

ARTICLE

Thin-bed reservoir characterisation using high-resolution ocean-bottom-node seismic data, seismic waveform indication inversion, and a Gaussian mixture model: A case study from the Guantao Formation, Bohai Bay Basin, China

Yanjia Wu¹, Zhiwen Dong¹, Yiyao Zhang¹, Zekun Li¹, Shihao Liu¹, Xiaoxiao Shi¹, Hongke Zhou¹, Qunhu Wu¹, Shina Yu¹, Ahmed Bakry^{2*}, and Qamar Yasin^{2*}¹Offshore Production Plant, Sinopec Shengli Oilfield Company, Dongying, Shandong, China²National Key Laboratory of Continental Shale Oil, Northeast Petroleum University, Daqing, Heilongjiang, China***Corresponding author:**Qamar Yasin
(qyasin@nepu.edu.cn)

Citation: Wu Y, Dong Z, Zhang Y, *et al.* Thin-bed reservoir characterisation using high-resolution ocean-bottom-node seismic data, seismic waveform indication inversion, and a Gaussian mixture model: A case study from the Guantao Formation, Bohai Bay Basin, China. *J Seismic Explor.* 2026;35(3):026060026.
doi: 10.36922/JSE026060026

Received: February 6, 2026**Revised:** March 25, 2026**Accepted:** March 31, 2026**Published online:** May 13, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract

Targeting reservoirs that fall below seismic resolution remains a primary challenge in reservoir characterisation. High-resolution ocean-bottom-node seismic data are essential for imaging oil-bearing thin-sand facies, particularly in the Guantao Formation of the Bohai Bay Basin. The basin is rich in hydrocarbon resources but underexplored due to the difficulty of identifying thin-sand bodies. This study investigated thin beds below the tuning thickness to determine reservoir quality and oil–water distribution within this formation. By applying seismic waveform indication inversion (SWII) at the reservoir level, we generated a high-frequency and high-resolution facies model validated against existing geological data. Simultaneously, a Gaussian mixture model (GMM) was utilised to detect anomalies in well-log data (gamma ray, velocity ratio [Vp/Vs], and density), leveraging inter-variable relationships to enhance geological interpretability. Results demonstrated an excellent correlation between GMM anomalies and high-resolution seismic data, significantly improving prediction accuracy in complex structural regimes. SWII effectively identified thin sand-shale layers, reflecting various stacking patterns and the effects of porosity and oil-bearing properties on sand-body velocity. Forward modelling analysis revealed the seismic-resolvable thicknesses of sand bodies, the seismic response characteristics associated with various stacking patterns, and the influence of physical properties and oil-bearing characteristics on sand-body velocity, among other factors. This integrated approach effectively resolves thin-bed distributions and identifies potential hydrocarbon traps, providing a robust foundation for future multilayer system deployment in the Bohai Bay Basin.

Keywords: Anomaly detection; Gaussian mixture model; Thin-bed reservoir characterisation; Well-log analysis, Seismic data inversion

1. Introduction

Well-log acquisition serves as the cornerstone of subsurface geological characterisation, providing high-resolution, in situ measurements of formation physical properties. The integration of nuclear logs (gamma ray [GR], bulk density [DEN], and neutron porosity [CNL]), acoustic logs (downhole travel-time of P-wave [DTP] and downhole travel-time of S-wave [DTS]), and electrical logs (deep resistivity) enables precise lithological discrimination and the quantification of storage capacity.^{1,2} For example, cross-validation of density and CNL is essential for fluid typing, while the DTP and DTS ratios are critical for delineating lithofacies classification. Consequently, the rigorous interpretation of these multi-physical datasets is a prerequisite for efficient reservoir management, serving as the basis for calculating reserves and optimising drilling trajectories.³ Moreover, advanced anomaly detection in these datasets is essential for preserving data integrity. Automated outlier identification enables effective reservoir modelling and facilitates early detection of geomechanical risks, such as overpressure zones and borehole instability.⁴⁻⁶

Gaussian mixture models (GMMs) have emerged as powerful tools for anomaly detection due to their capacity to model complex, multi-modal data distributions.⁷ As probabilistic models, GMMs assume data points are generated from a mixture of Gaussian distributions, allowing them to approximate any continuous probability density function with high precision, provided the components (means, covariances, and weights) are properly optimised.⁸ Unlike deterministic clustering approaches, GMMs provide probability-based membership estimates, enabling quantification of uncertainty and identification of low-probability observations that may correspond to subtle lithological variations, such as thin-sand interbeds.

While well logs and core data provide vertical detail, seismic data inversion is pivotal for understanding the lateral distribution of anomaly zones and identifying hydrocarbon-bearing formations.⁹⁻¹¹ As global energy demand grows, advanced inversion techniques are increasingly necessary to optimise well placement and minimise environmental impact.¹²⁻¹⁴ By providing detailed information on reservoir properties, such as reservoir facies distribution, porosity, permeability, and fluid saturation, seismic inversion significantly reduces uncertainties associated with oil and gas reservoir characterisation and management.¹⁵

However, traditional methods face significant challenges, e.g., P-wave impedance inversion using post-stack data often suffers from the “tuning effect,” where the lack of direct correspondence between seismic signals and thin beds prevents effective lateral prediction.¹⁶⁻¹⁸

While geostatistical inversion integrates diverse datasets to improve vertical resolution, its accuracy often falters in the high-frequency band, leading to inconsistencies with sedimentary models.¹⁹⁻²² However, this method's prediction accuracy is reduced in the high-frequency band, and the lateral prediction accuracy of the inversion results is often inconsistent with the sedimentary model.^{16,23,24}

To overcome these limitations, the seismic waveform indication inversion (SWII) method was developed. By utilising the lateral change information inherent in seismic waveforms, this method ensures reservoir predictions are less dependent on well distribution. This approach establishes an interpolation model aligned with geological and sedimentary characteristics, making it uniquely suited for high-precision predictions in areas characterised by fast lateral changes, strong heterogeneity, and thin interbeds.^{23,25,26}

Recent literature supports the efficacy of such advanced inversion workflows, with studies demonstrating improved accuracy in complex lithological environments and tight sandstone reservoirs. Helgesen *et al.*²⁷ found that the sparse pulse inversion method provides more accurate and detailed reservoir property predictions in a complex reservoir than traditional inversion methods. Li *et al.*²⁸ showed that the SWII method can accurately predict the spatial distribution of reservoir properties in a tight sandstone gas reservoir with complex geological structures. SWII establishes a relationship between seismic waveforms and high-frequency well logs by dynamically clustering seismic waveforms. This process aims to enhance the vertical and horizontal resolution of the inversion results, providing a more detailed and accurate depiction of subsurface geological features. Nizarul *et al.*²⁹ introduced the short-time Fourier transform half cepstrum method for enhancing thin-bed identification in the Sarawak Basin, emphasising its significance for improved hydrocarbon exploration and seismic interpretation. Babasafari *et al.*³⁰ proposed an approach for petro-elastic modelling that significantly improves the accuracy of elastic property estimation by simplifying the modelling process and optimising parameters. This ultimately enhances predictions and supports seismic inversion in carbonate reservoirs in Central Luconia, offshore Sarawak, Malaysia. Fateh *et al.*³¹ introduced an azimuthal deep resistivity technology and customised workflows for proactive geosteering in thin, heterogeneous carbonate reservoirs. Purnomo *et al.*³² introduced an improvement to the elastic inversion technique that allows the extraction of petrophysical characteristics from low-amplitude seismic data. This enhancement was implemented to calculate petrophysical parameters in a gas field characterised by

thin sand-shale layers in the Sarawak Basin, Malaysia. The study also demonstrated that the SWII method can help identify potential bypassed pay zones and optimise well placement and completion design.

Despite these advances, previous studies typically treat seismic inversion, rock physics, and machine learning techniques as parallel workflows rather than as a unified probabilistic framework. In many cases, machine learning methods are applied for supervised facies classification without explicitly constraining seismic inversion. In contrast, inversion studies often rely on deterministic impedance thresholds without incorporating statistical anomaly structure derived from multi-parameter log space. Consequently, a clear probabilistic linkage between anomaly detection in well-log data and waveform-controlled seismic inversion remains insufficiently explored—particularly for ultra-thin interbeds below seismic tuning thickness.

This study introduces a probabilistically constrained thin-bed characterisation framework that integrates unsupervised GMM, rock physics validation, forward seismic modelling, and SWII into a unified interpretive workflow. Unlike conventional approaches that treat anomaly detection and inversion as independent tasks, the proposed methodology uses GMM to quantify anomaly likelihood in multi-dimensional log space (GR, velocity ratio [V_p/V_s], DEN), identifies statistically significant thin-bed signatures, and cross-validates these signatures using rock-physics relationships and forward modelling. The inversion results are then evaluated within this probabilistic and physically consistent framework, thereby reducing interpretational ambiguity and enhancing lateral prediction reliability.

The primary methodological contributions of this study are: (i) the integration of probability-based anomaly quantification with waveform-controlled inversion; (ii) the explicit coupling of statistical log-space anomalies with rock physics consistent seismic responses; and (iii) the demonstration of enhanced thin-bed detectability (<10 m) using high-frequency ocean-bottom-node (OBN) data within a structurally complex fluvial system.

Building upon this framework, we applied the integrated methodology to the Guantao Formation in the Bohai Bay Basin. By merging probabilistic anomaly detection with SWII, we generated high-frequency and high-resolution imaging of thin-sand reservoirs in a tectonically complex setting. This approach not only improves thin-bed discrimination below the tuning thickness but also establishes a statistically robust, physically grounded workflow for reservoir characterisation in heterogeneous fluvial systems. The findings expand understanding

of the Guantao Formation and provide a transferable methodology for thin-bed exploration in similar geological environments worldwide.

2. Geological overview

The Guantao Formation is a Miocene-age stratigraphic unit that forms a foundational component of the Neogene succession within the Bohai Bay Basin (Figure 1A), one of China's most prolific hydrocarbon regions.³³ Its deposition occurred during the regional post-rift thermal subsidence phase that followed the intense Paleogene extensional rifting. This shift to a more stable tectonic regime led to widespread, blanket-like deposition of the Guantao Formation across the previously fractured basin architecture. It unconformably overlies the older Paleogene formations.

Lithologically, the Guantao Formation is characterised by coarse-grained, clastic sediments with a distinct vertical evolution. The lower section typically consists of thick, poorly sorted conglomerates and coarse sandstones deposited by high-energy braided river systems, which grade upward into finer-grained sandstones and interbedded mudstones deposited in meandering-river and flood-plain environments.³⁴ This fining-upward sequence reflects the gradual decrease in topographic relief and sediment supply as the basin continued to subside.

The primary significance of the Guantao Formation lies in its exceptional reservoir characteristics, making it a major exploration target for both conventional and heavy oil accumulations.³³ The coarse-grained, clean sandstones of the lower braided river channels exhibit high primary porosity and permeability, resulting in excellent reservoir quality (Figure 1B). These rocks play a crucial role in a highly effective petroleum system. The hydrocarbons generated from the underlying, organic-rich Palaeogene Shahejie Formation source rocks migrate vertically along faults and laterally along the regional unconformity at the base of the Guantao to charge these reservoirs. Trapping mechanisms are varied, including gentle anticlines and, most notably, stratigraphic and lithologic traps such as sand bodies pinching out against the basal unconformity or “buried hill” structures of the pre-Cenozoic basement. The widespread distribution and consistent reservoir properties of the Guantao Formation have therefore cemented its status as a key hydrocarbon-producing formation in the Bohai Bay Basin.

Figure 1 illustrates the regional geological framework and the location of the Bohai Bay Basin along China's eastern coast, within the broader North China Craton. The basin's proximity to the Yellow Sea and Bohai Sea provides the structural context for its evolution as a major

intracontinental rift system and hydrocarbon province. A regional summary structural map (Figure 1B) reveals the basin's characteristic graben-horst architecture developed during Cenozoic extension. The Guantao Formation (highlighted as dark brown in Figure 1A) unconformably overlies the fault-controlled Paleogene sequence, demonstrating the transition from syn-rift to post-rift sedimentation.

3. Methodology and data preparation

An extensive dataset of borehole well logs was compiled from six actively producing wells, accompanied by 3D OBN seismic data in the study area (Figure 1). The investigation encompassed a thorough analysis of key logging measurements, including calliper (CAL), GR, DTP, DTS, DEN, and CNL. This dataset comprises 10,000 measurements of GR (API), DEN (g/cm^3), and V_p/V_s over a depth range of 1,200 to 2,400 m (the upper and lower Guantao Formation). Two wells were specifically chosen for model training due to their high data quality and complete log coverage.

Before clustering and probabilistic modelling, several preprocessing steps were applied to ensure statistical consistency. Missing values were imputed using the mean. All variables were subsequently standardised using

z-scores so that each feature contributed equally to the distance metrics and likelihood calculations. An initial unsupervised partitioning of the standardised dataset was then carried out using K-means clustering with 10 clusters ($n_{\text{init}} = 10$; random seed = 42). Only the two clusters that displayed a clear lithological separation between sand-prone and shale-prone responses were retained for subsequent anomaly detection and GMM analysis.

3.1. Unsupervised anomaly detection

To uncover geologically meaningful patterns, the data were clustered using the K-means algorithm following standardisation. These clusters align with known lithofacies (sand and mudstone) and present a petrophysically diverse yet interpretable subset for anomaly detection.

Following the initial clustering phase, the isolation forest (IF) algorithm was deployed to identify potential anomalies within the selected lithological clusters. Unlike distance-based or density-based methods, IF is an ensemble learning approach that isolates anomalies by recursively partitioning the feature space into random binary trees.³⁵ The fundamental idea is that anomalous observations, being rare and distinct, are more easily isolated and therefore exhibit shorter average path lengths, whereas normal samples require more partitions.

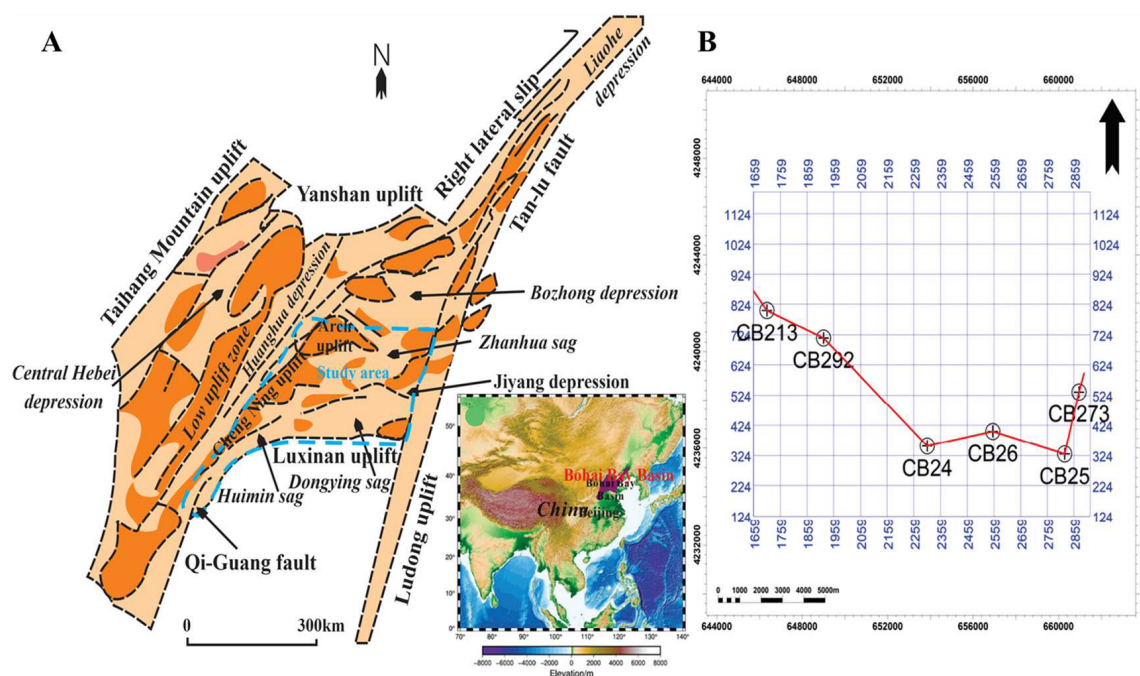


Figure 1. Regional geological framework of the study area. (A) Location of the Bohai Bay Basin. (B) Seismic and well location map displaying the spatial distribution of study wells across major structural elements.

In this study, the filtered dataset was divided into training (70%) and testing (30%) subsets using a fixed random seed (42) to ensure reproducibility. The IF model was built with 100 trees (the default in scikit-learn) and a contamination factor of 0.10, indicating that approximately 10% of the samples were expected to be anomalous. The algorithm outputs binary labels, where -1 indicates an anomaly and $+1$ denotes normal behaviour. It is important to stress that IF predictions were not treated as ground truth; rather, they provided an objective statistical reference for comparison with the probability-based anomalies derived from the GMM.

Figure 2 visualises the integration of clustering and anomaly detection across the GR, DEN, and Vp/Vs logs. The subplots contrast the measurements from both wells, delineating the distribution of training and testing samples. Notably, the visualisation highlights the distinct petrophysical separation between clusters 0 and 1, while revealing the stratigraphic and spatial distribution of the anomalies flagged by the algorithm.

3.2. Gaussian mixture model configuration

Data were collected from six wells that penetrated the Guantao Formation, which consists of mudstones and sandstones as the major lithologies, with minor conglomerates. Rigorous quality control procedures were implemented to minimise noise and bias, including outlier detection using both the three-sigma rule and interquartile range analysis, and verification of borehole conditions via CAL logs.

The GMM was configured with customised hyperparameters to optimise its performance. The GMM implementation followed the scikit-learn framework. A mixture of two Gaussian components was selected to represent the primary lithological populations in the feature space. Full covariance matrices were adopted to allow each component to capture anisotropic variance

and inter-variable correlations. Parameter estimation was performed using the expectation–maximisation algorithm with 10 independent initialisations ($n_{\text{init}} = 10$), and the solution with the highest log-likelihood was retained. The convergence tolerance and maximum number of iterations were kept at the default values of 10^{-3} and 100, respectively. A fixed random seed was applied to ensure reproducibility of the optimisation. To maintain stable geological interpretation across runs, the resulting mixture components were sorted by the ascending mean of the first feature dimension. This ordering prevents label switching and enables consistent comparison among wells. To ensure comparability across wells and lithologies, all features were standardised using z-score normalisation (Equation 1).

$$X_{\text{norm}} = \frac{X - \mu}{\sigma} \quad (1)$$

where x denotes an individual data point, μ denotes the sample mean of the input dataset, and σ denotes the sample standard deviation.

After training, each observation was evaluated against the learned probability density function. Samples located within the high-density core of the distributions were classified as background lithology, whereas points falling in the low-probability tails were considered anomaly candidates. Operationally, this corresponds to observations lying outside the three-sigma confidence envelope of the Gaussian components.

3.3. Seismic waveform indication inversion framework

The SWII method is a geostatistics-based approach that combines well logging, geological, and seismic data to establish probability models for reservoir attributes.^{23,27} Unlike conventional inversion, which relies on existing well configurations to initialise models, SWII focuses on lateral seismic waveform indicators derived from well data.

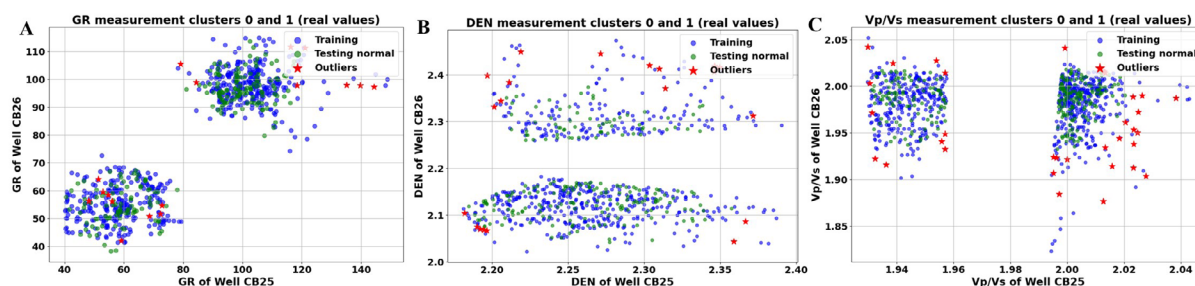


Figure 2. Clustering results and isolation forest anomaly detection for (A) gamma ray (GR), (B) density (DEN), and (C) velocity ratio (Vp/Vs) logs in wells CB25 and CB26. Red points indicate detected anomalies, highlighting their distribution relative to the main litho-fluid clusters.

It utilises a stochastic function model to obtain multiple high-resolution realisations aligned with both well and seismic data, inferring spatial distribution characteristics with high vertical resolution.

The objective function for waveform inversion is represented as (Equation 2):

$$E_{(m)} = \min_m (d - F_m)_2^2 \quad (2)$$

where d represents the observed seismic data (waveforms), and F is the forward modelling operator that generates synthetic data from the model m , which is the vector of subsurface properties (e.g., velocity and DEN) we aimed to

estimate. The objective of the $E(m)$ function is to minimise the misfit between the observed and synthetic data, ultimately refining the model parameters m .

The SWII utilises a deterministic method to derive impedance volumes, followed by the Markov Chain Monte Carlo algorithm to generate interwell impedance. The resulting impedance is converted into reflection coefficients and convolved with wavelets to produce synthetic traces. By comparing these synthetic traces to the original seismic data, we identified sandstone and shale layers through impedance matches. This process integrates information across multiple scales to achieve accurate inversion results that align with the inherent phase-contrast capabilities of seismic data (Figure 3).

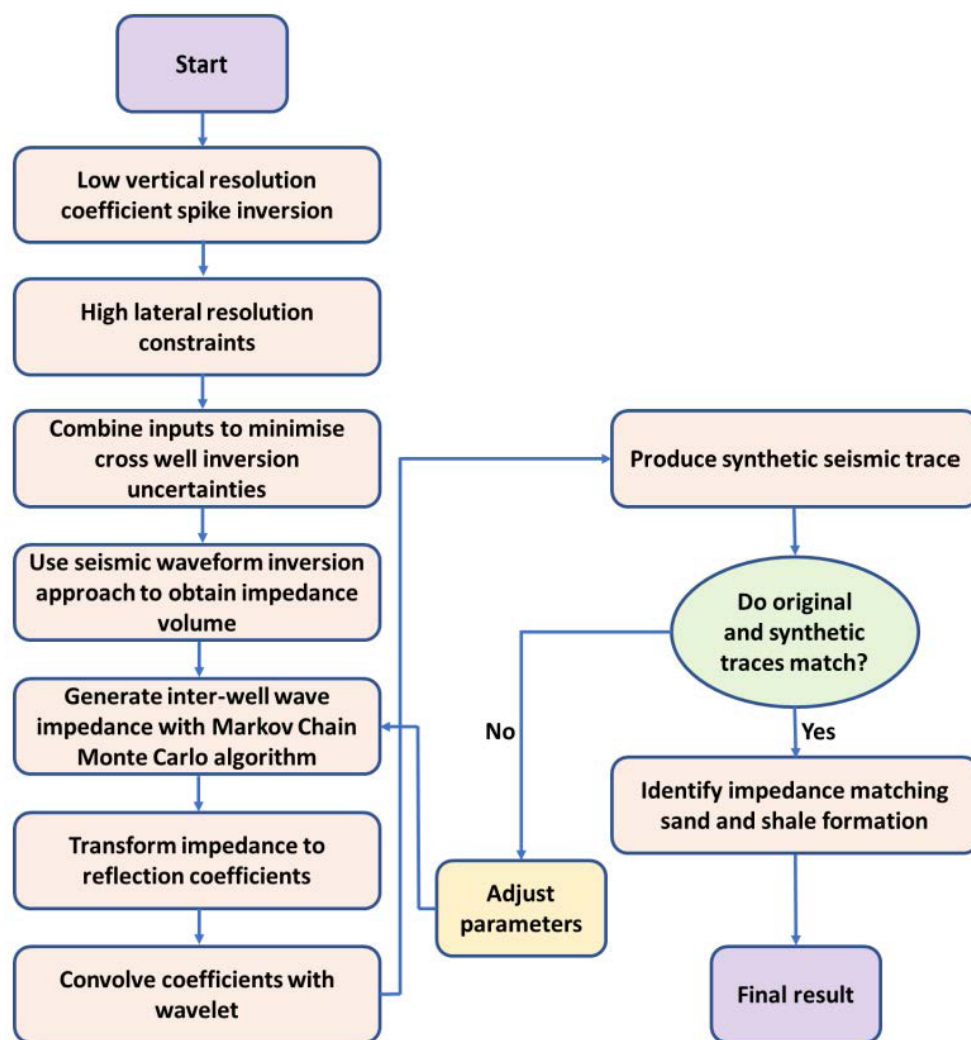


Figure 3. Simplified workflow illustrating the seismic inversion process used to construct a geological model containing thin interbeds

3.4. The challenges of seismic inversion in the Guantao Formation

After an extensive and detailed analysis of the reservoir's essential characteristics through the use of GMM, it became evident that applying the post-stack elastic parameter inversion technique to the Guantao Formation across the CB24, CB25, CB26, CB273, CB292, and CB213 wells facilitates an accurate distinction between sandstone and shale layers.

As illustrated in Figure 4, the V_p/V_s and Poisson's ratios derived from the sparse spike inversion method in the Guantao Formation revealed a consistent pattern. Notably, low V_p/V_s and low Poisson's ratios indicate the presence of sandstone beds. However, it is important to note the challenges presented by seismic data in effectively detecting all the thin interbedded sand layers. The variability in seismic waveforms, caused by the diverse configurations of these sand bodies, poses significant challenges to obtaining a clear and definitive identification in the Guantao Formation.

It is important to emphasise that over 80% of the individual sand layers within the Guantao Formation's

targeted strata tend to be thin, with thicknesses measuring less than 10 m, as depicted in Figure 5. This illustration plays a crucial role in analysing the difficulties encountered when trying to differentiate these complex and interbedded sand layers within the formation. The challenge is intensified by the predominantly thin profile of these sand layers, combined with the distinctive characteristics of the seismic data.

In high-resolution OBN seismic data, sand bodies exceeding 20 m in thickness are particularly notable. They are distinguished by complex waveforms and strong amplitudes that unmistakably signal their presence. Conversely, sand bodies that measure between 10 and 15 m are characterised by more streamlined waveforms and lower amplitudes. Meanwhile, smaller sand bodies measuring 5 m or less are often undetected, becoming nearly invisible in seismic data.

Thoroughly exploring the complexity beneath the tuning thickness distribution is essential to utilise SWII. This sophisticated technique unveils the hidden layers of thin sand with remarkable accuracy and insight, allowing us to uncover the complexities of these subtle geological features.

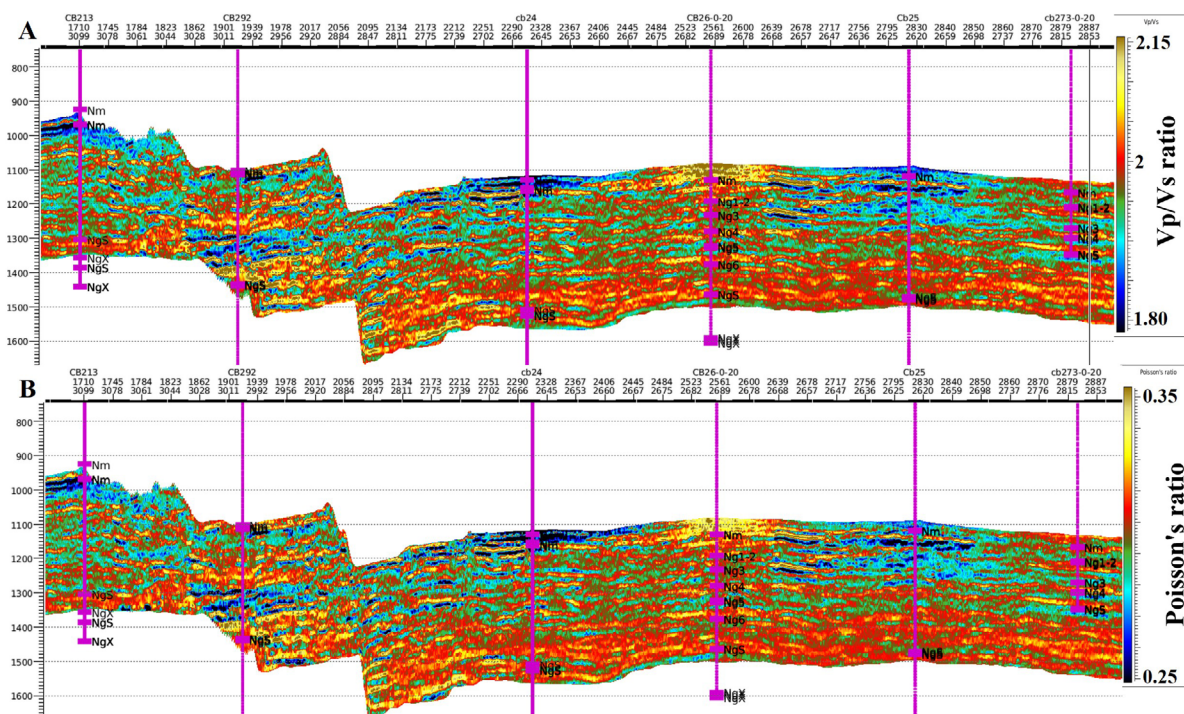


Figure 4. Seismic interpretation of the Guantao Formation. (A) Inverted velocity ratio (V_p/V_s) and (B) Poisson's ratio seismic sections of the Guantao Formation, showing consistent low-value anomalies associated with sandstone layers (for cross-section analysis, refer to Figure 1).

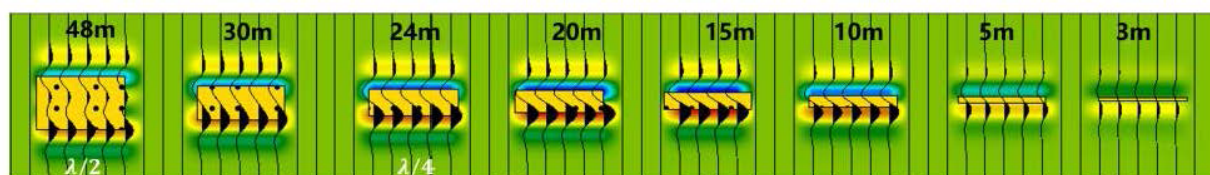


Figure 5. Thickness distribution of individual sand layers within the Guantao Formation

4. Results and discussion

4.1. Gaussian mixture model analysis of thin beds

Figure 6 illustrates the application of GMM to GR, Vp/Vs, and DEN logs. The first row shows the initial anomaly detection results from the IF algorithm. The IF method effectively delineated outliers (marked in red) from the nominal data clusters, revealing that anomalies are predominantly concentrated at the periphery of the data distribution. This step established a baseline for distinguishing between standard lithological responses and potential thin-bed occurrences.

The second row depicts the unsupervised training phase of the GMM. Because the dataset was unlabelled, the model approximated the underlying structure of the log data by fitting two distinct Gaussian components to the clusters. These components, visualised as ellipses, represented the high-probability regions of the dataset. By optimising parameters to fit these components, the GMM captured the intrinsic variance of the normal formation data, creating a probabilistic reference model against which deviations could be measured.

The third row presents the testing stage, visualising the model's predictions against the probability density functions of the learned distributions. Contour lines indicate the model's confidence levels; data points in high-density regions are classified as normal (blue), while those in lower-density regions that fall below a defined probability threshold are flagged as anomalies (red). Significantly, the concentration of anomalies within these low-probability tails correlated with the expected log responses of thin sand beds. This alignment confirms the GMM's efficacy in isolating subtle lithological features that deviate from the background trend.

To further clarify how the probabilistic structure governs the classification, the maximum-posterior cluster memberships were projected back into the original measurement space. The resulting Gaussian components formed compact, clearly separated populations, while their covariance matrices define ellipsoidal envelopes whose orientations reflect inter-variable dependence rather than isotropic spread. The majority of background samples fell

within the one- to two-sigma regions, demonstrating that the mixture adequately represents the dominant lithofacies.

In contrast, predicted anomalies were preferentially located near the margins of the distributions or beyond the three-sigma confidence limits, where the probability density decreased sharply. This geometric relationship provides a transparent explanation for the link between statistical rarity and geological distinctiveness: intervals that are difficult to represent with the learned background model correspond to thin, heterogeneous, or fluid-affected sand bodies.

The stability of the separation observed in Figure 6 further indicates that a two-component mixture is sufficient for the investigated interval. Tests with additional components yielded only marginal improvements in probability while reducing geological interpretability and were therefore not adopted.

4.2. Rock physics analysis

4.2.1. Fluid discrimination

Accurate reservoir characterisation and thin-bed detection rely on the identification of distinct fluid- and lithology-dependent signatures within well-log data. The porosity-resistivity crossplot, commonly referred to as the Pickett plot, provides a fundamental diagnostic framework based on the Archie equation, which relates porosity to resistivity and fluid saturation. In this study, the clean water-saturated trend (the “clean sand line”) established a baseline at a water saturation of 1, corresponding to fully saturated formations. Data points clustered along this line exhibited a predictable pattern: resistivity increased as porosity decreased, thereby calibrating the model to local reservoir conditions (Figure 7A).

Hydrocarbon-bearing zones deviated significantly from this water-saturated baseline due to the distinctively high resistivity of hydrocarbons relative to water.³⁶ On the Pickett plot, these zones appeared as clusters shifted toward higher resistivity values and lower water saturation (indicated by yellow points), representing the primary targets for completion and evaluation (Figure 7B). The analysis also delineated geologically complex formations, such as shaly sands and tight zones. Shaly sands, which

contain conductive clay minerals, can complicate fluid discrimination by occupying the region between the clean sand line and hydrocarbon clusters. Conversely, tight zones are characterised by low porosity and elevated resistivity; however, they are considered non-economic due to insufficient storage capacity and are typically bypassed during production planning.

4.2.2. Sensitivity analysis and crossplot validation

A comprehensive sensitivity analysis revealed the hierarchical importance of reservoir parameters in distinguishing sand from clay. The crossplot framework leveraged the differential sensitivity of elastic parameters to pore-fluid properties, utilising the relationship between DEN and Vp/Vs (Figure 8). Oil-saturated intervals (yellow clusters; DEN: 2,350–2,370 g/cm³, Vp/Vs: 1.90–2.0) displayed intermediate properties, reflecting high-porosity

(> 15%) sands. In contrast, water-saturated rocks (green clusters; DEN: 2,400–2,500 g/cm³, Vp/Vs: 2.00–2.05) maintained high elastic parameter values, consistent with the behaviour of incompressible pore fluids. Colour-coding by oil saturation revealed additional control mechanisms: zones with higher oil saturation exhibited enhanced fluid sensitivity due to increased pore-volume fractions, whereas tight rocks showed reduced discrimination capability.

The consistent clustering patterns observed across wells CB25 and CB26 validate the predictions for pore-fluid type and lithology discrimination (Figure 8A and 8B). Zones where fluid types overlap corresponded to intervals with mixed saturation (oil and water). Furthermore, production data from these wells confirms the presence of the identified water- and/or oil-saturated reservoirs. This quantitative framework enables probabilistic fluid

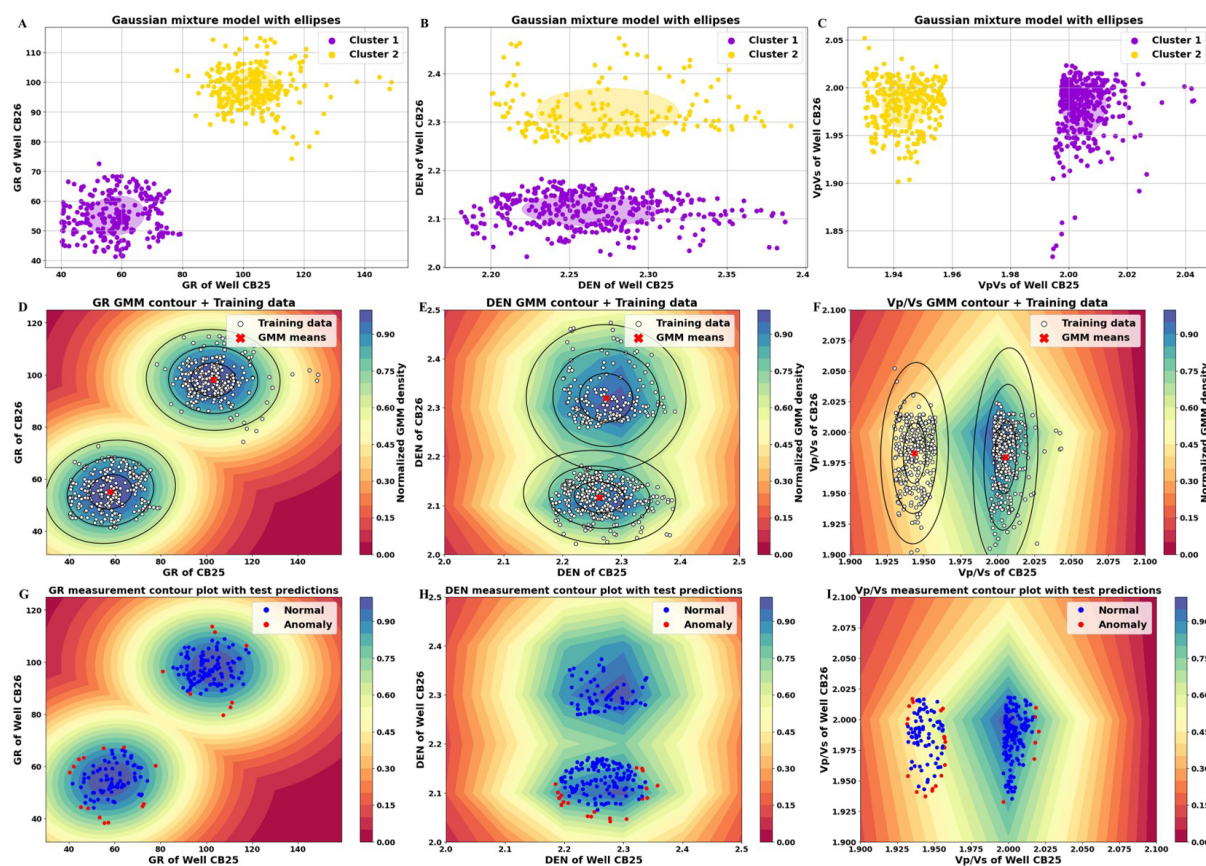


Figure 6. Unsupervised lithofacies clustering and anomaly detection using a Gaussian mixture model (GMM) applied to gamma ray (GR), velocity ratio (Vp/Vs), and density (DEN) logs. (A–C) Results of the isolation forest algorithm for (A) GR, (B) DEN, and (C) Vp/Vs, identifying initial statistical outliers (red) versus normal clusters. (D–F) GMM training phase showing the unsupervised fitting of Gaussian components (ellipses) to the data structure for (D) GR, (E) DEN, and (F) Vp/Vs. (G–I) Probability density function contours from the testing phase for (G) GR, (H) DEN, and (I) Vp/Vs. Points located in high-density regions are classified as normal background lithology (blue), while points falling below the probability threshold in low-density regions are flagged as anomalies (red), corresponding to thin sand beds.

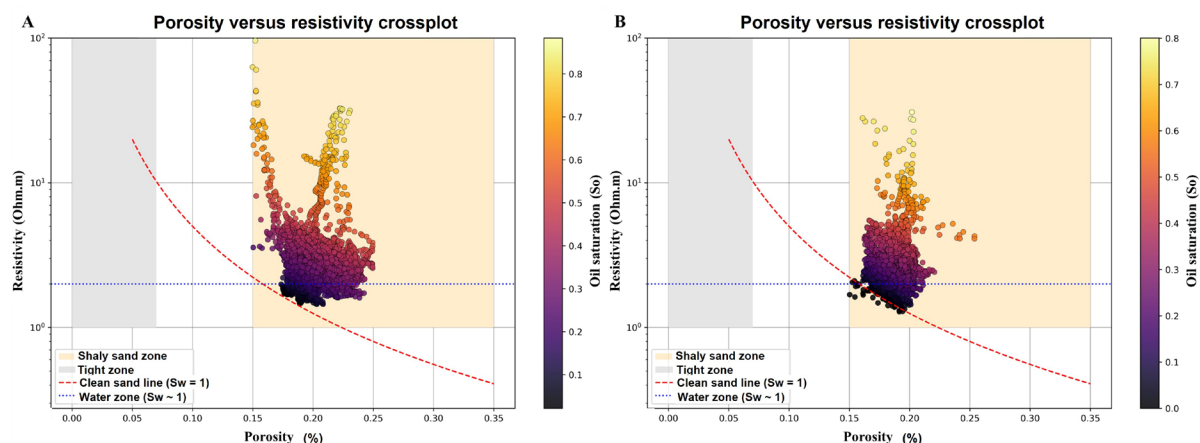


Figure 7. Porosity–resistivity (Pickett) crossplots used for fluid and lithology discrimination (CB25 and CB26; see Figure 1). (A) Establishment of the clean water-saturated baseline (water saturation $[S_w] = 1$), showing the inverse relationship between porosity and resistivity in water-bearing zones. (B) Identification of hydrocarbon potential. Hydrocarbon-bearing zones (yellow points) deviate significantly from the water line toward higher resistivities. The plot also distinguishes non-economic tight zones (low porosity and high resistivity) and shaly sands (intermediate resistivity) from the primary reservoir targets.

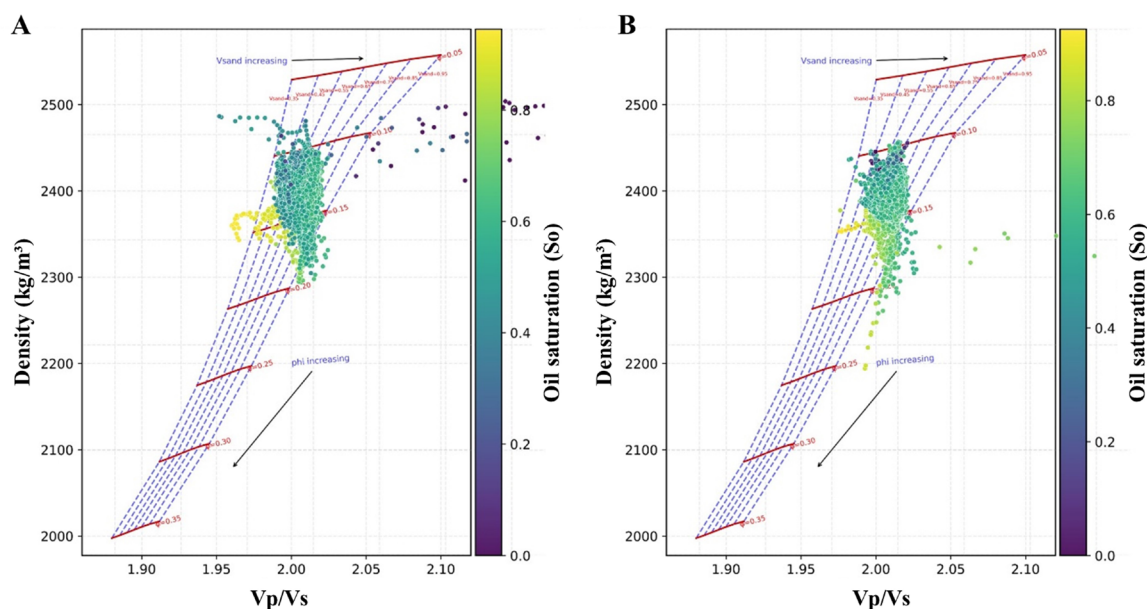


Figure 8. Crossplots of density versus velocity ratio (V_p/V_s) used for fluid and lithology discrimination in wells (A) CB25 and (B) CB26. The plots reveal distinct clustering based on pore fluid content. Oil-saturated sands (yellow clusters) are characterised by lower density and V_p/V_s values (1.90–2.0) compared to water-saturated zones (green clusters), which exhibit higher elastic parameters due to fluid incompressibility. Overlap regions between clusters indicate intervals of mixed fluid saturation.

prediction and supports the economic evaluation of potential well-completion intervals.

4.3. Forward seismic modelling

To evaluate the validity of the SWII approach and its effectiveness at identifying thin interbeds, we generated

synthetic seismic sections for different geological reservoir scenarios. Based on well profile analysis and existing literature on stacking patterns in the Guantao Formation, sedimentary facies were classified into channel sands, mudstones, and point bar sands, with thicknesses ranging from 3 to 50 m. The background velocities were defined as

2,720 m/s for mudstone and 2,670 m/s for sandstone.

Figure 9 shows the geological interpretations of seismic responses derived from forward seismic modelling. The orange colour suggests a mixed lithology or perhaps a facies change, while the yellow represents “cleaner” sand. Forward modelling was performed for various sand-body superposition patterns. Channel sand bodies were characterised by large-scale and regular morphology, reflecting the migration and amalgamation of river channels. In contrast, point bar sand bodies were smaller, isolated lens-shaped features resulting from the lateral accretion and abandonment of meandering rivers. Both facies consist of sandstone interbedded with flood-plain or abandoned-channel mudstones. These distinct lithological contrasts provide the physical basis for generating seismic reflections.

Different stacking patterns altered the interference of seismic waves, causing observable changes in attributes such as waveform symmetry, amplitude, and frequency. These variations serve as diagnostic criteria for distinguishing sand body patterns. Forward modelling at a dominant frequency of 120 Hz proved effective for identifying thin interbedded sands, as significant waveform variations were observed across different combinations of channel and point-bar sands (Figure 9).

Subsequently, the forward seismic modelling was performed for porosity, oil saturation, and velocity. The study accounted for the influence of physical properties on seismic response by modelling varying porosity and saturation levels. The forward seismic modelling results at 120 Hz are presented in Figures 10 and 11.

We observed that as porosity increased, sandstone velocity decreased. While low-porosity sandstone might exhibit velocities higher than mudstone, high-porosity sandstone dropped below the mudstone velocity. This transition creates a distinctive velocity contrast that is visually resolvable in the amplitude response (Figure 10).

Conversely, sandstone velocity decreased monotonically with increasing oil saturation. This amplifies the velocity difference relative to the surrounding mudstone, resulting in increased seismic amplitude (Figure 11).

In the study area, the baseline velocity of sandstone was slightly lower than that of mudstone. High porosity (25–35%) further reduced formation velocity and DEN, thereby enhancing impedance contrast and amplitude response. Oil-bearing sandstone exacerbated this velocity reduction, creating a strong impedance contrast with the surrounding mudstone, often resulting in “bright spot” reflections, as shown in Figure 12. Additionally, the dominant frequency tended to shift downward due to velocity attenuation.

Thus, the patterns derived from forward modelling are consistent with the well-log responses and rock physics analysis shown in Figure 8.

4.4. Seismic inversion and 3D visualisation

Previous studies have shown that SWII is highly effective for detecting thin interbedded sand layers. Specifically, the inverted acoustic impedance provides a robust tool for identifying oil-bearing sand layers within the target distribution. Figure 13 shows the results of waveform indication inversion of P-wave impedance. The lower P-impedance (pink and blue) revealed a remarkable capability to accurately identify thin sand layers in reservoirs to a depth of 3 m.

The SWII method offers notably superior vertical resolution, adapts well to lateral variations, and provides robust, precise prognostication.

P-impedance is an effective tool for identifying sand and shale and for locating the oil layer within a specific distribution range in the Guantao Formation. SWII was applied for lithofacies characterisation because the algorithm can better identify facies changes given the similarities and variations in these waveforms. Figure 14 visualises the spatial distribution and resource potential of these sand bodies. Overall, the results indicate significant potential for oil extraction in this area, suggesting that targeted exploration and drilling could yield promising results. These findings confirm that the inversion method is reliable for predicting the location and properties of subsurface formations.

The success of this approach lies in the methodology; waveform indication inversion is a phase-controlled method that does not rely on sparsity assumptions. It utilises waveform differences to constrain the inversion, facilitating the resolution of thinner layers. Because waveform disparity is a key parameter independent of seismic wavelet frequency, this method provides higher-resolution identification of thin beds within the seismic frequency band. Moreover, SWII uses a Markov Chain Monte Carlo approach to integrate well-log data with seismic texture, thereby resolving lithological units thinner than the tuning thickness ($1/4$ wavelength).

The analysis of lithofacies profiles across wells CB25, CB26, CB213, CB273, CB292, and CB24 revealed that sand and shale boundaries are clearly defined (Figure 15). These profiles confirm the extent of the sand bodies within the Guantao Formation oil layer.

The high concordance between the drilled well data and the inversion results further confirms the reliability of the inversion. The strong correlation between well points

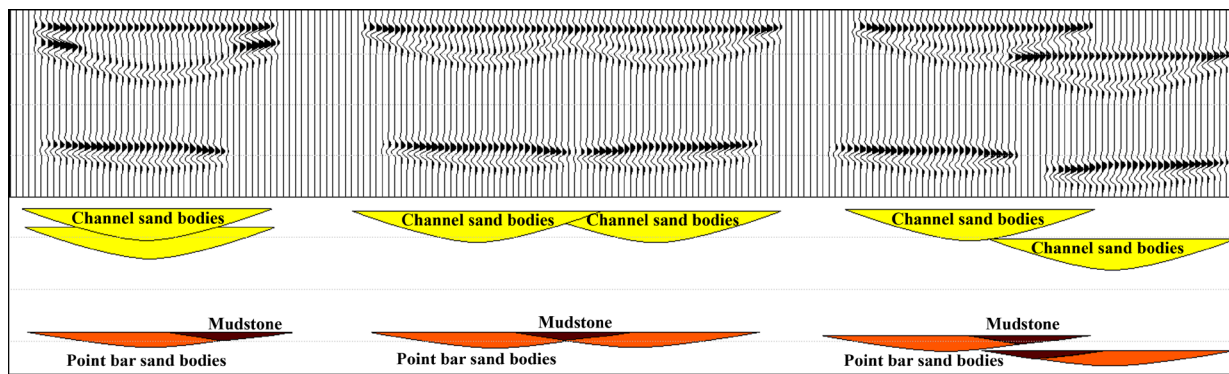


Figure 9. Forward geological model illustrating stacked channel and point-bar sand bodies within mudstone. Variations in sand-body geometry and thickness (3–50 m) generate distinct seismic waveform responses, providing a basis for evaluating the ability of waveform-indication inversion to resolve thin interbeds.

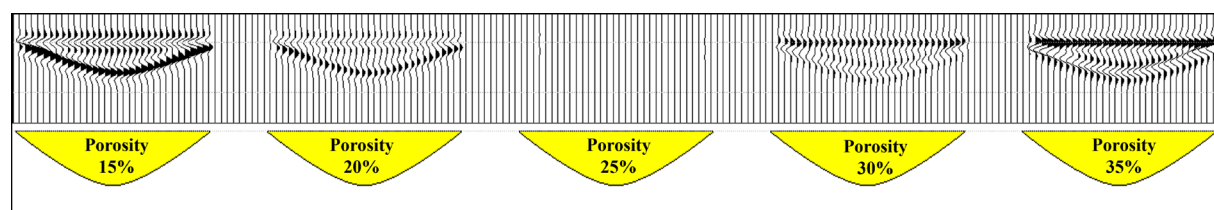


Figure 10. Focus on porosity sensitivity. Seismic amplitude response as a function of increasing porosity.

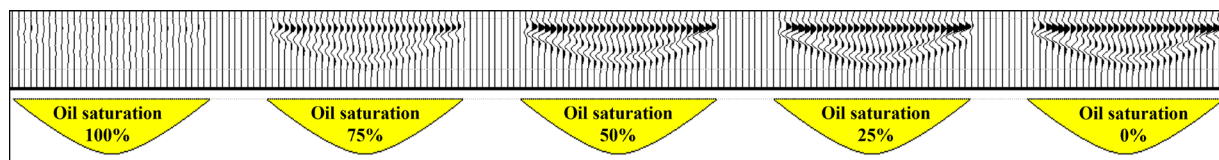


Figure 11. Focus on fluid sensitivity. Effect of oil saturation on sandstone velocity and seismic amplitude with different saturation levels.

and inverted parameters indicates that the model is both physically reasonable and highly predictive (Table 1).

Furthermore, 3D visualisation enables the identification of “sweet spots” and optimal drilling targets by illustrating the spatial continuity of reservoir properties. The lithofacies inversion elucidates structural and stratigraphic features, with colour-coded values highlighting lateral facies changes and depositional architecture (Figure 16A). The porosity distribution revealed variations in reservoir quality; prospective zones (porosity > 20% within sandstone lithofacies) are concentrated around the key wells CB24 and CB25, corresponding to channel sand systems (Figure 16B). In contrast, low-porosity zones (<

10%) associated with mixed sand–shale layers or mudstone lithofacies around wells CB273, CB26, CB213, and CB292 correspond to water-saturated or non-reservoir intervals.

This study demonstrates that integrating rock physics analysis with seismic inversion provides a robust, interpretable, and physically grounded framework for characterising thin-bed reservoirs. It enables the accurate estimation of reservoir characteristics, elastic parameters, and fluid discrimination across various lithologies. Consequently, this workflow has the potential to significantly improve exploration efficiency, optimise well-completion selection, and reduce uncertainties in subsurface modelling.

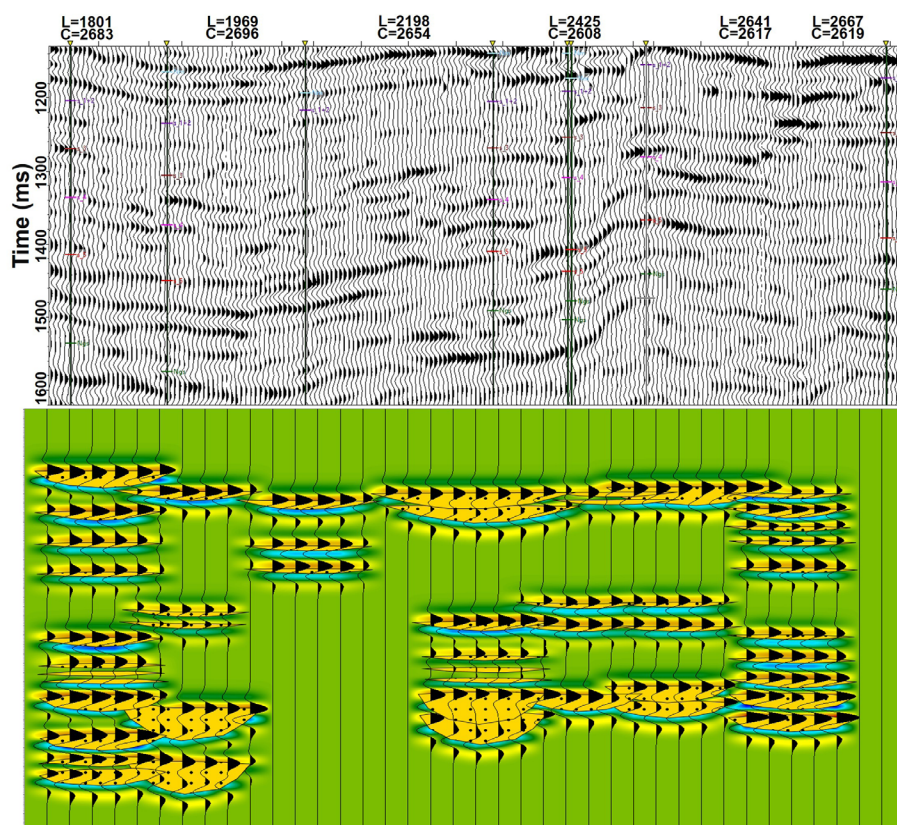


Figure 12. Comparison of forward-modelling results showing bright-spot characteristics in high-porosity, oil-bearing zones

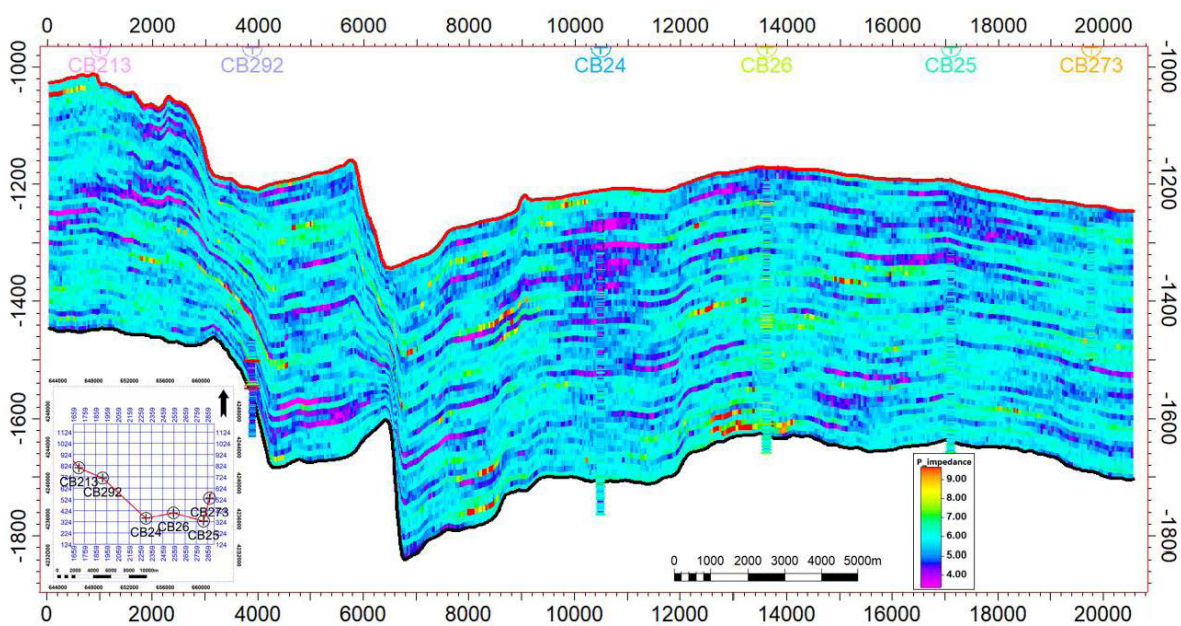


Figure 13. Cross-sectional view of P-impedance (g/cm³ × km/s) inversion profile in Guantao Formation

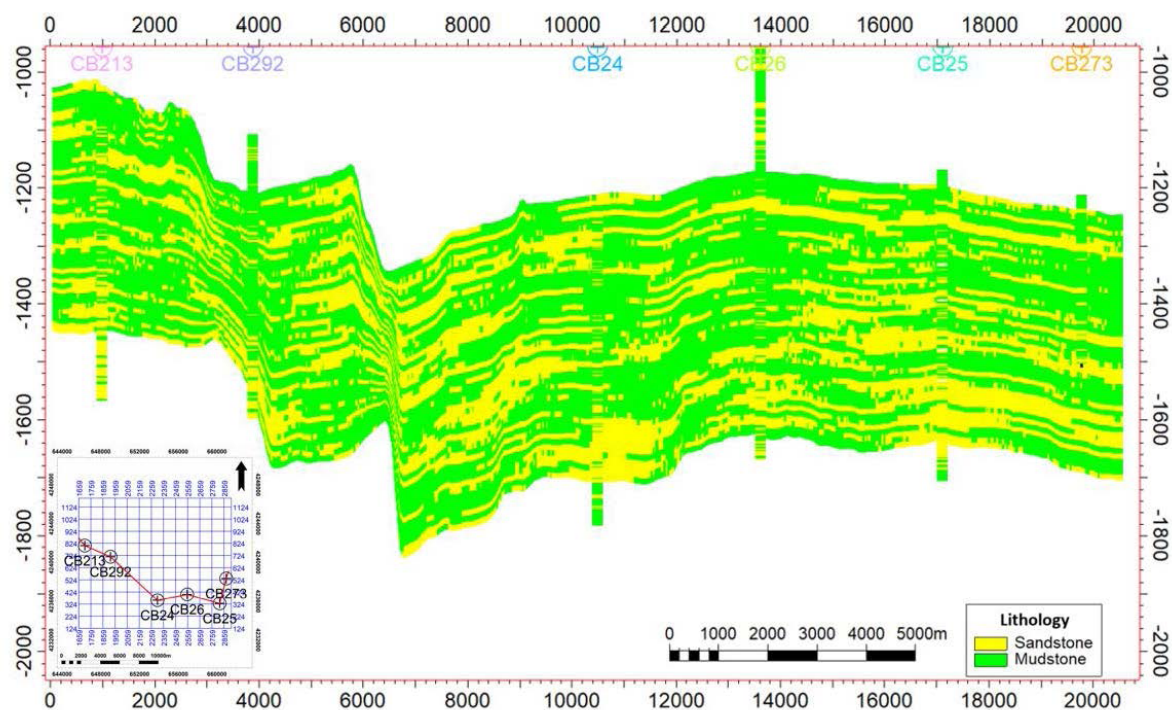


Figure 14. Visualisation of the spatial extent and hydrocarbon potential of the identified sand bodies. The distribution map highlights the favourable zones within the oil-bearing layers, suggesting significant extraction potential.

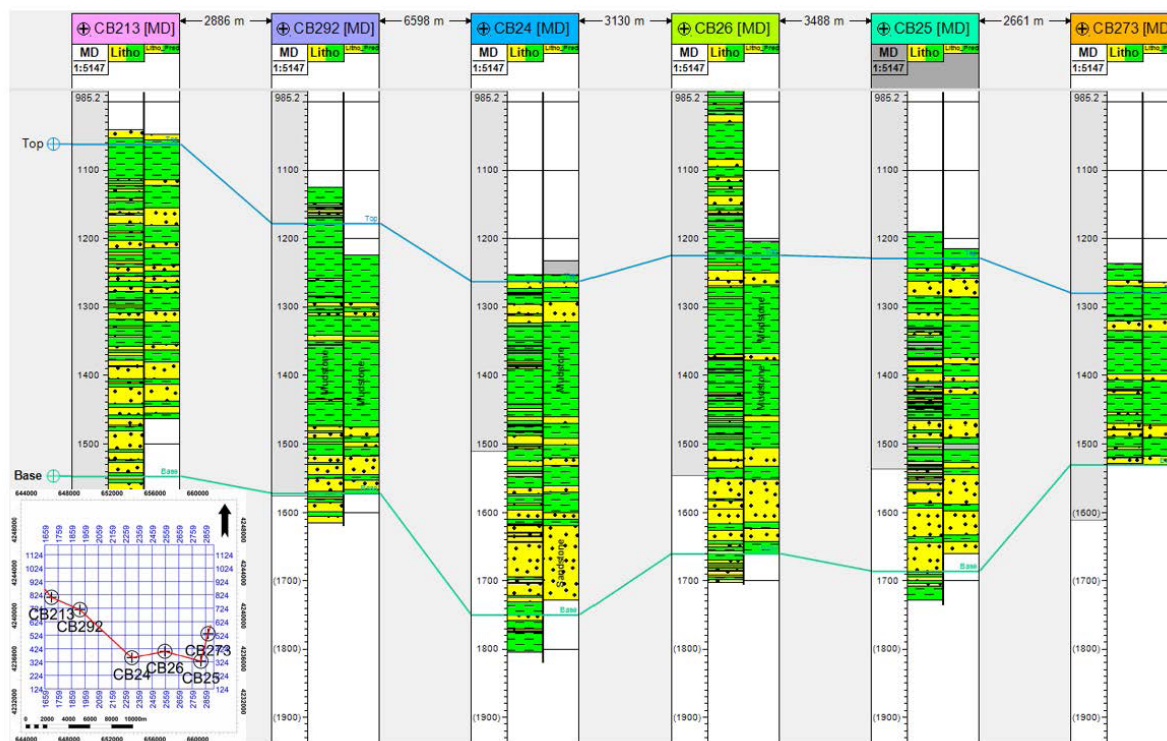


Figure 15. Seismically inverted lithofacies profiles across wells CB25, CB26, CB213, CB273, CB292, and CB24. The inversion results clearly delineate the boundaries between sand and shale, confirming the lateral extent and continuity of the sand bodies within the target Guantao Formation (for cross-section analysis, refer to Figure 1).

Table 1. Depth uncertainty through the studied wells

Well	Depth (Logs; m)	Depth (Seismic; m)	Difference (m)
CB213	-1,044.40	-1,043.90	-0.50
CB292	-1,160.40	-1,162.50	2.40
CB24	-1,241.50	-1,241.00	-0.50
CB26	-1,193.00	-1,192.90	-0.10
CB25	-1,208.00	-1,207.95	-0.05
CB273	-1,254.88	-1,254.18	-0.70
CB213	-1,529.40	-1,528.30	-1.10
CB292	-1,554.40	-1,554.40	0.00
CB24	-1,728.50	-1,728.50	0.00
CB26	-1,628.50	-1,628.50	0.00
CB25	-1,665.00	-1,664.80	-0.20
CB273	-1,504.88	-1,507.68	2.80

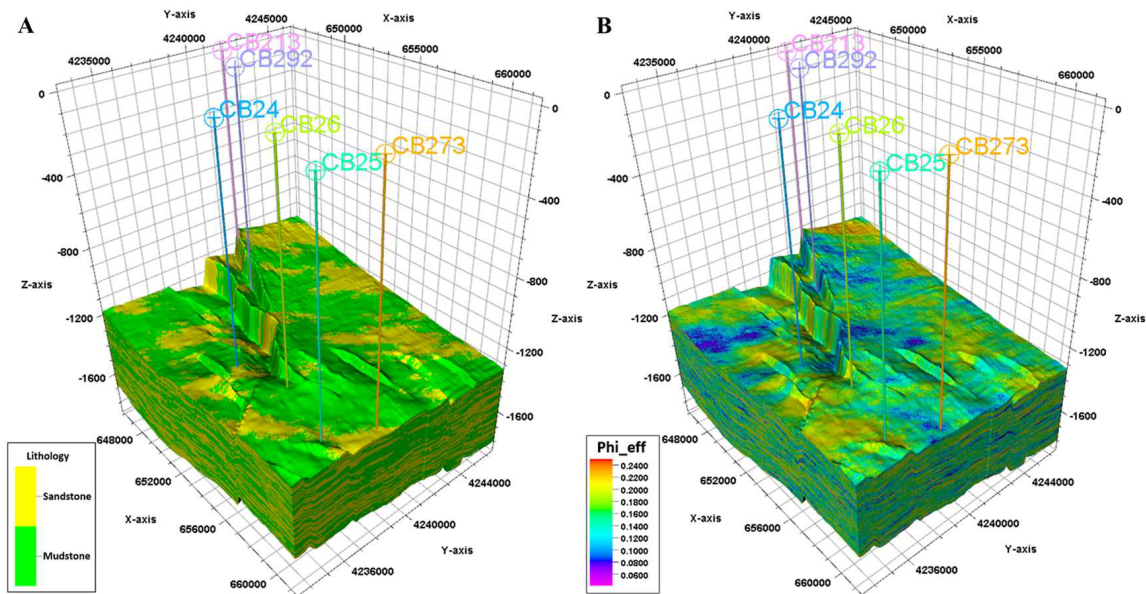


Figure 16. 3D visualisation of reservoir properties for “sweet spot” identification. (A) Lithofacies inversion volume illustrating the structural features, depositional architecture, and lateral facies changes. (B) 3D porosity distribution showing variations in reservoir quality. High-quality prospective zones (porosity > 20%) are concentrated in sandstone channels near wells CB24 and CB25, while low-porosity zones correspond to mudstone-rich intervals.

5. Conclusion

This study provides critical insights into the distribution of the Guantao Formation, emphasising the challenges of identifying thin sand layers less than 10 m thick. By employing SWII, we addressed the difficulty in distinguishing these layers, achieving high-precision reservoir prediction. The SWII effectively identifies thin sand-shale layers, reflecting various stacking patterns and the effects of porosity and oil saturation on sand-body velocity.

The integration of GMM provides a promising alternative for anomaly detection, leveraging inter-variable relationships (GR, Vp/Vs, and DEN) to improve geological interpretability and facies classification. The results demonstrate reliable effects and considerable predictive power, with inversion parameters coinciding closely with rock physics and GMM analysis. Consequently, this research establishes a robust foundation for future multilayer system deployment and efficient oil and gas exploration in the Bohai Bay Basin and similar geological settings.

Acknowledgments

None.

Funding

This research was financially supported by the offshore production plant, Sinopec Shengli Oilfield Company, Dongying, China (30200012-25-ZC0699-0008).

Conflict of interest

Yanjia Wu, Zhiwen Dong, Yiyao Zhang, Zekun Li, Shihao Liu, Xiaoxiao Shi, Hongke Zhou, Qunhu Wu, Shina Yu are employees of Sinopec Shengli Oilfield Company; however, they were not involved in any activities that could constitute a conflict of interest in relation to this study. The authors confirm that they have no conflicts of interest, financial or non-financial, associated with the material presented in this manuscript.

Author contributions

Conceptualization: Yanjia Wu, Qamar Yasin

Formal analysis: Zhiwen Dong, Hongke Zhou, Shina Yu

Investigation: Qunhu Wu, Shihao Liu, Xiaoxiao Shi

Methodology: Yanjia Wu, Qamar Yasin

Validation: Zekun Li, Ahmed Bakry, Qamar Yasin

Visualization: Yanjia Wu, Qamar Yasin, Ahmed Bakry

Writing—original draft: Yanjia Wu, Yiyao Zhang

Writing—review & editing: Shina Yu, Yiyao Zhang, Ahmed Bakry

Availability of data

All datasets, trained models, and analysis code will be made available upon publication through a code repository (<https://github.com/qamarquaidian/Pore-Structure-Analysis>).

References

1. Darling T. *Well Logging and Formation Evaluation*. Gulf Professional Publishing; 2005.
doi: 10.1016/B978-0-7506-7883-4.X5000-1
2. Lai J, Su Y, Xiao L, *et al*. Application of geophysical well logs in solving geologic issues: Past, present and future prospect. *Geosci Front*. 2024;15(3):101779.
doi: 10.1016/j.gsf.2024.101779
3. Zhu J, Qu Y, Duan W, Zhang Q, Yang C. Frequency Domain Acoustic Wave Full Waveform Inversion Based on Average-Derivative Optimization. *J Seism Explor*. 2024;33(5):21-38.
4. Pe'er U, Dy JG. Automated target detection for geophysical applications. *IEEE Trans Geosci Remote Sens*. 2016;55(3):1563-1572.
doi: 10.1109/TGRS.2016.2627245
5. Struminskiy K, Klenitskiy A, Reshytko A, *et al*. Well log data standardization, imputation and anomaly detection using hidden Markov models. In: *Petroleum Geostatistics 2019*; September 2-6, 2019; Florence, Italy. European Association of Geoscientists & Engineers; 2019:1-5.
doi: 10.3997/2214-4609.201902208
6. Al-Fakih A, Kaka SI, Koeshidayatullah AI. Reservoir property prediction in the North Sea using machine learning. *IEEE Access*. 2023;11:140148-140160.
doi: 10.1109/ACCESS.2023.3336623
7. Liu J, Zhu H, Liu Y, Wu H, Lan Y, Zhang X. Anomaly detection for time series using temporal convolutional networks and Gaussian mixture model. In: *Journal of Physics: Conference Series*. 2018 International Symposium on Power Electronics and Control Engineering (ISPECE 2018); December 28-30, 2018; Xi'an, China. IOP Publishing; 2019;1187(4):042111.
doi: 10.1088/1742-6596/1187/4/042111
8. Reynolds D. Gaussian mixture models. In: Li SZ, Jain AK, eds. *Encyclopedia of Biometrics*. Springer; 2015:827-832.
doi: 10.1007/978-1-4899-7488-4_196
9. Jia J, Zheng X, Wang Y, *et al*. The effect of artificial intelligence evolving on hyperspectral imagery with different signal-to-noise ratio, spectral and spatial resolutions. *Remote Sens Environ*. 2024;311:114291.
doi: 10.1016/j.rse.2024.114291
10. Sun M, Yasin Q, Daniel E, *et al*. Swin-UNETR: A transformer-based model for 3D pore network segmentation

- in low-permeability sedimentary rocks. *Int J Coal Geol.* 2026;315:104951.
doi: 10.1016/j.coal.2026.104951
11. Ashraf U, Zhang H, Anees A, *et al.* Controls on reservoir heterogeneity of a shallow-marine reservoir in Sawan Gas Field, SE Pakistan: Implications for reservoir quality prediction using acoustic impedance inversion. *Water.* 2020;12(11):2972.
doi: 10.3390/w12112972
 12. Yasin Q, Ding Y, Baklouti S, Boateng CD, Du Q, Golsanami N. An integrated fracture parameter prediction and characterization method in deeply-buried carbonate reservoirs based on deep neural network. *J Pet Sci Eng.* 2022;208:109346.
doi: 10.1016/j.petrol.2021.109346
 13. Jackson CA, Larsen E. Temporal constraints on basin inversion provided by 3D seismic and well data: a case study from the South Viking Graben, offshore Norway. *Basin Res.* 2008;20(3):397-417.
doi: 10.1111/j.1365-2117.2008.00359.x
 14. Chetia B, Nagaraju J, Sharma SK, Kumar A. Reservoir characterization through seismic inversion: A case study. In: *SEG Technical Program Expanded Abstracts 2019*. 89th Society of Exploration Geophysicists International Exposition and Annual Meeting 2019; September 15-20, 2019; San Antonio, TX. SEG; 2019:3215-3219.
doi: 10.1190/segam2019-3216159.1
 15. Ali M, Zhu P, Jiang R, *et al.* Data-driven lithofacies prediction in complex tight sandstone reservoirs: a supervised workflow integrating clustering and classification models. *Geomech Geophys Geo-Energy Geo-Resour.* 2024;10(1):70.
doi: 10.1007/s40948-024-00787-5
 16. Zhang R, Sen MK, Phan S, Srinivasan S. Stochastic and deterministic seismic inversion methods for thin-bed resolution. *J Geophys Eng.* 2012;9(5):611-618.
doi: 10.1088/1742-2132/9/5/611
 17. Ashraf U, Zhang H, Anees A, *et al.* Application of unconventional seismic attributes and unsupervised machine learning for the identification of fault and fracture network. *Appl Sci.* 2020;10(11):3864.
doi: 10.3390/app10113864
 18. Yakovleva IP, Filippova KE, Kuznetsov VI, Keller TE. Comparison of deterministic and geostatistical inversion results-a case study for a gas-saturated reservoirs with coals. In: *Proceedings of the 78th EAGE Conference and Exhibition 2016*; May 30 - June 2, 2016; Online. European Association of Geoscientists & Engineers; 2016:1-5.
doi: 10.3997/2214-4609.201601337
 19. Haas A, Dubrule O. Geostatistical inversion-a sequential method of stochastic reservoir modelling constrained by seismic data. *First Break.* 1994;12(11).
doi: 10.3997/1365-2397.1994034
 20. Liu X, Li J, Chen X, *et al.* Stochastic inversion of facies and reservoir properties based on multi-point geostatistics. *J Geophys Eng.* 2018;15(6):2455-2468.
doi: 10.1088/1742-2140/aac694
 21. Li X, Shen Y, Chen W, Feng S, Zhang Z, Hong H. Seismic response of ice-water-covered two-layer seabed under obliquely incident P/SV-wave considering imperfect interface condition. *Ocean Eng.* 2025;342:122872.
doi: 10.1016/j.oceaneng.2025.122872
 22. Ullah J, Luo M, Ashraf U, *et al.* Evaluation of the geothermal parameters to decipher the thermal structure of the upper crust of the Longmenshan fault zone derived from borehole data. *Geothermics.* 2022;98:102268.
doi: 10.1016/j.geothermics.2021.102268
 23. Gao J, Bi JJ, Zhao HS, Fu ZF. Seismic waveform inversion technology and application of thinner reservoir prediction. *Prog Geophys.* 2017;32(1):142-145. [In Chinese].
doi: 10.6038/pg20170119
 24. Ashraf U, Zhang H, Thanh HV, *et al.* A robust strategy of geophysical logging for predicting payable lithofacies to forecast sweet spots using digital intelligence paradigms in a heterogeneous gas field. *Nat Resour Res.* 2024;33(4):1741-1762.
doi: 10.1007/s11053-024-10350-4
 25. Cheng Z, Liang Z, Han J, *et al.* Application of pre-stack geostatistical inversion in prediction of thin high-quality conglomerate reservoir in Shawan Depression, Junggar Basin. *Front Earth Sci.* 2025;19(4):498-504.
doi: 10.1007/s11707-025-1159-9
 26. Wang J, Li Z, Chen C, Meng C. Predicting gas content in coalbed methane reservoirs using seismic waveform indication inversion: a case study from the Upper Carboniferous Benxi Formation, eastern Ordos Basin, China. *Acta Geophys.* 2022;70(2):623-638.
doi: 10.1007/s11600-022-00745-3
 27. Helgesen J, Magnus I, Prosser S, *et al.* Comparison of constrained sparse spike and stochastic inversion for porosity prediction at Kristin Field. *Leading Edge.* 2000;19(4):400-407.
doi: 10.1190/1.1438620
 28. Li B, Bi J, Wei S, Liu M, Han S. Application of waveform inversion technique in prediction of beach bar sandstone reservoir. In: *Proceedings of the International Geophysical Conference*; April 17-20, 2017; Qingdao, China. Society of Exploration Geophysicists and Chinese Petroleum Society; 2017:686-689.

- doi: 10.1190/IGC2017-174
29. Nizarul O, Hermana M, Bashir Y, Ghosh DP. Improving thin bed identification in Sarawak Basin field using short time Fourier transform half Cepstrum (STFTHC) method. In: *IOP Conference Series: Earth and Environmental Science*. 3rd International Conference on Geological, Geographical, Aerospace and Earth Science 2015 (AeroEarth 2015); September 26-27, 2015; Jakarta, Indonesia. IOP Publishing; 2016;30(1):012003.
doi: 10.1088/1755-1315/30/1/012003
 30. Babasafari AA, Bashir Y, Ghosh DP, *et al.* A new approach to petroelastic modeling of carbonate rocks using an extended pore-space stiffness method, with application to a carbonate reservoir in Central Luconia, Sarawak, Malaysia. *Leading Edge*. 2020;39(8):592a1-592a10.
doi: 10.1190/tle39080592a1.1
 31. Fateh A, Al-Haji A, Elmasry H, Bikchandaev E. Customized workflow for proactive geosteering in thin heterogeneous carbonate reservoirs utilizing deep resistivity inversion. In: *Proceedings of the SPWLA 64th Annual Logging Symposium*; June 10-14, 2023; Lake Conroe, TX. Society of Petrophysicists and Well Log Analysts; 2023:D041S008R003.
doi: 10.30632/SPWLA-2023-0046
 32. Purnomo EW, Abdul Latiff AH, Elsaadany MA. Predicting reservoir petrophysical geobodies from seismic data using enhanced extended elastic impedance inversion. *Appl Sci*. 2023;13(8):4755.
doi: 10.3390/app13084755
 33. Liu H, Lin C, Zhang X, Wang H, Fu X. Sedimentological characteristics and controlling factors on fluvial styles of the Guantao Formation, Kongdian Oilfield, China. *J Earth Syst Sci*. 2019;128(2):37.
doi: 10.1007/s12040-018-1058-5
 34. Wang J, Li ZG. Lacustrine Facies Sedimentary Characteristics of Guantao Formation in Zhanhua Sag. *Appl Mech Mater*. 2014;522-524:1337-1340.
doi: 10.4028/www.scientific.net/AMM.522-524.1337
 35. Liu FT, Ting KM, Zhou ZH. Isolation forest. In: Giannotti F, Gunopulos D, Turini F, Zaniolo C, Ramakrishnan N, Wu X, eds. *Proceedings of the 2008 Eighth IEEE International Conference on Data Mining*; December 15-19, 2008; Pisa, Italy. IEEE; 2008:413-422.
doi: 10.1109/ICDM.2008.17
 36. Xue S, Li Z, Feng G, *et al.* Visualization characterization of void structure evolution in broken limestone and its influence on permeability. *Fuel*. 2026;405:136412.
doi: 10.1016/j.fuel.2025.136412