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, 2015. p. 1556-1566.

doi: 10.3115/v1/P15-1150

30. Sathasivam S, Abdullah WATW. Logic learning in Hopfield networks. *Modern Applied Science*, 2008, 2(3):57.

doi: 10.5539/mas.v2n3p57

31. Nair V, Hinton GE. Rectified linear units improve restricted Boltzmann machines. In: *Proceedings of the 27th International Conference on International Conference on Machine Learning. Haifa, Israel*; Omnipress; 2010. p. 807-814.

32. Zhang L, An H, Dan G, Liang G, Gao X. Geostatistical inversion application of carbonate reservoir prediction in Lungu oilfield. *Petrol Geol Eng*. 2016;30(2):1-4.