Petroleum Science xxx (xxxx) xxx

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science

Original Paper

Genesis of lamina combinations and intelligent well-logging interpretation in the upper Xiaganchaigou Formation, Yingxi area, Qaidam Basin, China

Jia-Lin Fu [a,b,c], Da-Li Yue [a,b,c,*], Wu-Rong Wang [a,b,c,**], Kun-Yu Wu [d,e], Han Wang [b,c,f], Ying-Hai Jiang [d,e], Shu-Qi Zhang [a,b,c], Zi-Mo Xu [a,b,c], Wei Li [a,b,c]

[a] Hainan Institute of China University of Petroleum (Beijing), Sanya, 572025, Hainan, China
[b] College of Geosciences, China University of Petroleum (Beijing), Beijing, 102249, China
[c] State Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum (Beijing), Beijing, 102249, China
[d] Qinghai Provincial Key Laboratory of Plateau Saline-Lacustrine Basinal Oil&Gas Geology, Dunhuang, 736202, Gansu, China
[e] Research Institute of Exploration and Development, PetroChina Qinghai Oilfield Company, Dunhuang, 736202, Gansu, China
[f] Chengdu Exploration and Development Research Institute of PetroChina Daqing Oilfield Company Ltd., Chengdu, 610051, Sichuan, China

ARTICLE INFO

Article history:
Received 11 July 2025
Received in revised form
13 January 2026
Accepted 10 April 2026
Available online xxx

Edited by Xiu-Fang Hu

Keywords:
Lamina combination identification
Transformer intelligent recognition
Sedimentary origin
E_3^2 segment
Qaidam Basin

ABSTRACT

The laminar sedimentary structures of saline lacustrine mixed rocks affect both organic matter enrichment and reservoir storage performance. However, due to the small-scale nature of laminae, large-scale identification using well-logging data during reservoir exploration and development remains challenging. It is necessary to introduce a research method to identify and characterize the development of different types of laminae. Based on analyses of typical cores, XRD data, and well-logging curves from the upper member of the Xiaganchaigou Formation in the Yingxi area, five main types of laminae and six lamina combinations were classified. A Transformer-based intelligent recognition method was then applied to identify these lamina combinations from well-log data, with the Random Forest algorithm used as a comparative benchmark. Verification results show that the Transformer model achieves a higher total accuracy of 84% in lamina combination recognition. This study proposes a new approach for the conventional well-log characterization of laminae, in which the classification is established from the perspective of laminae genesis. It reflects the development patterns of lamina combinations driven by paleoenvironmental changes, and selects an appropriate intelligent recognition method to address the challenges in well-log characterization of such reservoirs. In terms of engineering applications, this study can accurately indicate the positions of high-quality reservoirs within sedimentary cycles during field development. It provides a sedimentary facies-controlled basis for the three-dimensional characterization of reservoir quality, thereby offering a valuable reference for the exploration and development of reservoirs formed under similar sedimentary conditions.

1. Introduction

Laminae in lacustrine mixed lithologies have become a key focus in the exploration and development of unconventional hydrocarbon reservoirs (Feng et al., 2011; Song et al., 2024). Similar to shale reservoirs, laminae are generally regarded as favorable targets (Xin et al., 2022) and play a significant guiding role in evaluating reservoir properties such as storage capacity and compressibility (Pollastro, 2007; Zhao et al., 2020). Laminae represent a common microscopic structural feature in sedimentary rocks, typically formed through the slow deposition of fine-grained sediments under stratified water conditions. These structures typically occur as millimeter-scale layers and only rarely reach the centimeter scale, and they may consist of a single lithology or alternating lithologies. The interfaces between laminae are often well-defined and may appear planar, wavy, or

* Corresponding author.
** Corresponding author.
E-mail addresses: yuedali@cup.edu.cn (D.-L. Yue), wwr@cup.edu.cn (W.-R. Wang).
Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2026.04.014

lenticular in form (Jiang et al., 2013; Mingram, 1998; Wang et al., 2019). As the smallest observable stratification units, lamina combinations reflect short-term fluctuations in depositional environments, whereas stratification (bedding) typically consists of multiple lamina combinations and indicates longer-term sedimentary processes and environmental changes (Anderson and Dean, 1988; Bai et al., 2018; He et al., 2025a). Some researchers have proposed that laminae exhibit a multi-scale structural nature, encompassing a wide range of scales from macroscopic to microscopic. These multi-scale structures are present across stratigraphic intervals and within individual rock units, comprising various lithologies and superimposed mineralogical cycles. Together, these structures form a hierarchical framework linking laminae, lithologic units, and sedimentary facies (Chen et al., 2017; Shi et al., 2018; Wang et al., 2019). The structural elements within this system are often interrelated, with millimeter-scale laminae frequently reflecting geological information about decimeter-to meter-scale facies (Li et al., 2018, Li et al., 2019). Compared to their marine counterparts, lacustrine laminae are less affected by oceanic circulation and other external disturbances, and therefore more directly record depositional environments and sedimentary mechanisms (Sun and Wang, 2005; Wang and Zhong, 2004). Some scholars have employed astronomical tuning methods to establish high-resolution chronostratigraphic frameworks of lacustrine deposits, revealing the Milankovitch cyclicity of lacustrine sedimentary sequences and indicating the controlling effect of climate change on lithological variations (Shi et al., 2019). Lacustrine sedimentation is typically more localized and is often influenced by a variety of terrestrial clastic sources. Along basin margins, depositional processes are affected by distal supplies from alluvial fan termini and deltaic systems. Additionally, structural uplifts on the basin margins frequently lead to the incorporation of gypsum clasts derived from platform-scale evaporative settings. In the above situation, lacustrine mixed rock reservoirs are usually formed (Moreira-Lima and De-Ros, 2019; Olivito and Souza, 2020). Within these reservoirs, laminae exhibit diverse mineral compositions, including carbonates, quartz, gypsum, and clay minerals. The morphology, stacking relationships, and distribution of the laminae vary significantly, resulting in strong vertical heterogeneity and the presence of multiple lamina combinations (Peng et al., 2022; Yang et al., 2020). The depositional processes of different lamina combinations are particularly sensitive to changes in external conditions such as temperature and salinity (Pan et al., 2021; Zeng et al., 2023), leading to diverse genetic mechanisms. However, current research on laminae in lacustrine mixed lithologies has primarily focused on the formation mechanisms of individual mineral laminae and their relationship with paleoclimate indicators (Guo et al., 2018; He et al., 2025b; Li et al., 2020; Park and Fürsich, 2001; Wang et al., 2014). Studies on more complex lamina combinations remain limited. The division scheme lacks genetic basis, and the paleoenvironmental evolution at the scale of lamina combination is not fully understood. However, the diagenetic variations of lacustrine organic-rich shales at the lamina scale exert a significant influence on pore structure and hydrocarbon generation potential. Such diagenetic heterogeneity at this scale is a key factor controlling reservoir quality, and accurately characterizing lamina combinations is therefore of great significance for the development of these reservoirs (Liang et al., 2018).

In practical development and production research, laminae are a key focus in microscopic reservoir studies; however, large-scale characterization at the macroscopic level has progressed relatively slowly. Current research remains largely concentrated in the domain of analytical and testing techniques. At present, most laminae characterization relies on core description, thin-section petrography, and a series of well-logging identification methods. Conventional logging tools, such as natural gamma ray, acoustic, and resistivity logs, are typically limited to identifying laminae structures at meter-scale resolution (Pang et al., 2023; Song et al., 2023). LithoScanner logging techniques—commonly used in the development of mixed lithologic reservoirs—enable vertical analysis of mineral composition and allow for the interpretation of decimeter-scale lithological variations. New-generation logging tools, such as elemental capture spectroscopy (ECS), can identify the concentration of major rock-forming elements and their corresponding minerals with high vertical resolution, thereby facilitating decimeter-scale mineralogical discrimination (Fan et al., 2024; Lai et al., 2022). Imaging logging provides comprehensive two-dimensional representations of rock physical features, making it possible to identify sedimentary structures, including laminae, at millimeter to micrometer scales (Lai et al., 2018; Wang et al., 2013). These techniques enable the characterization of laminae development across multiple scales. However, their widespread application in actual field development is often limited by the high cost associated with specialized logging tools. Although tools like LithoScanner can be widely applied, their effectiveness is limited by the complex diagenetic mineral assemblages and strong heterogeneity of lacustrine mixed rock reservoirs. As a result, the responses of different lamina combinations to various logging measurements are typically non-linear (Song et al., 2017; Tian et al., 2023; Zhang et al., 2018). Traditional interpretation techniques—such as rock physics modeling, cross-plotting, and multiple linear regression—often fail to account for such complexity and thus struggle to deliver accurate well-log interpretations (Ran et al., 2016).

With the advancement of machine learning, artificial intelligence (AI)-based recognition methods have been widely applied to well-log interpretation. These methods offer high precision and efficiency, making them a current hotspot in lithofacies identification and other interpretation tasks (Gama et al., 2025; Liu et al., 2026; Wang et al., 2015, 2025; Zhang et al., 2024, 2025). A variety of mature machine learning algorithms have been integrated into geological workflows and are suitable for addressing various challenges in exploration and development. For instance, Convolutional Neural Networks (CNNs) excel at extracting local features and are particularly effective in learning from high-dimensional data such as images. They also support multi-source data integration, thereby significantly enhancing interpretation accuracy (Haritha et al., 2025; Li et al., 2024c; Yue et al., 2025; Zhang et al., 2021). For one-dimensional data types such as well logs, Random Forest (RF) algorithms are commonly used. These algorithms process data in parallel by constructing multiple decision trees and aggregating their predictions, which substantially improve the model's generalization ability and robustness while maintaining high computational efficiency (Fan et al., 2025; Feng et al., 2021; Song et al., 2023). However, the aforementioned machine learning algorithms still exhibit limitations when applied to the identification of lamina combinations in lacustrine mixed lithologic reservoirs. For example, CNNs are highly dependent on large volumes of multidimensional data. When applied solely to low-dimensional datasets such as conventional well logs, CNNs often fail to achieve satisfactory performance. RF, on the other hand, is an ensemble method that integrates multiple decision trees, but it is inherently a "black box" model. Given the complexity of laminae types in such reservoirs, the model can be easily misled by erroneous information during training, and lacks transparency for correction, thereby compromising interpretation reliability (Feng et al., 2021; Liu et al., 2025; Wang et al., 2023). Accurate lamina combination identification requires models that integrate multiple logs and capture long-range vertical

dependencies. Attention mechanisms are therefore well suited for suppressing low-importance features and enhancing sequence modeling capability (Ren and Wang, 2021).

This study focuses on the upper member of the Xiaganchaigou Formation in the western Yingxiongling structural belt of the Qaidam Basin, northwestern China (Yingxi area), which represents a saline lacustrine mixed sedimentary reservoir. The objectives are to clarify the genetic mechanisms and classification of lamina combinations, infer the vertical development patterns and their correspondence with paleoenvironmental evolution, and to optimize intelligent identification methods while evaluating their reliability and advantages. This research proposes a novel approach for characterizing laminae in saline lacustrine mixed rocks using conventional well logs. From a genetic perspective, the study establishes a classification scheme for lamina combinations, revealing the development patterns driven by paleoenvironmental controls. Furthermore, it endows intelligent algorithms with "geological reasoning", yielding promising results in practical reservoir interpretation. The proposed method offers an effective solution to the challenge of characterizing complex lithological laminae and provides valuable insights for lamina characterization and logging interpretation in similar depositional environments.

2. Geological settings

The Qaidam Basin, situated in northwestern China on the northeastern margin of the Tibetan Plateau, is the largest Cenozoic basin in the region (Fig. 1(a)), covering an area of 258,000 km^2 with an average elevation of 2800 m (Li et al., 2021; Xiong et al., 2018). It represents a critical area for investigating tectonic evolution and hydrocarbon accumulation. The basin is bounded by the Altyn Tagh Fault to the northwest, the Qilian Mountains to the northeast, and the Kunlun Mountains to the south. The complex structure of the Qaidam Basin is primarily attributed to the Himalayan orogeny (Cenozoic, ca. 65 Ma to present) and earlier tectonic events during the Indosinian (Triassic, Mesozoic) and Yanshanian (Jurassic–Cretaceous, Mesozoic) orogenies. These successive orogenies, driven by plateau uplift and peripheral compression, shaped the basin's distinctive intermontane features (Song et al., 2010; Zhang et al., 2018; Zhu et al., 2022).

The Yingxiongling structural belt, situated in the southwestern Qaidam Basin, exhibits a large arcuate anticline with steep southern and gentle northern limbs. Its internal structure is highly fragmented, showing intense fold deformation (Fu et al., 2015; Qiao et al., 2020). The Yingxi area, located in the western part of this structural belt, represents the most prolific oil-producing block in the basin. During the Cenozoic period, the sedimentary center of the Qaidam Basin underwent continuous migration. In the Paleogene, the main lake body shifted progressively northeastward along the Altyn Tagh Fault, while in the Neogene, the uplift of the Altyn Mountains caused the main lake body to migrate southward (Guo et al., 2018). Throughout the Paleogene–Neogene, the lacustrine basin experienced predominantly semi-saline to saline conditions. In the Oligocene, the uplift of the Himalayas blocked the inflow of marine moisture, leading to an increasingly arid environment within the basin (Sun and Wang, 2005). During the early Oligocene, the depositional environment transitioned from lacustrine to salt lake settings. In the middle to late Oligocene, semi-saline lakes were developed, with fan delta deposits forming in structurally uplifted areas along the basin margins (Li et al., 2021). Yingxi area, adjacent to the Altyn Mountain structural uplift and located along the basin margin, is primarily characterized by the development of fan delta, littoral–shallow lake, and semi-deep lake deposits (Zhang et al., 2018). As a result, the Yingxi area encompasses sedimentary facies belts ranging from shallow- to deep-lake environments, as well as distal parts of deltas, resulting in complex sedimentary compositions (Fig. 1(b)).

The stratigraphy of this area includes two major sequences: the Paleogene and Neogene. The Paleogene sequence comprises, from bottom to top, the Lulehe Formation (E_{1+2}) and the Xiaganchaigou Formation (E_3), while the Neogene sequence includes the Shanganchaigou Formation (N_1), the Xiayoushashan Formation (N_2^1), and the Shangyoushashan Formation (N_2^2) (additional Neogene strata overlie these units but are omitted here). Among them, the upper member of the Xiaganchaigou Formation (E_3^2) represents the principal reservoir in this area. It mainly consists of carbonates and mudstone, interbedded with fine-grained clastic rocks, and a gypsum layer is present at the top (Wang et al., 2020) (Fig. 1(c)). The reservoir types in the study area are typical carbonate-rich lacustrine mixed rocks, exhibiting certain features associated with saline lacustrine basins.

3. Methods and data

3.1. Data sets

The samples used in this study were provided by the Research Institute of Exploration and Development, PetroChina Qinghai Oilfield Company. They were collected from 13 vertical cored wells from the Yingxi area. The dataset includes thin-section petrographic observations, X-ray diffraction (XRD), X-ray fluorescence (XRF), and petrophysical analyses of 311 samples (including 215 with XRD results). In addition, multiple well-log datasets are available for each well. Based on these data, the types of laminae and lamina combinations were classified, their corresponding well-log response characteristics were identified, and an intelligent interpretation model was trained. The model's reliability was evaluated using blind wells. Geochemical parameters were further used to clarify the genesis and developmental patterns of each lamina combination, thereby validating the classification scheme and analyzing the strengths of the intelligent interpretation model.

The well-logging experimental data consist of well-log curves and lamina combination interpretation results from the E_3^2 interval of 13 cored wells in the Yingxi area. Five types of conventional well logs were included: gamma ray (GR), acoustic (AC), shallow lateral resistivity (LLS), deep lateral resistivity (LLD), and density (DEN), as well as LithoScanner data. A total of 4,368 well-log data points were selected from 13 wells (well-log data points were derived from the discretization of logging data from the lamina combinations development section), with 3,276 data points from 10 wells used as the training set, and 1,092 data points from 3 wells used as the testing set. The sampling interval for both training and testing sets was 0.125 m. Given the imbalance in the number of samples between lamina combinations and non-laminated intervals, which could lead to overfitting, the Synthetic Minority Over-sampling Technique (SMOTE) algorithm was employed to generate synthetic samples for underrepresented classes, thus balancing the distribution (Verbiest et al., 2014). As a result, the final training set included 13,759 samples.

3.2. Lamina combinations intelligent recognition model

3.2.1. Intelligent model principle and modules

Given the lithological complexity and strong heterogeneity of the lacustrine mixed lithologic reservoir in the study area, it is essential to address challenges such as diverse mineral compositions, strong interlayer heterogeneity, and significant variation in interpretation scales. Moreover, the interpretation needs to integrate multiple logging curves, including both conventional logs

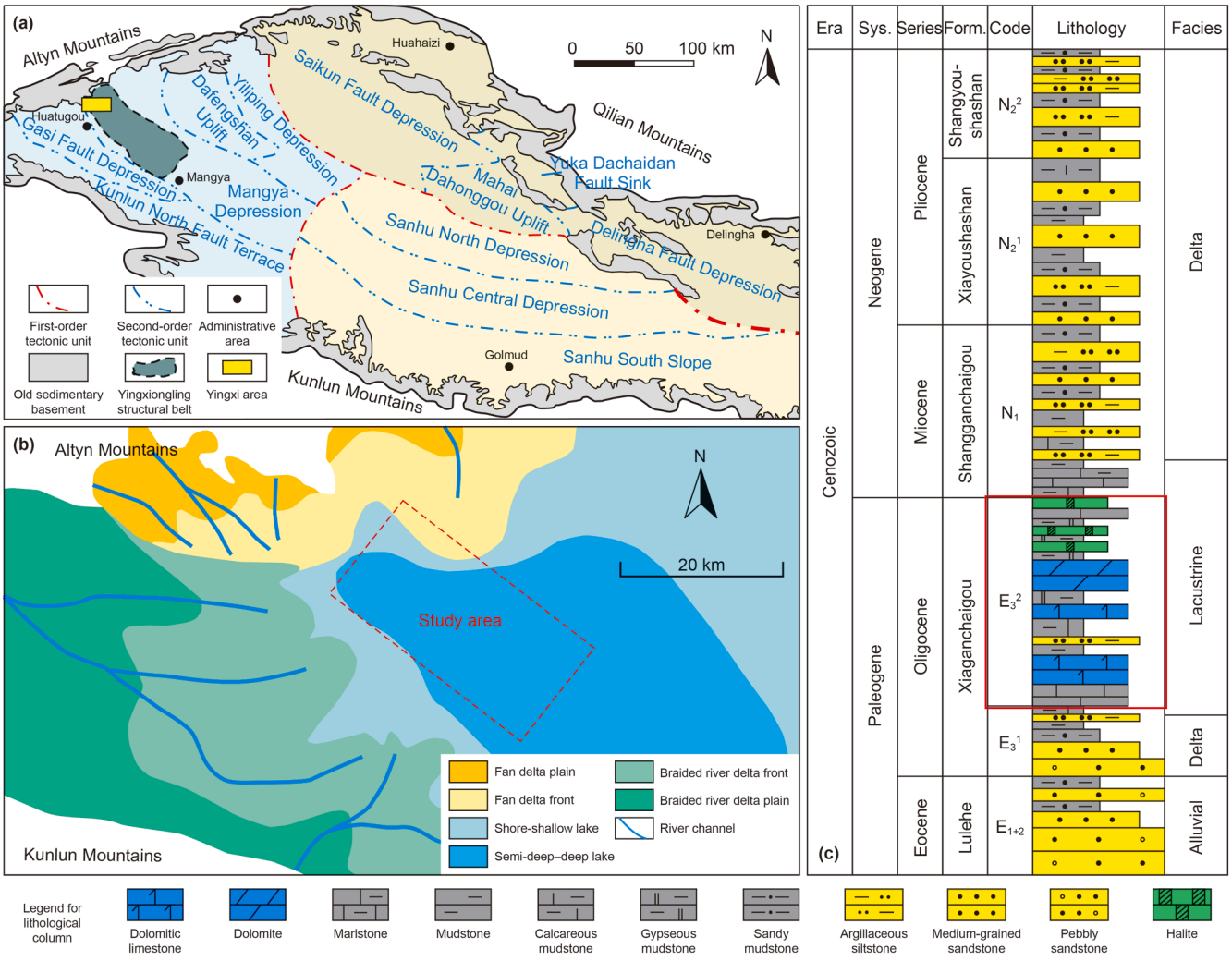

Fig. 1. Overview of the study area (modified from Qiao et al., 2020; Wang et al., 2020; Zhang et al., 2018). **(a)** Structural unit map of Qaidam Basin; **(b)** paleoenvironmental map of the study area during the Xiaganchaigou Formation period; **(c)** schematic stratigraphic column of the study area.

and LithoScanner data. Therefore, the new method must be capable of handling multi-source data fusion and long-sequence data modeling. In this study, the Transformer algorithm is applied for intelligent well log interpretation (Sun et al., 2024; Vaswani, 2017), while the widely used Random Forest algorithm is employed as a benchmark for comparison (Feng, 2021; Vijayakumar et al., 2025).

The Transformer model consists of two main components: the encoder and the decoder (Fig. 2(a)). The encoder converts the input sequence into context-aware embeddings, mapping each data point into a vector space that captures its relationship within the sequence. The decoder then establishes connections between the right-shifted input sequence (Shift Right) or class-labeled time-series data and the encoded features to generate the final prediction. Typically, multiple encoder–decoder pairs are employed in a complete Transformer framework (Fig. 2(b)). Before being fed into the encoder and decoder, both the input and output data require positional encoding. This is because the Transformer algorithm relies on parallel computation, and positional encoding is necessary to provide the model with information about the sequence order of the data in time (or depth), enabling the model to capture positional dependencies (Li et al., 2023, 2024a). The encoder comprises a Multi-Head Attention layer, a Feed Forward layer, and Residual Connection and Layer Normalization (Add & Norm). The decoder includes a Masked Multi-Head Attention layer, an Encoder–Decoder Multi-Head Attention layer, and a Feed Forward layer (Fig. 2(c)).

It should be noted that the Multi-Head Attention mechanism in the decoder differs from that in the encoder. One type of Self-Attention in the decoder masks future information to ensure that the generation of the current output depends only on the current and previous outputs, and not on any future outputs. This mechanism is therefore referred to as Masked Multi-Head Attention. Another type is known as Encoder–Decoder Multi-Head Attention, which, in addition to performing self-attention within the decoder, also interacts with the encoder's outputs. This forms encoder–decoder attention, also called cross-attention, and is responsible for aligning the features extracted by the encoder with the target labels from the decoder. Furthermore, each Multi-Head Attention layer is followed by a Feed Forward layer, which performs a nonlinear transformation on the output at each position. The computation is defined in Equation (1), where $\mathbf{W}_1$ and $\mathbf{W}_2$ are weight matrices, $\mathbf{b}_1$ and $\mathbf{b}_2$ are bias terms.

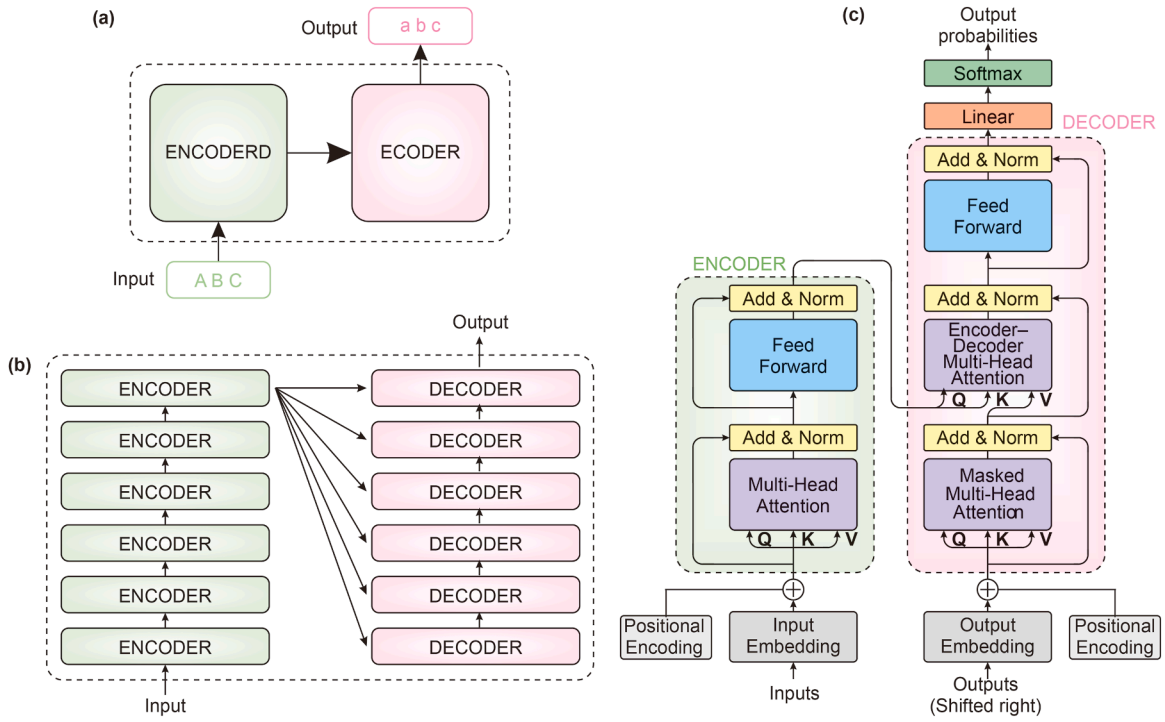

Fig. 2. Transformer model architecture and schematic diagram. **(a)** Single set of encoder and decoder; **(b)** multiple sets of corresponding encoders and decoders; **(c)** schematic diagram of the transformer architecture.

$$\mathrm{FFN}(\mathbf{x}) = \max(0, \mathbf{x}\mathbf{W}_1 + \mathbf{b_1})\mathbf{W}_2 + \mathbf{b_2} \tag{1}$$

A Residual Connection and Layer Normalization is applied after each sub-layer—either a Multi-Head Attention layer or a Feed Forward layer. This design facilitates gradient propagation and prevents model degradation, thereby accelerating the training process. The computation is defined in Equation (2), where Sublayer($\mathbf{x}$) denotes the output of the corresponding sub-layer.

$$\mathrm{LayerNorm}(\mathbf{x} + \mathrm{Sublayer}(\mathbf{x})) \tag{2}$$

It is important to note that the core of the Transformer model is the Self-Attention Mechanism, which is used to compute the dependencies between all positions in the input sequence. This allows the model to capture relationships between any two positions. The computation process is shown in Equation (3), where **Q** (Query), **K** (Key), and **V** (Value) are matrices obtained through linear transformations of the input sequence, and d_k represents the dimension of the Key vector. This equation calculates the similarity between **Q** and **K**, and then uses the resulting similarity scores to perform a weighted sum over **V**, thus generating the final output, where the superscript T denotes matrix transposition.

$$\mathrm{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \mathrm{Softmax}\left(\frac{\mathbf{Q}\mathbf{K}^{\mathrm{T}}}{\sqrt{d_k}}\right)\mathbf{V} \tag{3}$$

Building upon this, the Multi-Head Attention Mechanism is introduced. It consists of a set of h independent Self-Attention modules, where each independent Self-Attention operation is referred to as a head. Each head applies different linear transformations to the Query, Key, and Value matrices. The outputs from these h attention heads are then concatenated and passed through a final linear transformation. The entire computation process is shown in Equations (4) and (5), where $\mathbf{W}_i^{\mathbf{Q}}, \mathbf{W}_i^{\mathbf{K}}, \mathbf{W}_i^{\mathbf{V}}, \mathbf{W}^{\mathrm{O}}$ denote trainable weight matrices. $\mathbf{W}^{\mathrm{O}}$ is the output projection matrix.

$$\mathrm{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \mathrm{Concat}(\mathrm{head}_1, \ldots, \mathrm{head}_h)\mathbf{W}^{\mathrm{O}} \tag{4}$$

$$\mathrm{head}_i = \mathrm{Attention}\left(\mathbf{Q}\mathbf{W}_i^{\mathbf{Q}}, \mathbf{K}\mathbf{W}_i^{\mathbf{K}}, \mathbf{V}\mathbf{W}_i^{\mathbf{V}}\right) \tag{5}$$

3.2.2. Intelligent recognition process

Based on the Multi-Head Self-Attention mechanism and the principles of the Transformer model, the training workflow of the intelligent recognition model proposed in this study is illustrated in Fig. 3. First, the standardized logging response data are fed into the encoder, where positional encoding is applied to inform the model of the sequential order of each logging data point. Upon entering the encoder, the multi-head self-attention mechanism establishes relationships between well logging segments with similar features but located at different depths. The output at each position is then processed through a Feed Forward neural network for nonlinear transformation. Each Multi-Head Attention and Feed-Forward sublayer undergoes residual connection and layer normalization before passing to the next encoder layer. After passing through multiple encoder layers, the original logging inputs are transformed into position-wise multidimensional feature embeddings that retain depth correspondence while incorporating global contextual information. These encoder outputs are directly passed to the decoder as the Key and Value inputs of the encoder–decoder Multi-Head Attention layer, forming the information bridge between the encoder and decoder modules.

In the decoder, the output data generated by the model (i.e., inputs with lamina combination classification labels) are likewise positionally encoded. A Masked Multi-Head Attention layer extracts temporally enhanced features representing contextual

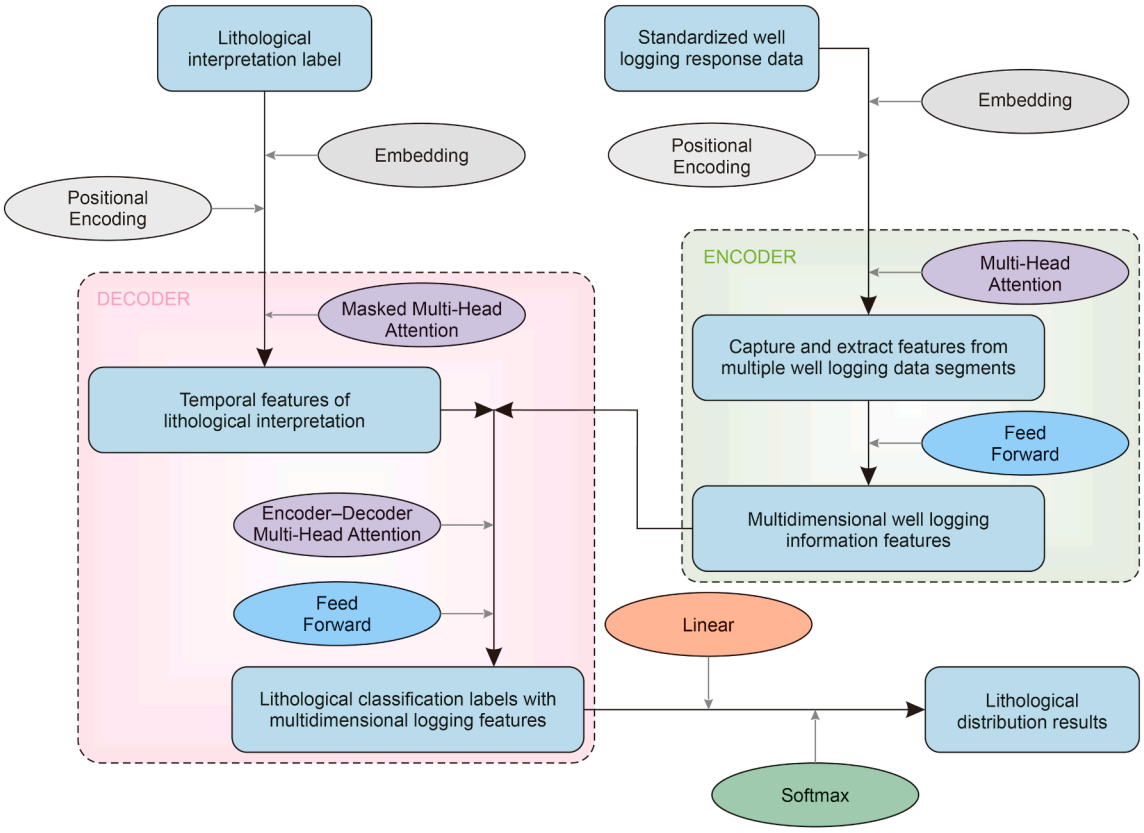

Fig. 3. Training process of lamina combination intelligent recognition based on Transformer model.

information, such as dependencies and transition patterns between lithologies, suitable for characterizing vertically varying stratigraphic features (Han et al., 2024). Subsequently, the multidimensional logging features generated by the encoder are integrated with lamina combination predictions through the Encoder–Decoder Multi-Head Attention mechanism and further processed by a Feed-Forward with residual connections and layer normalization. The decoder outputs are then linearly transformed and passed through a Softmax function to produce the final probability distribution of lamina combination types.

This architecture enables end-to-end mapping from multidimensional logging data to lamina combination sequences with improved accuracy and robustness compared with conventional classification approaches.

4. Results

4.1. Identification of lamina combinations

4.1.1. Single lamina types and characteristics

The upper member of the Xiaganchaigou Formation in the Yingxi area represents a typical lacustrine mixed lithologic reservoir, characterized by complex and diverse mineral compositions. The reservoir rocks are primarily composed of carbonate minerals, terrigenous clastic minerals, clay minerals, minor evaporite minerals, and trace amounts of other components (Fig. 4(a)). Among them, dolomite and calcite exhibit the highest average content, collectively accounting for 49.2% of the total mineralogical composition. Clay minerals constitute approximately 22.6%, while terrigenous clastic minerals—mainly quartz and feldspar—contribute relatively less, with a combined proportion of no more than 20%.

In this study, a lithological classification ternary diagram was established based on the relative mineral content distribution in the study area and previous lithological classification schemes (Chen et al., 2015; Liu et al., 2021; Sun et al., 2023; Zhang, 2000). A total of 215 sample points were statistically analyzed for the content of carbonate minerals, clay minerals, and terrigenous clastic minerals (Fig. 4(b)). The majority of the samples fall within the carbonate- and clay-dominated regions of the ternary diagram, with only a few samples classified as clastic rocks.

(1) Characteristics of lamina mineral components

In sedimentological terminology, laminae are the most fundamental units of stratification and are internally structureless at the macroscopic scale. They form through the contemporaneous deposition of sediments with broadly similar lithologic properties under relatively uniform depositional conditions, and their thickness is typically on the order of a few millimeters. Beds (or bedded structures) are sedimentary units thicker than individual laminae, commonly at the centimeter scale or above, and may consist of multiple laminae or internally homogeneous material (Zhu, 2008). In contrast, non-laminated intervals lack identifiable lamination at the observation scale and therefore do not preserve diagnostic lamina-scale architecture; consequently, they are excluded from lamina combination classification and genetic analysis in this study. In most cases, the development of laminae at different scales can be clearly observed in core samples. Moreover, microscopic examination of casting thin sections allows for the classification of individual laminae according to their mineral compositions. In the upper member of the Xiaganchaigou Formation in the Yingxi area, five major types of laminae are

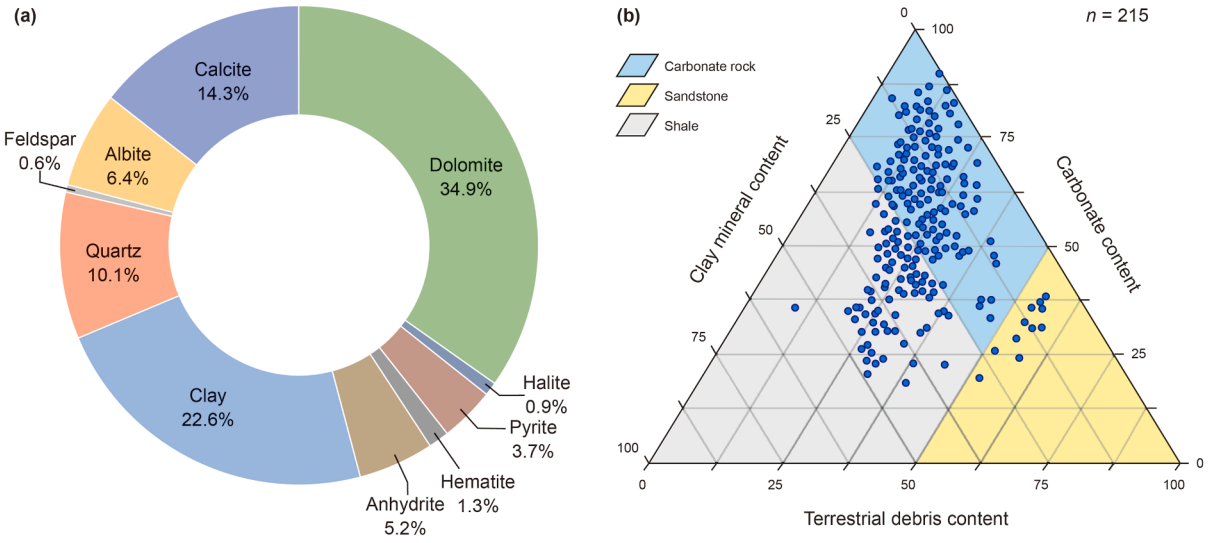

Fig. 4. Petrological characteristics of E_3^2 in the Yingxi area. **(a)** Proportion of average rock mineral content; **(b)** lithological classification triangle.

recognized in the study interval, consisting of carbonate laminae, felsic laminae, clay laminae, and two mixed laminae types (Fig. 5).

Carbonate laminae are primarily composed of calcite and dolomite, with some displaying micritic textures. These laminae are generally thin and often interbedded with clay minerals. At the microscopic scale, clear interlamination between carbonate and clay minerals can be observed due to the significant difference in grain size. Additionally, small amounts of pyrite are commonly present in carbonate laminae samples.

Felsic laminae mainly consist of feldspar, quartz, and minor detrital grains, with grain sizes generally below 0.03 mm, typically forming siltstone. Gradational lamination may develop with vertical variations in felsic content. Well-sorted laminae exhibit sharp boundaries, whereas poorly-sorted ones show diffuse interfaces with dispersed clay minerals. Pyrite commonly occurs along silt-enriched bands.

Clay laminae are predominantly composed of clay minerals, appearing dark brown under the microscope. These laminae are usually thicker and more compact, occasionally containing micritic carbonate minerals and felsic detritus. Black organic matter particles are also observed. The lamina boundaries vary depending on the contact mineralogy, exhibiting straight, continuous, or wavy interfaces.

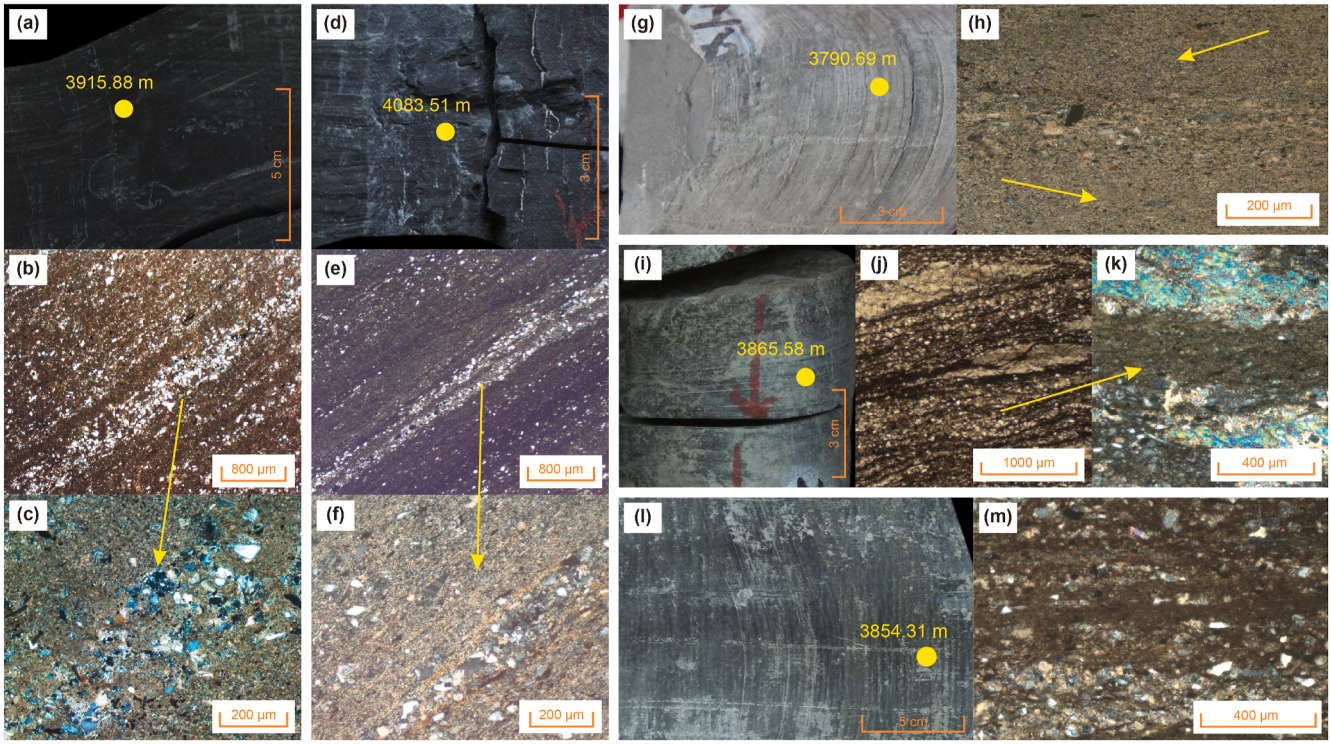

Fig. 5. Single lamina types in the study area. **(a)–(c)** Carbonate laminae, well B43, 3915.88 m; **(d)–(f)** felsic laminae, well B41-2, 4083.51 m; **(g) (h)** clay laminae, well B49-1, 3790.69 m; **(i)–(k)** gypsiferous dolomite laminae, well B41-6-1, 3865.58 m; **(l) (m)** sandy dolomite laminae, well B41-6-1, 3854.31 m. **(a), (d), (g), (i), (l)** Core image; **(b), (e), (j)** microscopic photograph, plain-polarized light (PPL); **(c), (f), (h), (k), (m)** microscopic photograph, cross-polarized light (XPL).

Mixed laminae exhibit diverse diagenetic mineral compositions, including quartz, feldspar, calcite, and gypsum, with no single dominant mineral. Based on mineral assemblages, mixed laminae can be further divided into gypsiferous dolomite laminae and sandy dolomite laminae. The gypsiferous dolomite laminae have complex mineral compositions, with micritic calcite and dolomite distributed in banded patterns. Gypsum crystals are interbedded within the laminae composed of these two minerals, with highly variable development. The lamina boundaries are generally distinct, making their morphology easily observable under the microscope. The sandy dolomite laminae are composed of micritic calcite mixed with siltstone, where the silt content commonly exceeds 30%. These laminae exhibit diverse morphologies, with irregular internal arrangements of various minerals. Clay minerals often appear as flaky aggregates, and the lamina boundaries are typically indistinct.

(2) Characteristics of lamina structural pattern

As a sedimentary structure, laminae record depositional environments and hydrodynamic conditions, with their morphology controlled by variations in paleogeomorphology and paleohydrology. Core observations and borehole imaging logs (FMI_DYN) show that laminae in the upper Xiaganchaigou Formation are well developed, mainly at the centimeter to decimeter scale, and occur predominantly as horizontal layers with locally wavy geometries and scour features.

In terms of sedimentary morphology, the structural patterns of laminae in the study area can be classified into two types based on their density and individual thickness per unit length: laminated layer and bedded layer. Morphologically, laminated layers are relatively thin and display densely packed, regular alternating light and dark banding on borehole imaging logging. These bands represent the interbedding of high-resistivity carbonate rocks (or sandstone) with low-resistivity mudstone layers in thin, closely spaced successions. Individual laminae within this type are typically less than 10 mm thick, with a density generally exceeding 100 laminae per meter.

In contrast, bedded layers are thicker, more variable in form, and often lack clear lithological boundaries. On borehole imaging logging, they appear as abrupt changes in brightness or darkness on a relatively uniform background color, accompanied by noticeable transitional zones. These patterns indicate less distinct interbedding between lithologies with contrasting resistivities. Individual laminae in this category typically range 10–100 mm in thickness, with a lamina density of approximately 10–100 laminae per meter (Fig. 6).

Based on previous studies on the genesis of lacustrine laminae (Guo et al., 2018; He et al., 2025a; Jackson, 1985; Li et al., 2022) and observation of core samples in the study area, the depositional characteristics of the aforementioned two types of laminae can be summarized (Table 1). Under conditions of similar diagenetic mineral compositions, laminated layer is typically associated with periodic high-frequency depositional events, such as seasonal wet–dry alternations and episodic climatic fluctuations in temperature and humidity. The hydrodynamic conditions are generally of moderate to low energy with distinct cyclicity. These laminae are typically deposited in environments such as shallow evaporative platforms or high-salinity settings. In contrast, bedded layer are deposited more continuously, primarily controlled by low-frequency events on centennial to millennial scales. Some turbiditic processes induced by flood events are also observed. The overall hydrodynamic conditions are persistently low-energy. Typical depositional environments include semi-deep to deep lacustrine settings with low salinity or distal turbidite systems. Each single lamina type may correspond to two structural patterns, reflecting differences in genesis under varying depositional environments. Therefore, intervals characterized by multiple single laminae forming distinct lamina combinations often indicate a specific type of depositional environment.

4.1.2. Lamina combination types and characteristics

As an integrated response of sedimentation and diagenesis, laminae exhibit significant differences in reservoir properties—such as petrophysical characteristics and hydrocarbon-bearing capacity—due to variations in mineral composition and subsequent diagenetic alterations. These differences are further manifested in the production performance of different lamina combination intervals during reservoir development (Gao et al., 2013; Liang et al., 2018; Song et al., 2020; Wilhelms et al., 1989). Therefore, the classification and identification of lamina combination types are essential in reservoir studies.

Based on core and thin section observations of the vertical stacking relationships of laminae, combined with macroscopic differences in laminae structural patterns, six lamina combinations were identified (Fig. 7): laminated clay–carbonate rock combination (LCC); bedded clay–carbonate rock combination (BCC); laminated clay–felsic rock combination (LCF); bedded clay–felsic rock combination (BCF); laminated clay–gypsiferous dolomitic rock combination (LCGC); and bedded clay–sandy dolomitic rock combination (BCSC).

Laminated clay–carbonate rock combination (LCC) (Fig. 7(a)): The core exhibits densely spaced, horizontally laminated structures with noticeable color contrasts, indicative of laminated laminae. Under the microscope, distinct interbedding of dark yellow carbonate laminae and deep brown clay laminae is observed, with sharp and continuous boundaries. Carbonate laminae are generally thicker than the clay-rich ones. Bedded clay–carbonate rock combination (BCC) (Fig. 7(b)): The core shows poorly developed horizontal bedding, with varying thickness and muted color contrasts, typical of bedded laminae. Under the microscope, the boundaries between carbonate and clay laminae are indistinct, and carbonate laminae appear discontinuous. Laminated clay–felsic rock combination (LCF) (Fig. 7(c)): This type exhibits distinct, fine, and uniform laminae in the core, showing typical strong layering. Under the microscope, felsic mineral laminae (mainly quartz and feldspar) appear in well-developed, banded forms with clear and stable boundaries against clay minerals. Bedded clay–felsic rock combination (BCF) (Fig. 7(d)): The core reveals scoured surfaces and wedge-shaped cross-laminations, with irregular laminae arrangements lacking clear interbedding patterns, representing bedded features. Microscopically, felsic laminae are discontinuously developed and often interbedded with clay minerals. Lenticular fine-grained sandstone structures are visible, interpreted as the result of high-energy depositional processes. Laminated clay–gypsiferous dolomitic rock combination (LCGC) (Fig. 7(e)): The core displays densely spaced fine horizontal stripes, typical of laminated structures. Microscopically, multiple minerals are arranged in banded distributions. Despite the mineralogical complexity within the gypsiferous dolomitic laminae, the boundaries with adjacent clay-rich laminae are clear, and overall laminae continuity is high. Bedded clay–sandy dolomitic rock combination (BCSC) (Fig. 7(f)): The core shows thick, irregular, light-gray bands without consistent patterns. Under the microscope, the sandy dolomitic laminae appear weakly banded and discontinuous, with dispersed mineral grains and indistinct boundaries with adjacent clay minerals, often forming transitional contacts.

To reduce interpretational subjectivity in lamina combination classification, Table 2 incorporates quantifiable mineralogical

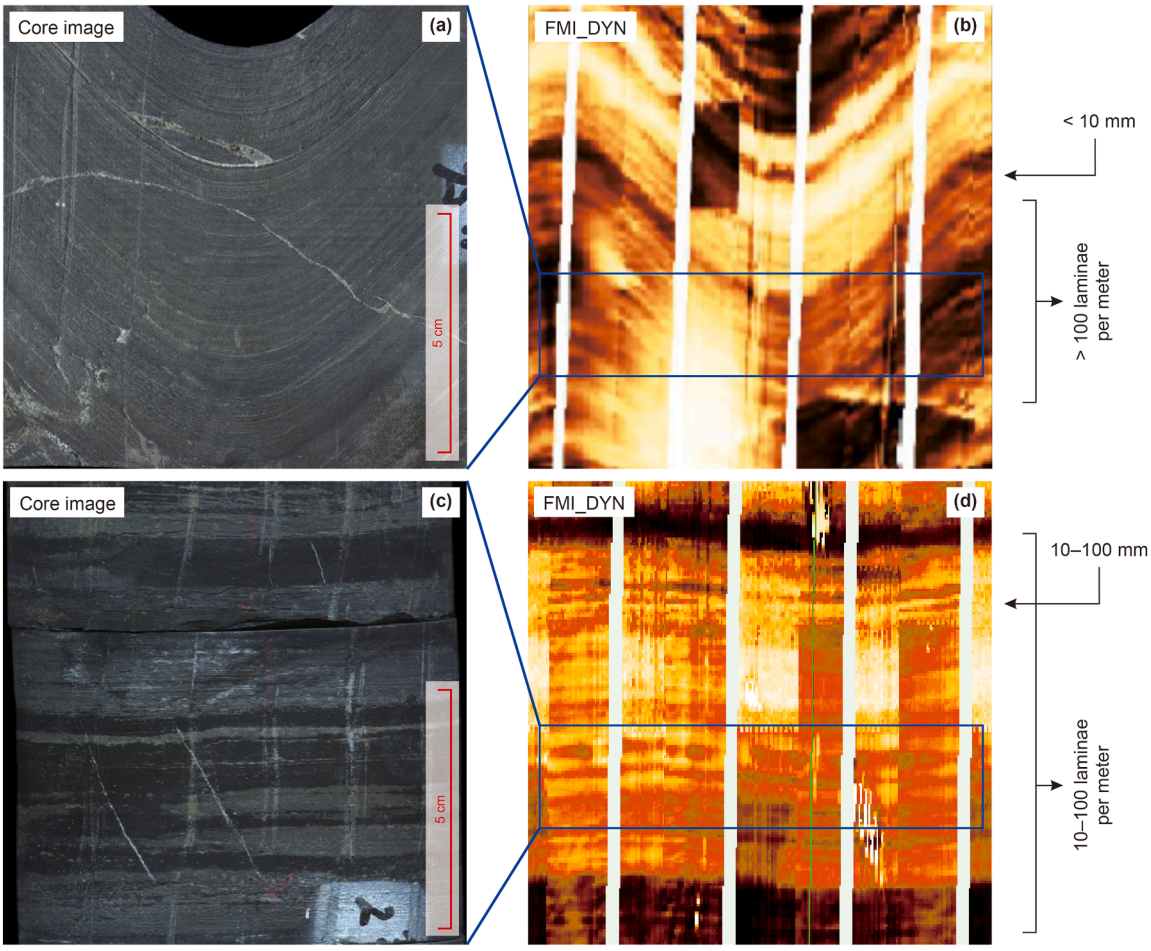

Fig. 6. Classification of lamina structural patterns at different scales. **(a)** Well B53-1, 4046.19–4046.27 m, core image (surface rolling scan); **(b)** well B53-1, same position, FMI_DYN; **(c)** well B41-2, 4077.89–4077.97 m, core image; **(d)** well B41-2, same position, FMI_DYN.

Table 1
Comparison of laminated structures under identical diagenetic mineral assemblages.

Structural type	Laminated layer	Bedded layer
Thickness	Millimeter–centimeter scale	Centimeter–decimeter scale
Lamina boundaries	Sharp and distinct, marked by abrupt changes in grain size or mineralogy	Gradational and diffuse, with gradual grain size transitions
Depositional frequency	High-frequency (seasonal to interannual scale)	Low-frequency (centennial to millennial scale) or flood events
Hydrodynamic conditions	Moderate to low energy, periodic	Persistently low energy
Typical environment	Shallow-water (evaporative) platform, high-salinity settings	Semi-deep to deep lacustrine environments, low salinity, distal turbidites

parameters, including the relative abundance of carbonate, felsic, clay, evaporite-related, and sandy dolomitic components. The reservoir development parameters also vary systematically among the six types of lamina combination intervals, as reflected by the distributions of porosity, mineral brittleness index (Huo et al., 2018; Jarvie et al., 2007), and TOC content summarized in Table 2 (Zhang et al., 2026). LCC and LCF exhibit relatively high brittleness (0.80–0.97) and moderate to high TOC contents (0.75%–1.60%), indicating favorable fracturing potential and hydrocarbon generation capacity. BCC and BCF show comparatively higher porosity but lower organic matter enrichment. In contrast, LCGC displays the lowest porosity, brittleness, and hydrocarbon generation capacity, reflecting poor reservoir quality under intense evaporation and diagenetic compaction. The classification from a sedimentary-genetic perspective demonstrates the controlling influence of depositional origin on reservoir evaluation and development potential, thereby supporting the rationality of the proposed lamina combination scheme.

In lacustrine mixed rocks, laminae commonly exhibit complex vertical stacking patterns, ranging from simple binary alternations to compositionally mixed assemblages. In this study, lamina combinations are classified based on the dominant association between two lamina types under specific lamina structural patterns (laminated or bedded). Mixed lamina types, such as gypsiferous dolomite laminae and sandy dolomite laminae, inherently incorporate multiple mineral components and thus partially capture higher-order compositional complexity. At the resolution of well logs, intervals with complex alternations are therefore represented by their dominant lamina combination, with subordinate components treated as internal variations.

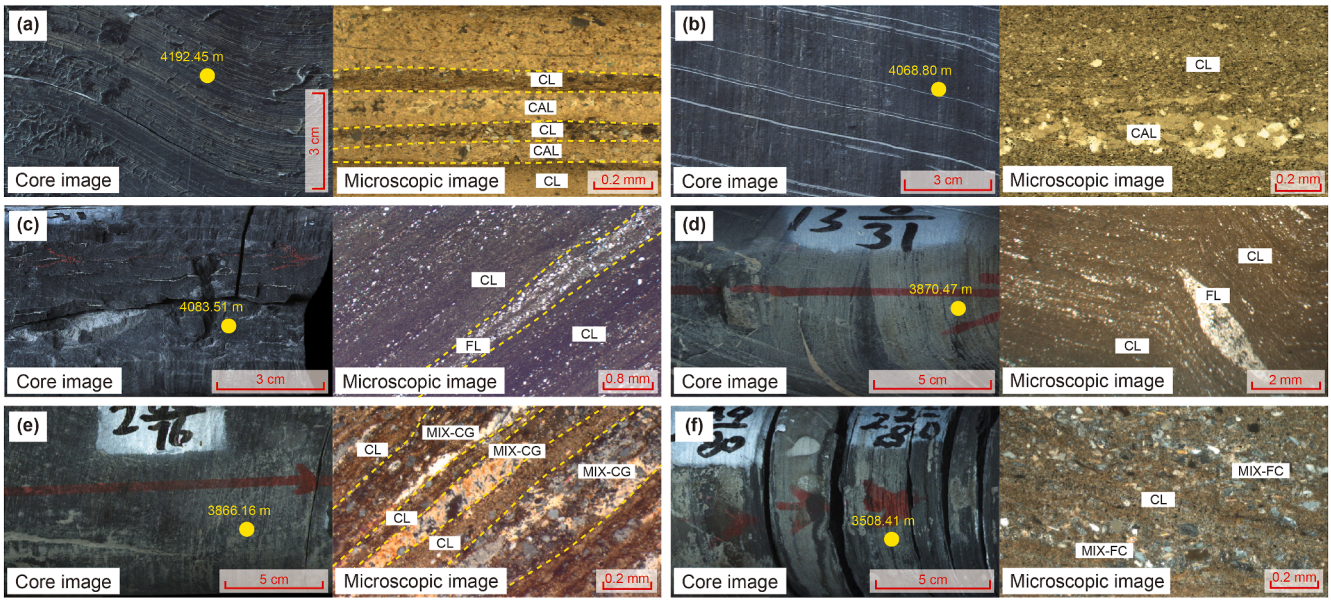

Fig. 7. Lamina combination types of the study area. **(a)** LCC, well B41-2, 4192.45 m, XPL; **(b)** BCC, well B41-2, 4068.80 m, PPL; **(c)** LCF, well B41-2, 4083.51 m, PPL; **(d)** BCF, well B49-1, 3870.47 m, PPL; **(e)** LCGC, well B41-6-1, 3866.16 m, XPL; **(f)** BCSC, well B38-2, 3508.41 m, XPL. Note: CL: clay laminae; FL: felsic mineral laminae; CAL: carbonate laminae; MIX-CG: gypsiferous dolomitic laminae; MIX-FC: sandy dolomitic laminae.

Table 2
Mineralogical and reservoir properties of different lamina combination types.

Type	Mineralogical characteristics, vol%	Porosity, %	Mineral brittleness, dimensionless	TOC, %
LCC	Dolomite ≥ 25; calcite ≥ 10; gypsum ≤ 10	8.5–12.0	0.80–0.95	1.10–1.60
BCC	Clay ≥ 25; carbonate (dolomite + calcite) ≥ 35	8.0–15.0	0.60–0.75	0.50–0.77
LCF	Quartz ≥ 20; clay ≥ 25; gypsum ≤5	5.0–8.5	0.80–0.97	0.75–0.85
BCF	Quartz ≥ 18; clay ≥ 28; carbonate (dolomite + calcite) ≤ 40	6.5–10.0	0.60–0.80	0.30–0.50
LCGC	Gypsum ≥ 20; carbonate (dolomite + calcite) ≥ 25; quartz ≤20	2.0–5.0	0.40–0.60	0.15–0.30
BCSC	Dolomite ≥ 18; quartz ≥ 18; gypsum ≤ 10	5.5–8.0	0.55–0.75	0.35–0.75

4.2. Response characteristics of lamina combination logging

Based on the classification of lamina combinations, the logging characteristics of intervals with different lamina combinations were statistically analyzed and summarized. It should be noted that individual laminae are typically only a few millimeters to centimeters thick, making them difficult to characterize using conventional logging data. However, the thickness of lamina combination intervals generally ranges 0.4–1.5 m, a scale that can be effectively captured by both conventional well-logging and LithoScanner measurements. The logging dataset includes both conventional well logs and core scanning log data. In the domain of conventional logging, five types of logs were utilized: GR, AC, LLD, LLS and DEN. The core scanning log data provide the percentage variations of five mineral components: gypsum, dolomite, calcite, quartz, and clay. Through statistical analysis, the logging response characteristics are shown in the boxplot below (Fig. 8).

The figure presents the distribution characteristics of both conventional well-logging data and LithoScanner data, with each subplot displaying the values for different lamina combinations along the x-axis. In the conventional well logs, for the GR curve, the LCGC type is mainly distributed within the range of 65–98 API, which is significantly lower than other lamina combinations due to its high gypsum content. LCC and BCC, both carbonate-dominated lamina combinations, exhibit slightly lower GR values than LCF and BCF, whereas BCSC falls between the two groups. In the AC curve, the LCC and BCC types show a similar distribution, primarily within the range of 185–208 μs/m. LCF and BCF range slightly higher, 190–218 μs/m. LCGC and BCSC are distributed in the ranges 188–200 μs/m and 185–210 μs/m, respectively. Both ranges fall largely within the intervals of the other four types, indicating poor differentiation. For the DEN curve, there are noticeable differences between laminated and bedded combinations. The three laminated combinations—LCC, LCF, and LCGC—have density values range of 2.32–2.54 g/cm^3, whereas the other three bedded types are generally lower, ranging from 2.24 g/cm^3 to 2.32 g/cm^3. In terms of resistivity (LLS and LLD), BCC, LCF, and BCF exhibit broader ranges, typically 5–40 Ω m. The remaining three types fall within a narrower range of 5–22 Ω m. Moreover, for all lamina combinations, LLD values are generally higher than LLS values.

For the LithoScanner data: In the clay component, LCGC shows a lower clay content (approximately 15%–30%), while the other five types are relatively similar, ranging 20%–42%. In the Quartz component, LCF and BCF show significantly higher contents (23%–35%), followed by BCSC (19%–28%). LCC, BCC, and LCGC exhibit lower quartz contents, generally ranging 9%–21%. In the calcite component, BCC stands out with a higher content range of 17%–31%, while the remaining five types range 4%–24% with smaller differences. In the dolomite component, LCC exhibits a clearly higher content (30%–47%), whereas the other five types range between 15% and 35%. In the gypsum component, LCGC shows a notably higher content of 18%–32%, with distinct separability. The remaining five combinations have relatively low gypsum contents, ranging 3%–12%.

The above analysis demonstrates that the well-log response values of different lamina combinations vary significantly. Moreover, the degree of differentiation between lamina combinations

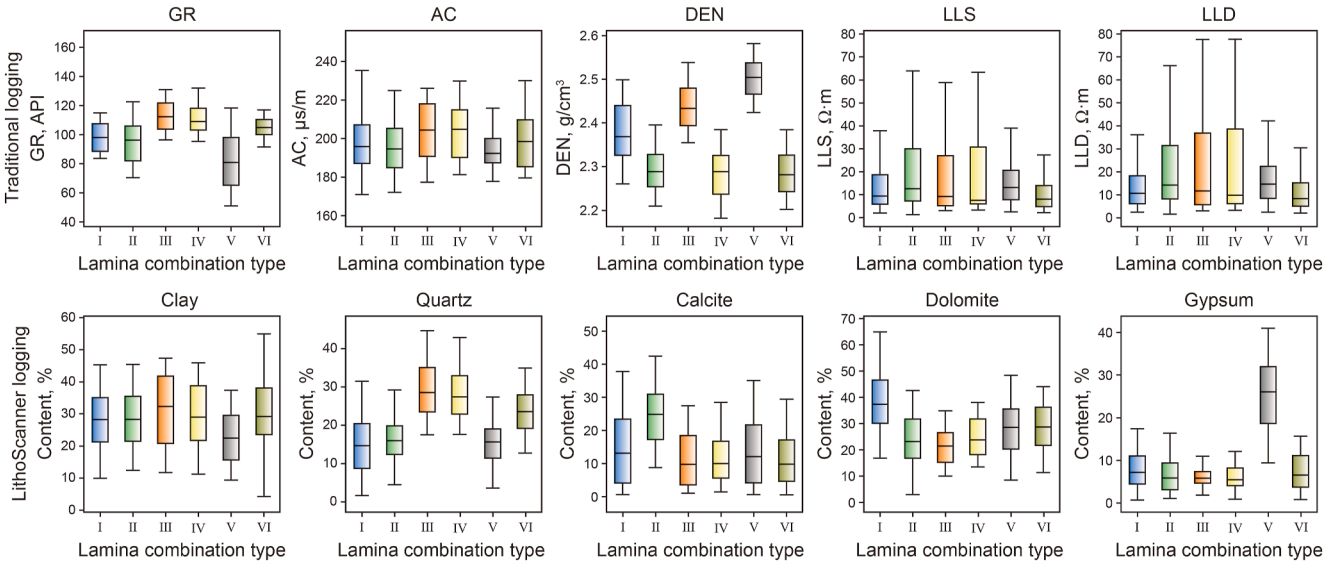

Fig. 8. Boxplots of various logging responses for different lamina combination types. I–LCC; II–BCC; III–LCF; IV–BCF; V–LCGC; VI–BCSC.

depends on the type of logging data, with most logs only reflecting part of the variability among the combinations. This highlights the limitations of traditional interpretation methods, such as cross-plot classification, in accurately distinguishing between lamina types. These findings further underscore the necessity of employing multi-parameter fusion-based intelligent algorithms for effective well-log interpretation in this study.

To quantitatively verify the discriminability of the logging responses, the mean, standard deviation, and separation index (D-value, defined as the ratio of inter-class mean difference to standard deviation) were calculated for each parameter. As shown in Table 3, the D-values of GR, clay, and gypsum parameters all exceed 2.7, indicating that they are the most sensitive curves and can serve as key indicators for lamina combination identification. The D-values of AC, DEN, LLS, LLD, and carbonate components (calcite and dolomite) range between 2.3 and 2.6, suggesting moderate sensitivity. Overall, the dataset exhibits good separability and is suitable for intelligent identification.

4.3. Intelligent interpretation of lamina combination

After training the Random Forest and Transformer algorithms using the training dataset, to further evaluate the cross-well generalization stability of the proposed model, the trained algorithms were applied to two representative cored wells for independent validation. Their performance was evaluated on the testing dataset for comparison. In the process of Transformer algorithm recognition, a grid search with 500 iterations was conducted to automatically optimize the hyperparameters. The search space included learning rate (0.01–0.0001), batch size (16, 32, 64), dropout rate (0.1, 0.2, 0.3, 0.5), and weight decay (0, 0.01, 0.1). The objective was to maximize the F1 score on the test dataset. The optimal hyperparameters obtained were as follows: learning rate = 0.001, batch size = 32, dropout rate = 0.5, weight decay = 0.01, with 500 training epochs. The Random Forest algorithm was implemented using its default parameter settings. Due to the presence of non-laminated intervals within the strata of the study area, it was first necessary to distinguish between laminated and non-laminated intervals. The recognition accuracies for this task differed between the two models: the Random Forest achieved an accuracy of 84%, while the Transformer model achieved 89%, indicating that the latter was more effective in identifying laminated intervals.

Regarding the overall classification accuracy of lamina combinations, the Random Forest achieved 76%, whereas the Transformer reached 84%, once again demonstrating the superior performance of the Transformer model in identifying all types of lamina combinations. The recognition results for two cored wells from the test dataset are illustrated in Fig. 9 (in Fig. 9 and the subsequent single-well identification results, blank intervals in the lamina interpretation track represent extremely weakly laminated or massive intervals in which lamina combinations are not

Table 3
D-value and sensitivity assessment of logging parameters.

Parameter	Unit	The mean	Standard deviation	Range of variation	D-value	Sensitivity evaluation
GR	API	100.58	11.64	33	2.84	High
AC	μs/m	198.33	5.12	12.5	2.44	Medium
DEN	g/cm^3	2.36	0.09	0.23	2.43	Medium
LLS	Ω·m	9.83	2.34	5.5	2.35	Medium
LLD	Ω·m	11.22	2.6	6.5	2.5	Medium
Clay	%	28.17	3.46	10	2.89	High
Quartz	%	21.4	5.99	13.5	2.25	Medium
Calcite	%	13.87	5.87	15.2	2.59	Medium
Dolomite	%	27.02	6.26	16	2.56	Medium
Gypsum	%	9.58	7.59	20.5	2.7	High

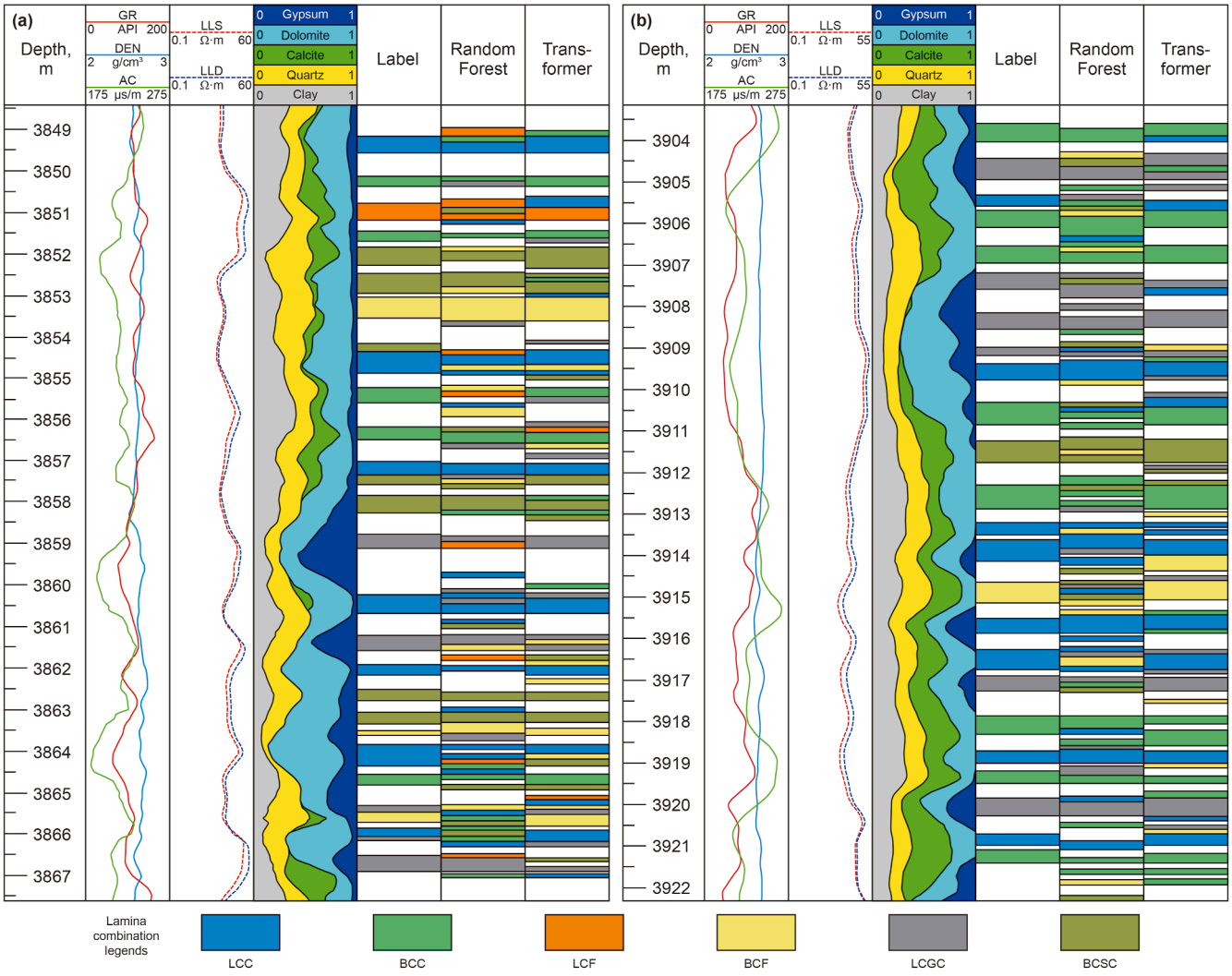

Fig. 9. Comparison of lamina combinations identified by different models and actual lamina combinations with logging curves in two representative cored wells from the testing dataset. **(a)** Well B41-6-1, 3848.45–3867.56 m; **(b)** well B43, 3903.10–3922.20 m.

developed. As this study focuses on the development characteristics and intelligent recognition of lamina combinations, these intervals are outside the scope of the present work). The consistent recognition performance observed across different wells indicates that the Transformer model maintains stable performance across different wells rather than overfitting to a single test case.

Overall, the Transformer model exhibits markedly superior interlayer continuity and boundary-matching accuracy compared with the Random Forest algorithm, particularly in intervals with fine-scale and highly variable laminae. It demonstrates greater stability and consistency, while the Random Forest shows evident discontinuities and misclassifications, especially in the middle to lower sections with recurrent lamina combinations. These results confirm that the Transformer provides higher recognition accuracy and robustness in handling heterogeneous and rapidly varying laminae within lacustrine mixed sedimentary reservoirs.

To further assess cross-well generalization, we additionally evaluated the model under different well training–testing configurations, where each well was held out in turn as an independent test set. Performance was summarized by the range of accuracy and macro-F1 across these configurations (Table 4). The accuracy and macro-F1 ranges indicate that, regardless of which well is used as the blind well, the performance variations remain within an acceptable range, suggesting that both algorithms exhibit a certain degree of generalization capability for this task.

Table 4
Cross-well generalization performance based on leave-one-well-out cross-validation.

Model	Accuracy range	Macro-F1 range
Random Forest	0.73–0.78	0.57–0.61
Transformer	0.81–0.86	0.60–0.66

Note: Accuracy and macro-F1 ranges reflect performance variability observed across independent cored wells under different well training–testing configurations. These ranges are summarized to illustrate cross-well generalization stability.

5. Discussion

5.1. The genesis and vertical evolution of laminated combination

Following extensive analysis of lamina combinations from numerous cored wells in the study area, it became possible to investigate the relationship between different lamina combinations and their corresponding depositional environments on a single-well scale. Maximum Entropy Spectral Analysis (MESA) was

applied to the GR logging curve to generate the INEPFA-GR curve, a reconstructed signal that enables the identification of sedimentary cycles through the segmentation of trend intervals and inflection points (Liu et al., 2024; Lu et al., 2007).

In the saline lacustrine setting of the study area, the types of mixed sediments are primarily controlled by water salinity, and are less influenced by other environmental factors such as temperature (Jiang et al., 2019). The commonly used paleosalinity proxy, the Sr/Ba ratio, has certain limitations. Strontium can originate not only from evaporative concentration in saline or brackish environments but also from terrigenous input (e.g., feldspars and clay minerals). Moreover, the enrichment of carbonate minerals may increase Sr concentrations while simultaneously decreasing Ba levels (Wei et al., 2025; Zhu et al., 2023). Additionally, this ratio lacks a linear scale response to salinity, making it ineffective for distinguishing subtle variations between mildly and highly saline conditions (Li et al., 2024b). Therefore, relying solely on the Sr/Ba ratio for salinity reconstruction is not accurate in this context. Given the stronger correlation between redox conditions and salinity in this region (Jiang et al., 2017, 2019; Tribovillard et al., 2006), the ratios of redox-sensitive elements—V/(V + Ni) and Th/U—were used to infer depositional environments. Lower values of these two ratios indicate oxic conditions, while higher values suggest anoxic or reducing environments. Combined with sedimentary cycle division derived from the INEPFA-GR curve, this proxy enables the interpretation of paleowater depth and paleosalinity conditions. By integrating well-log characteristics and lamina combination classifications, the relationships between different lamina combinations and depositional environments were established (Fig. 10). The vertical alternation of these lamina combinations closely corresponds to the intermittent humid intervals within the overall aridification process that occurred in the western Qaidam Basin during the late Oligocene. Such climatic fluctuations led to the periodic expansion and contraction of the lake, thereby forming the cyclic sequences of carbonate, siliciclastic, and gypsiferous laminae within the Xiaganchaigou Formation.

At the tops of progradational (normal) cycles or the bases of retrogradational (reverse) cycles, samples generally exhibit higher redox-sensitive parameter values, indicating more reducing conditions typically associated with deeper lakewater and wetter climatic settings. In contrast, at the tops of retrogradational cycles or the bases of progradational cycles, the redox parameter values are lower, suggesting more oxidizing environments linked to shallower lakewater and drier climatic conditions. From the perspective of lamina combination types, LCF and BCF—both characterized by higher clastic content—are more commonly developed under humid conditions. In such settings, increased external detrital input supplies clastic particles into the lake system. LCC and BCC, which are rich in carbonate content, are typically formed in semi-humid to semi-arid transitional zones. In some downward (retrogradational) cycle segments, there is a clear stratigraphic succession from BCC to LCC. Additionally, LCGC tends to develop under strongly arid conditions, where intense evaporation facilitates its formation. In contrast, BCSC usually forms in environments representing a transitional overlap between humid and arid conditions.

By integrating lamina distribution patterns within single-well sedimentary cycles with previous studies on carbonate-dominated lacustrine mixed rocks (Li et al., 2022; Wang et al., 2020), the genesis of each lamina combination type can be reasonably interpreted. The genetic models of the six lamina combination types in the study area are illustrated in Fig. 11.

Fig. 11(a) represents the depositional environment of LCGC, characterized by the shallowest water depth and arid climatic conditions. There is no clastic input, and the water is highly susceptible to seasonal salinity stratification. High salinity causes gypsum precipitation and the formation of gypsiferous dolomitic laminae, whereas clay-rich laminae accumulate during cooler or lower-salinity periods.

Fig. 11(b) illustrates the environment of LCC. The water depth is shallow and the climate semi-arid, with negligible external clastic input. Similar to LCGC, seasonal salinity stratification is prevalent, though the overall salinity is lower. Carbonate minerals dominate the precipitation under high-salinity conditions, while clay minerals accumulate during more balanced periods.

Fig. 11(c) depicts the BCC environment, where the water is relatively deeper and the climate transitions from semi-arid to semi-humid. Clastic input remains minimal. Due to the greater water depth, seasonal salinity stratification is less pronounced. Instead, extended periods of either dry or humid conditions dominate. During drier periods, high salinity promotes carbonate precipitation; during wetter intervals, lower salinity favors clay deposition. Longer, more gradual climatic transitions produce bedded layer.

Fig. 11(d) corresponds to the depositional environment of BCSC. The water is relatively deep, and the climate is semi-humid, with minor clastic input. In such deep settings, seasonal salinity stratification is again unlikely. During phases of stronger clastic influx, fine-grained siliciclastic minerals dominate deposition; during weaker influx stages, clay minerals are predominant. Minor carbonate precipitation occurs under varying salinity conditions, leading to a mixed clay–sandy carbonate lamina combination.

Fig. 11(e) and (f) represent the depositional environments of LCF and BCF, respectively. Both are associated with deeper water, humid climates, and ample clastic supply. During the rainy season, abundant sediment influx results in deposition of fine-grained siliciclastic laminae, while the dry season is dominated by clay-rich laminae. The consistently low salinity due to heavy precipitation suppresses carbonate mineral precipitation. LCF reflects seasonal variations, characterized by thin, rapidly changing individual laminae. In contrast, BCF reflects flood events, where lamina thickness varies greatly and contacts between laminae are more complex.

In summary, lamina development in hypersaline lacustrine basins is strongly controlled by depositional environment changes. Ideally, a full sequence of lamina combinations can be recognized within a short-term cycle, allowing inference of their spatial distribution. This relationship not only supports the interpretability of the lamina combinations within a sedimentary framework but also indirectly demonstrates the rationality and geological significance of the proposed classification scheme. Moreover, the vertical variation of lamina combination characteristics throughout the Xiaganchaigou Formation, from bottom to top, reflects a continuous evolutionary process of the lake—from an open freshwater stage to a closed evaporative stage—predominantly controlled by the progressive aridification during the Oligocene.

5.2. Advantages of transformer method

5.2.1. Evaluation of intelligent recognition methods

In addition to overall accuracy, the classification performance of machine learning methods is commonly evaluated using precision, recall, and F1-score, which assess the classification accuracy for each individual label. Precision (Equation (6)) refers to the proportion of correctly predicted samples of a specific class out of all samples predicted to belong to that class in the test set, indicating the accuracy of the prediction for that class. Recall (Equation (7)) refers to the proportion of correctly predicted samples of a specific class out of all actual samples of that class, where a higher

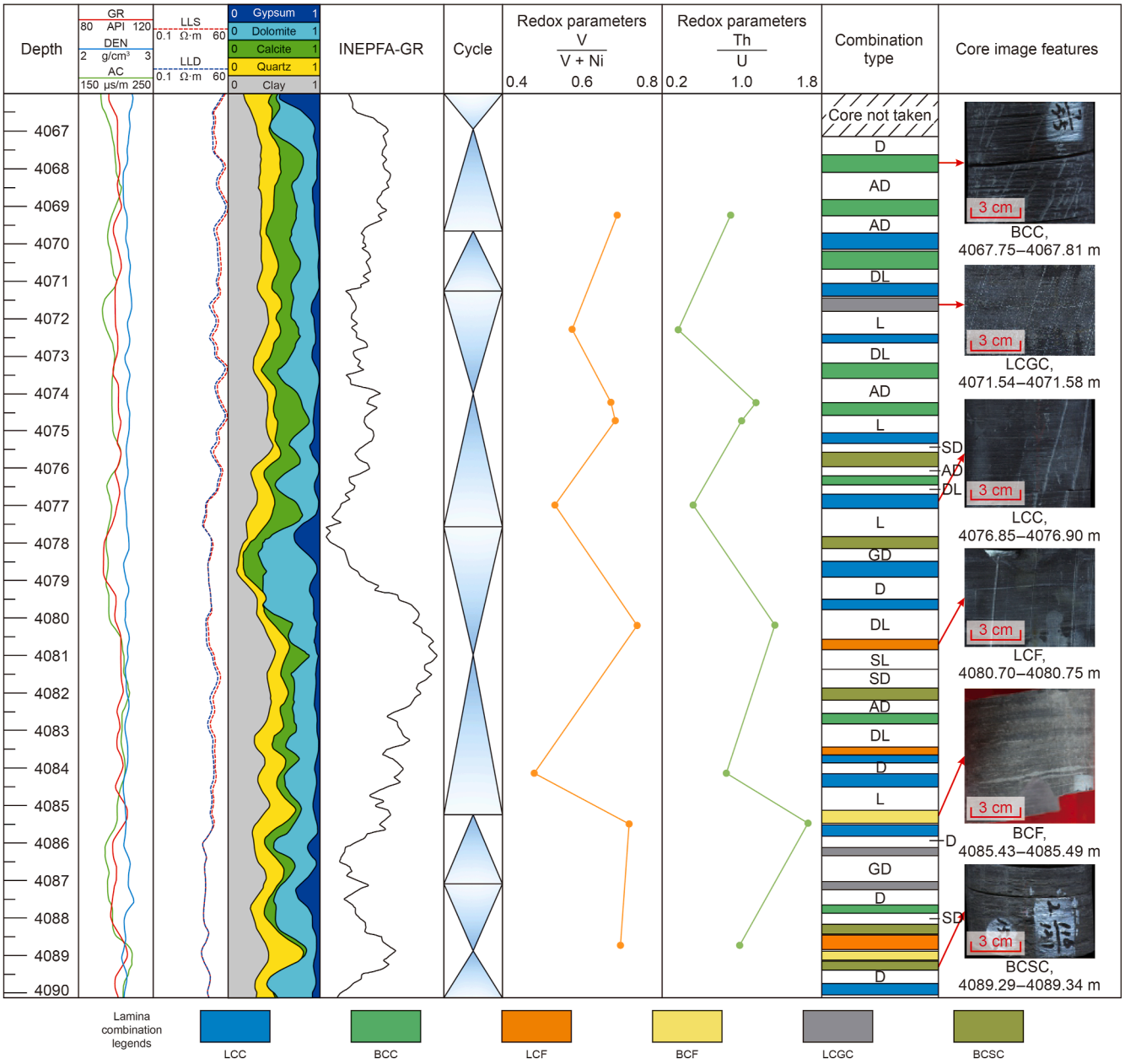

Fig. 10. The relationship between vertical distribution, sedimentary cycle, and paleoenvironment of single-well lamina combinations (taking well B41-2 as an example, 4066–4090.15 m). Lithology abbreviations for non-laminated intervals are as follows: D–dolomite; L–limestone; DL–dolomitic limestone; AD–argillaceous dolomite; SD–sandy dolomite; SL–sandy limestone; GD–gypsiferous dolomite. These lithologic labels are shown only to ensure the completeness of sedimentary cycle representation and are not included in lamina combination classification.

value implies fewer misclassified instances. The F1-score (Equation (8)) represents the harmonic mean of precision and recall, and reflects the overall performance of the classifier in balancing both metrics. Where TP represents the number of samples in the test set that are correctly identified as the target class, TN denotes the number of samples that do not belong to the target class and are correctly classified as such, FP represents the number of samples that are incorrectly identified as the target class, and FN denotes the number of samples belonging to a given class that are incorrectly predicted.

$$P = \frac{TP}{(TP + FP)} \tag{6}$$

$$R = \frac{TP}{(TP + FN)} \tag{7}$$

$$F1 = 2 \times \frac{(P \times R)}{(P + R)} \tag{8}$$

Fig. 12 illustrates the performance of different lamina combination types in terms of precision, recall, and F1-score using two intelligent algorithms. In terms of precision, the Transformer algorithm achieved values above 80% for all lamina combination types. Except for LCC, the Transformer significantly outperformed the Random Forest in the precision of the remaining five types,

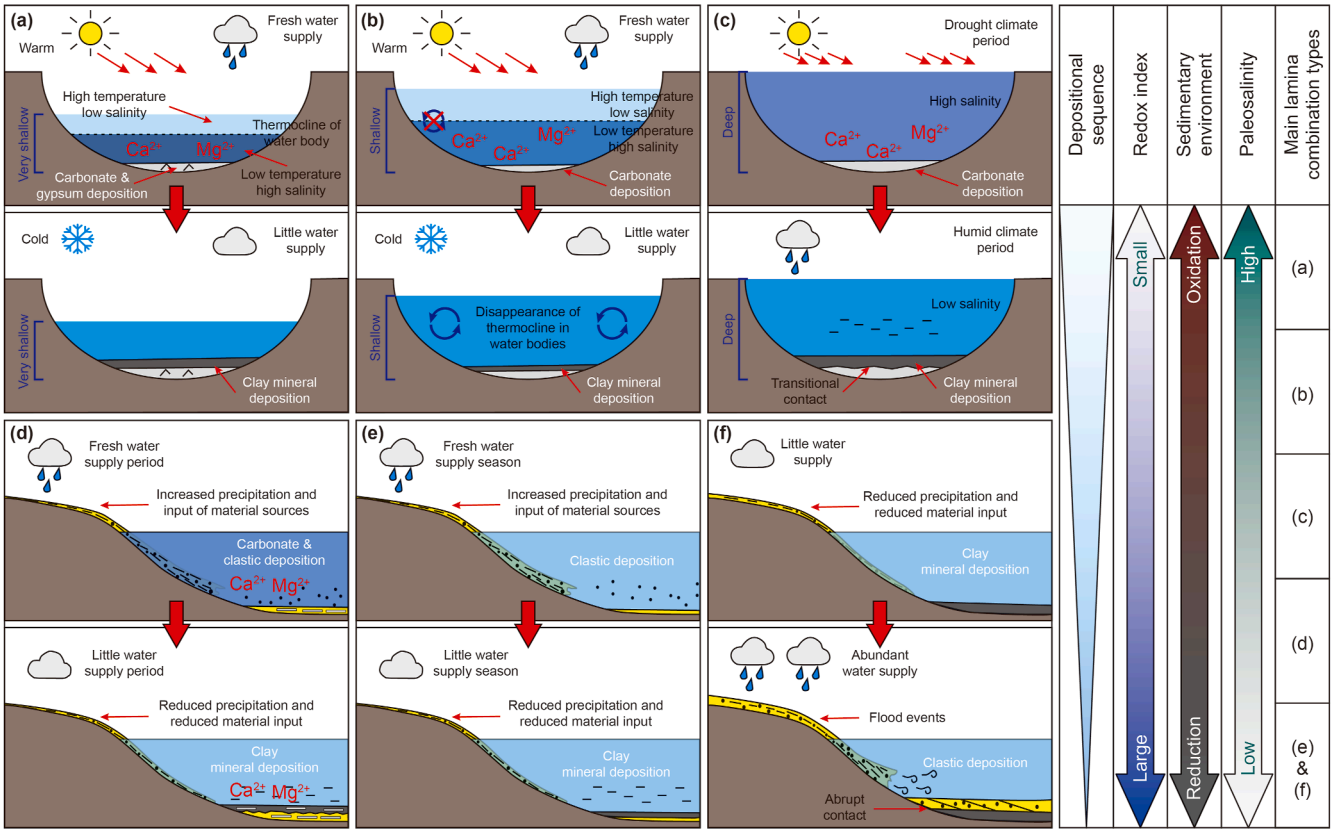

Fig. 11. Depositional settings and vertical stacking patterns of the six lamina combination types. **(a)** Generation environment of LCGC; **(b)** generation environment of LCC; **(c)** generation environment of BCC; **(d)** generation environment of BCSC; **(e)** generation environment of LCF; **(f)** generation environment of BCF.

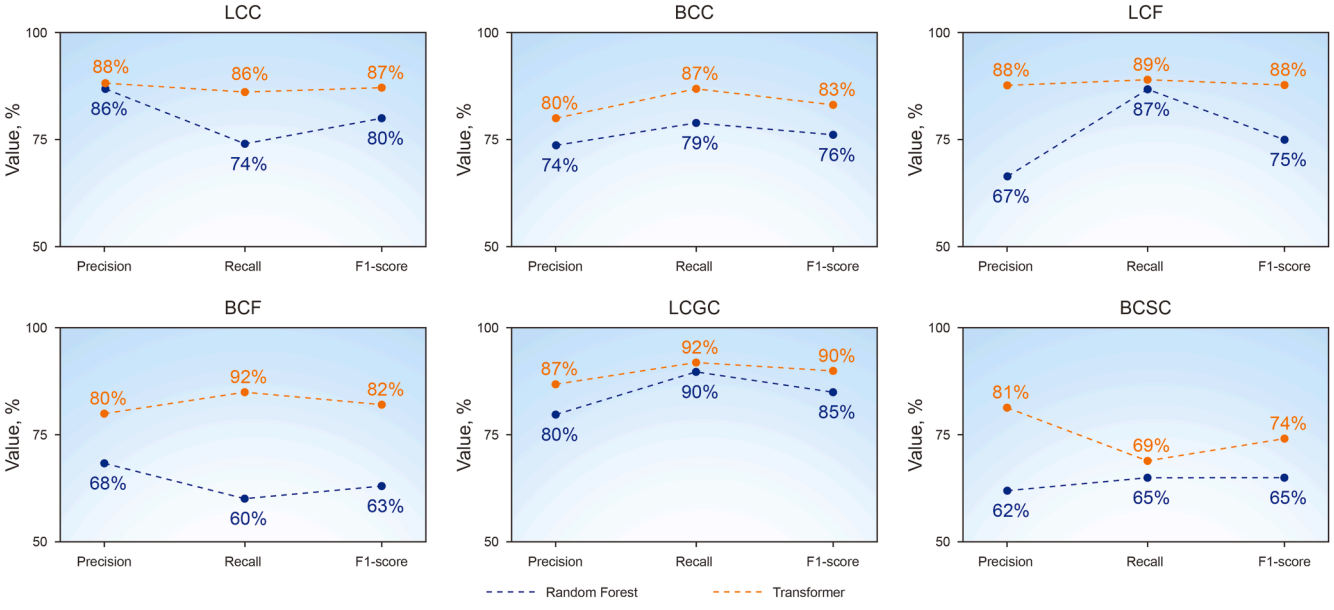

Fig. 12. Comprehensive evaluation of lamina combination identification under different models.

with LCF showing a remarkable improvement of 21 percentage points. This may be attributed to the relatively larger number of LCC samples, while the other types had limited data. Although data augmentation was applied, it could not fully mitigate the issue of data scarcity. Nevertheless, the Transformer exhibited a strong ability to extract information from limited samples, resulting in better training performance and higher classification accuracy. With regard to recall, the Transformer achieved values between 86% and 92% for most lamina combinations, except for BCSC, which had a relatively low recall of 69%. This may be due to the indistinct logging response characteristics and limited sample size of this combination type. Overall, the Transformer outperformed Random Forest in recall, with improvements exceeding 10 percentage points for some types. This indicates that the Transformer

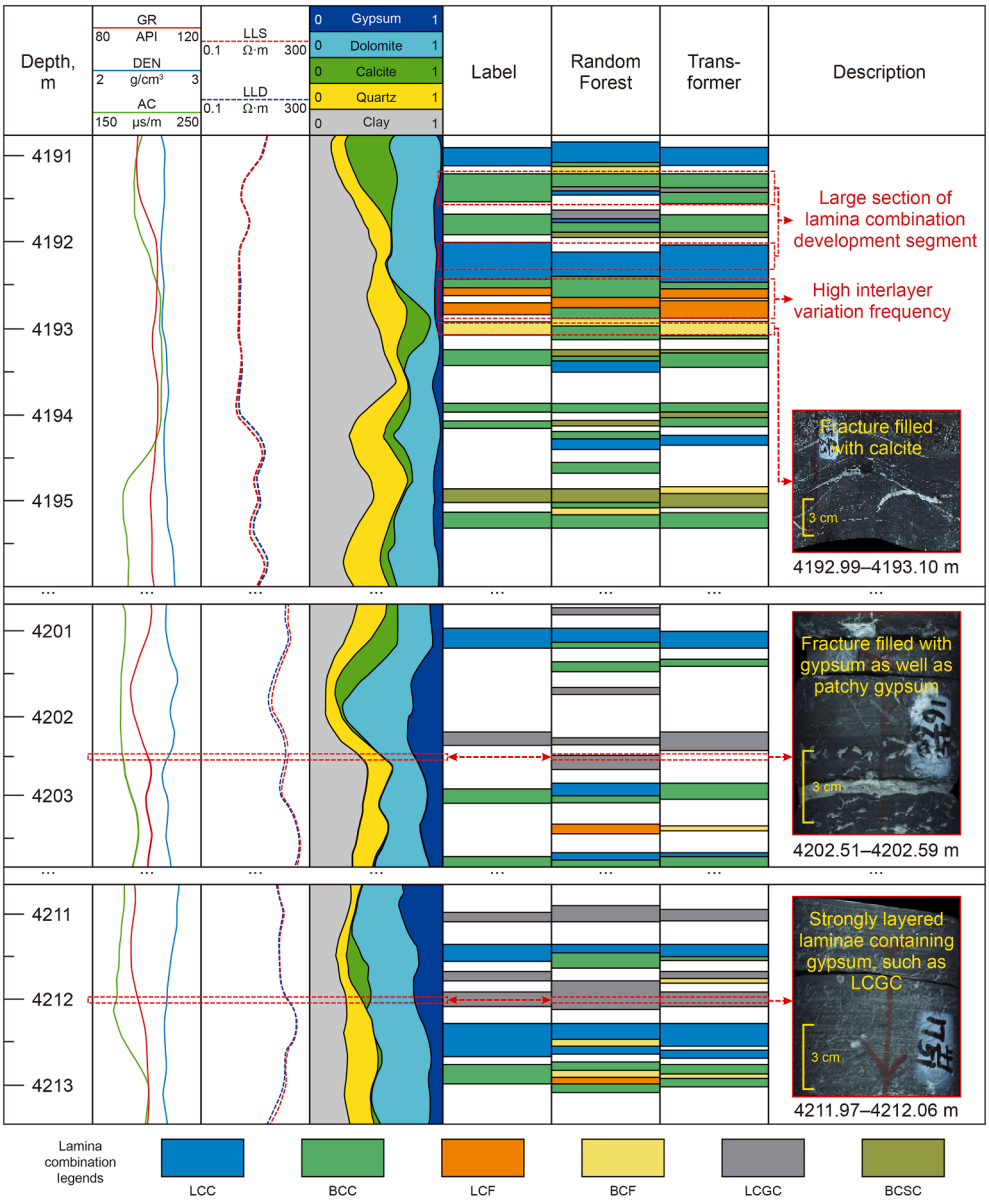

Fig. 13. The advantages of intelligent recognition methods for cored wells (taking well B41-2 as an example).

missed fewer target labels, thereby offering broader coverage. In terms of the F1-score, all values except for BCSC (74%) exceeded 82%, with a variation range of 74%–90%. In contrast, the F1-scores of Random Forest were mostly below 80% and showed a notable gap compared to the Transformer. These results suggest that the Transformer model produced fewer false positives and false negatives, demonstrating superior recognition and classification capabilities for diverse and complex lamina combination types.

5.2.2. Analysis of algorithm advantages

Compared with traditional machine learning algorithms, the Transformer shows significant improvements in the quantitative analysis results on the test dataset. A key advantage of the Transformer is its ability to capture both short-term patterns and long-term trends in time-series data (Ahmed et al., 2023; Wen et al., 2023). This makes it well-suited for learning the variations in logging responses of the same lamina combination type at different depths, effectively addressing the mineralogical complexity inherent in lamina combinations. Moreover, the Transformer demonstrates enhanced learning capability on limited-sample labeled data. The Transformer's hierarchical attention mechanism adaptively focuses on key intervals across multiple scales, overcoming the Random Forest's need for manual feature tuning. This leads to improved accuracy in identifying both variable-scale lamina combinations and high-frequency interbedded sequences. Additionally, the study area is characterized by a large number of fault zones formed by intense tectonic activity. Strong geostress has led to the development of structural and low-angle fractures in the reservoir. These interconnected fractures often act as conduits for surface acidic fluids, which infiltrate underground during diagenesis and dissolve carbonate minerals to form caverns (Bastesen et al., 2009; Lyu et al., 2024; Wu et al., 2017). In late-stage diagenesis, these fractures and cavities are often filled with gypsum or calcite (Cerva-Alves et al., 2017; Zhao et al., 2025), resulting in abnormally high log responses for gypsum or calcite. However, these minerals likely postdate deposition, weakening the link between logging response and the original sedimentary environment.

For instance, the LCGC often shows similar logging features to those segments filled with gypsum. In non-cored wells, this can easily lead to misinterpretation. Here, the Transformer's advantage in context-enhanced modeling becomes apparent. Its ability to model long-sequence dependencies enables it to learn the vertical developmental patterns of lamina combinations associated with sedimentary cycles in the study area. The decoder component of the Transformer can uncover and apply such patterns to correct prediction results, effectively imparting a form of "geological reasoning" to the model. In blind well validation, the Transformer showed higher identification accuracy for these core segments than the Random Forest, minimizing misclassification caused by similar logging responses.

These advantages are further demonstrated in the recognition comparison shown in Fig. 13. In well B41-2, from 4191.22 m to 4192.30 m, several lamina combination segments approximately 0.5 m thick were identified, and a small-scale, high-frequency interbedding sequence was observed from 4192.41 m to 4192.89 m. The Random Forest algorithm produced multiple misclassifications, particularly in scale prediction, whereas the Transformer accurately recognized both intervals in scale and type. In the fracture-filled interval at 4192.99–4193.10 m, which should correspond to BCF, the logging responses were distorted by calcite-filled fractures, causing Random Forest to misclassify the interval as BCC. However, the Transformer still correctly identified the primary lamina combination as BCF with only minor misclassification. Similarly, the interval at 4202.59–4205.51 m featured gypsum-filled fractures that led to anomalies in the logging data, confusing the Random Forest but not affecting the Transformer's prediction. At 4211.97–4212.06 m, where LCGC was developed, the Transformer once again accurately identified the segment and avoided the scale-related mispredictions made by Random Forest.

In conclusion, the Transformer algorithm demonstrates superior performance in scale prediction, anomaly handling caused by gypsum or calcite filling, and sequence context modeling. Its accuracy and robustness clearly surpass those of traditional methods, further validating its advantages in complex lithological identification tasks.

6. Conclusions

(1) Based on the petrological characteristics of the upper member of the Xiaganchaigou Formation in the Yingxi area, five single lamina types were identified: carbonate laminae, felsic laminae, clay laminae, gypsiferous dolomite laminae, and sandy dolomite laminae. In addition, six lamina combination types were classified, including laminated clay–carbonate rock combination, bedded clay–carbonate rock combination, laminated clay–felsic rock combination, bedded clay–felsic rock combination, laminated clay–gypsiferous dolomitic rock combination, and bedded clay–sandy dolomitic rock combination. These lamina combination types show significant differences in reservoir development parameters and well-logging responses, indicating the practical relevance and operational feasibility of the classification scheme.

(2) The genesis of the six lamina combination types in the Yingxi area is closely linked to paleoenvironmental factors such as paleosalinity and paleowater depth. Under ideal conditions, a short-term sedimentary cycle can be summarized, from top to bottom, as: LCF or BCF → BCSC → BCC → LCC → LCGC, providing a reference model for the lamina development patterns of saline lacustrine mixed siliciclastic–carbonate reservoirs.

(3) The Transformer-based intelligent recognition model demonstrates significantly higher accuracy than the traditional Random Forest algorithm in identifying various lamina combinations, with an overall accuracy reaching 84%. This model effectively handles the nonlinear relationships among multiple parameters, making it particularly suitable for mixed lithological reservoirs with high interlayer heterogeneity. Its self-attention mechanism enhances the recognition accuracy of gypsum- and calcite-filled fracture and dissolution zones by learning vertical developmental patterns of lamina combinations. This method offers a novel approach for conventional well-log-based lamina characterization and provides valuable insights for the identification and interpretation of complex laminae in similar sedimentary environments.

CRediT authorship contribution statement

Jia-Lin Fu: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Da-Li Yue:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Wu-Rong Wang:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Kun-Yu Wu:** Validation, Supervision, Resources, Project administration, Investigation, Data curation. **Han Wang:** Validation, Software, Resources, Methodology. **Ying-Hai Jiang:** Supervision, Resources, Project administration, Methodology, Conceptualization. **Shu-Qi Zhang:** Validation, Methodology, Investigation. **Zi-Mo Xu:** Investigation, Formal analysis, Data curation. **Wei Li:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was financially supported by National Major Science and Technology Projects of China (No. 2024ZD1406601), National Natural Science Foundation of China (Nos. 42272186, 42302128, 42472179, and 42202109). We would also like to thank the editors and anonymous reviewers, as well as the Research Institute of Exploration and Development, PetroChina Qinghai Oilfield Company for providing samples, technical support, and assistance with this study, and for their agreement to publish the results.

References

Ahmed, S., Nielsen, I.E., Tripathi, A., et al., 2023. Transformers in time-series analysis: A tutorial. Circ. Syst. Signal Process. 42 (12), 7433–7466. https://doi.org/10.1007/s00034-023-02454-8.

Anderson, R.Y., Dean, W.E., 1988. Lacustrine varve formation through time. Paleogeogr. Paleoclimatol. Paleoecol. 62 (1), 215–235. https://doi.org/10.1016/0031-0182(88)90055-7.

Bai, C.Y., Yu, B.S., Liu, H.M., et al., 2018. The genesis and evolution of carbonate minerals in shale oil formations from Dongying depression, Bohai Bay Basin, China. Int. J. Coal Geol. 189, 8–26. https://doi.org/10.1016/j.coal.2018.02.008.

Bastesen, E., Braathen, A., Nottveit, H., et al., 2009. Extensional fault cores in micritic carbonate—case studies from the Gulf of Corinth, Greece. J. Struct. Geol. 31 (4), 403–420. https://doi.org/10.1016/j.jsg.2009.01.005.

Cerva-Alves, T., Dorneles-Remus, M.V., Dani, N., et al., 2017. Integrated field, mineralogical and geochemical characteristics of Caçapava do Sul alvikite and beforsite

intrusions: A new ediacaran carbonatite complex in Southernmost Brazil. Ore Geol. Rev. 88, 352–369. https://doi.org/10.1016/j.oregeorev.2017.05.017.
Chen, D.Q., Shen, X.S., Cui, J., et al., 2015. Reservoir characteristics and controlling factors of deep diamictite in Yingxi area, Qaidam Basin. Lithol. Reserv. 27 (5), 211–217. https://doi.org/10.3969/j.issn.1673-8926.2015.05.034 (in Chinese).
Chen, S.Y., Zhang, S., Liu, H.M., et al., 2017. Discussion on mixing of fine-grained sediments in lacustrine deep water. J. Palaegeogr. 19 (2), 271–284. https://doi.org/10.7605/gdlxb.2017.02.021 (in Chinese).
Fan, J.L., Chen, W.H., Yue, A.Z., et al., 2024. An enhanced accuracy method for monitoring formation water salinity utilizing elemental spectroscopy logging. Geoenergy Sci. Eng. 233, 212521. https://doi.org/10.1016/j.geoen.2023.212521.
Fan, X.Y., Meng, F., Deng, J., et al., 2025. Transformative reconstruction of missing acoustic well logs using multi-head self-attention BiRNNs. Geoenergy Sci. Eng. 245, 213513. https://doi.org/10.1016/j.geoen.2024.213513.
Feng, J.L., Hu, K., Cao, J., et al., 2011. A review on mixed rocks of terrigenous clastics and carbonates and their petroleum-gas geological significance. Geol. J. China Univ. 17 (2), 297–307. https://doi.org/10.16108/j.issn1006-7493.2011.02.016 (in Chinese).
Feng, R.H., 2021. Improving uncertainty analysis in well log classification by machine learning with a scaling algorithm. J. Pet. Sci. Eng. 196, 107995. https://doi.org/10.1016/j.petrol.2020.107995.
Feng, R.H., Grana, D., Balling, N., 2021. Imputation of missing well log data by Random Forest and its uncertainty analysis. Comput. Geosci. 152, 104763. https://doi.org/10.1016/j.cageo.2021.104763.
Fu, S.T., Ma, D.D., Guo, Z.J., et al., 2015. Strike-slip superimposed Qaidam Basin and its control on oil and gas accumulation, NW China. Petrol. Explor. Dev. 42 (6), 778–789. https://doi.org/10.1016/S1876-3804(15)30074-4.
Gama, P.H.T., Faria, J., Sena, J., et al., 2025. Imputation in well log data: A benchmark for machine learning methods. Comput. Geosci. 196, 105789. https://doi.org/10.1016/j.cageo.2024.105789.
Gao, X.Z., Wang, Q., Zhang, R., et al., 2013. The characteristics of lacustrine mixed deposit in lower Xiaganchaigou formation and its effect on reservoir property: a case in the central areas of northern Qaidam Basin. Acta Sedimentol. Sin. 31 (4), 724–729. https://doi.org/10.14027/j.cnki.cjxb.2013.04.014 (in Chinese).
Guo, P., Liu, C.Y., Yu, M.L., et al., 2018. Paleosalinity evolution of the paleogene perennial Qaidam Lake on the Tibetan plateau: Climatic vs. tectonic control. Int. J. Earth Sci. 107 (5), 1641–1656. https://doi.org/10.1007/s00531-017-1564-8.
Han, X.H., He, Y.S., Chen, J., et al., 2024. Research on rock lithology intelligent identification based on swin transformer. Mod. Electron. Tech. 47 (7), 37–44. https://doi.org/10.16652/j.issn.1004-373x.2024.07.006 (in Chinese).
Haritha, D., Satyavani, N., Ramesh, A., 2025. Generation of missing well log data with deep learning: CNN-BI-LSTM approach. J. Appl. Geophys. 233, 105628. https://doi.org/10.1016/j.jappgeo.2025.105628.
He, H.L., Guo, P., Li, C.Z., et al., 2025a. Formation of carbonate laminae in the Oligocene Shangganchaigou formation of the Qaidam Basin and its paleoclimatic significance. Acta Sedimentol. Sin. 43 (3), 1019–1036. https://doi.org/10.14027/j.issn.1000-0550.2023.065 (in Chinese).
He, J.X., Chen, X., Li, Y.Y., et al., 2025b. Bedding structure types, combinations and reservoir significances of fine-grained rocks of the Permian Lucaogou Formation in Jimsar Sag, Junggar Basin. Mar. Geol. Quat. Geol. 45 (1), 178–187. https://doi.org/10.16562/j.cnki.0256-1492.2023121301 (in Chinese).
Huo, Z.P., Zhang, J.C., Li, P., et al., 2018. An improved evaluation method for the brittleness index of shale and its application: A case study from the southern north China Basin. J. Nat. Gas Sci. Eng. 59, 47–55. https://doi.org/10.1016/j.jngse.2018.08.014.
Jackson, M.J., 1985. Mid-proterozoic dolomitic varves and microcycles from the McArthur Basin, northern Australia. Sediment. Geol. 44 (3), 301–326. https://doi.org/10.1016/0037-0738(85)90017-X.
Jarvie, D.M., Hill, R.J., Ruble, T.E., et al., 2007. Unconventional shale-gas systems: The Mississippian Barnett shale of north-central texas as one model forthermogenic shale-gas assessment. AAPG (Am. Assoc. Pet. Geol.) Bull. 91 (4), 475–499. https://doi.org/10.1306/12190606068.
Jiang, Q.C., Liu, B., Guo, R.T., et al., 2017. Microbial mechanism of lacustrine primary dolomite. J. Palaegeogr. 19 (2), 257–269. https://doi.org/10.7605/gdlxb.2017.02.020 (in Chinese).
Jiang, Q.C., Ma, Y.S., Shen, Y.C., et al., 2019. High-frequency redox variations of the eocene cyclic lacustrine sediments in the Yingxi area, western Qaidam Basin, China. J. Asian Earth Sci. 174, 135–151. https://doi.org/10.1016/j.jseaes.2018.11.025.
Jiang, Z.X., Liang, C., Wu, J., et al., 2013. Several issues in sedimentological studies on hydrocarbon-bearing fine-grained sedimentary rocks. Acta Petrol. Sin. 34 (6), 1031–1039. https://doi.org/10.7623/syxb201306001 (in Chinese).
Lai, J., Pang, X.J., Zhao, X., et al., 2022. Typical misunderstandings and scientific ideas in well logging geology research. Nat. Gas. Ind. 42 (7), 31–44. https://doi.org/10.3787/j.issn.1000-0976.2022.07.004 (in Chinese).
Lai, J., Wang, G.W., Wang, S., et al., 2018. A review on the applications of image logs in structural analysis and sedimentary characterization. Mar. Petrol. Geol. 95, 139–166. https://doi.org/10.1016/j.marpetgeo.2018.04.020.
Li, G.X., Zhu, R.K., Zhang, Y.S., et al., 2022. Geological characteristics, evaluation criteria and discovery significance of Paleogene Yingxiongling shale oil in Qaidam Basin, NW China. Petrol. Explor. Dev. 49 (1), 18–31. https://doi.org/10.11698/PED.2022.01.02 (in Chinese).
Li, L.H., Huang, B.X., Li, Y.Y., et al., 2018. Multi-scale modeling of shale laminas and fracture networks in the Yanchang formation, southern Ordos Basin, China. Eng. Geol. 243, 231–240. https://doi.org/10.1016/j.enggeo.2018.07.010.
Li, L.H., Huang, B.X., Li, Y.Y., et al., 2019. Multi-scale 3-D modeling of Yanchang shale geological structure considering laminas and fracture networks. J. Eng. Geol. 27 (1), 69–79. https://doi.org/10.13544/j.cnki.jeg.2019-061 (in Chinese).
Li, Q.L., Wang, Y., Shao, Y.D., et al., 2023. A comparative study on the most effective machine learning model for blast loading prediction: From GBDT to transformer. Eng. Struct. 276, 115310. https://doi.org/10.1016/j.engstruct.2022.115310.
Li, W., Zhu, X.M., Duan, H.L., et al., 2020. Characteristics and forming mechanism of laminae fine-grained sedimentary rock of the Paleogene Funing Formation in Gaoyou and Jinhu sags, Subei Basin. J. Palaegeogr. 22 (3), 469–482 (in Chinese).
Li, W.Z., Liu, C.S., Xu, Y.Y., et al., 2024a. An interpretable hybrid deep learning model for flood forecasting based on transformer and LSTM. J. Hydrol. Reg. Stud. 54, 101873. https://doi.org/10.1016/j.ejrh.2024.101873.
Li, X., Wang, J.G., Li, F., et al., 2021. Sedimentary characteristics of Eocene lacustrine microbialites in western Qaidam Basin: A case study from Xiaganchaigou Formation in Xichagou and Liangdong areas. Lithol. Reserv. 33 (3), 63–73. https://doi.org/10.12108/yxyqc.20210306 (in Chinese).
Li, X.B., Xiao, K., Sun, L., et al., 2024b. Paleosalinity characteristics of the 1st member in the lower cretaceous prosopis Formation, Baobab North Sag, Bongor Basin. Front. Earth Sci. 12, 1450001. https://doi.org/10.3389/feart.2024.1450001.
Li, Z.J., Deng, S.G., Hong, Y.Z., et al., 2024c. A novel hybrid CNN–SVM method for lithology identification in shale reservoirs based on logging measurements. J. Appl. Geophys. 223, 105346. https://doi.org/10.1016/j.jappgeo.2024.105346.
Liang, C., Cao, Y.C., Liu, K.Y., et al., 2018. Diagenetic variation at the lamina scale in lacustrine organic-rich shales: Implications for hydrocarbon migration and accumulation. Geochem. Cosmochim. Acta 229, 112–128. https://doi.org/10.1016/j.gca.2018.03.017.
Liu, L., Li, W., Gao, J., et al., 2026. An interpretable attention-guided generative adversarial network framework with dual-domain learning for multi-condition constrained sedimentary facies modeling. Pet. Sci. 23 (4). https://doi.org/10.1016/j.petsci.2026.02.018 (in press).
Liu, Y.F., Li, H., Peng, P., et al., 2024. High-resolution sequence stratigraphy research based on continuous wavelet transform and INPEFA: A case study in the Kalashayi Formation, Tarim Basin, China. Unconv. Resour. 4, 100099. https://doi.org/10.1016/j.uncres.2024.100099.
Liu, Y.J., Lai, F.Q., Xu, H., et al., 2025. Intelligent identification methods for shale lithology based on the coupling deeply of logging curves. Bull. Geol. Sci. Technol. 44 (1), 308–320. https://doi.org/10.19509/j.cnki.dzkq.tb20230361 (in Chinese).
Liu, Z.G., Zhang, Y.S., Song, G.Y., et al., 2021. Mixed carbonate rocks lithofacies features and reservoirs controlling mechanisms in the saline lacustrine basin in Yingxi area, Qaidam Basin, NW China. Petrol. Explor. Dev. 48 (1), 68–80. https://doi.org/10.11698/PED.2021.01.06 (in Chinese).
Lu, S.X., Zhang, H.Z., Meng, N., et al., 2007. Application of INPEFA technique to carry out sequence-stratigraphic study. Oil Geophys. Prospect. 42 (6), 703–708. https://doi.org/10.16562/j.cnki.issn.1000-7210.2007.06.001, 733, 609 (in Chinese).
Lyu, X.R., Ju, B.S., Wang, B., et al., 2024. Mechanism, mode, and prediction of karst caves collapse in the deep marine carbonate fracture-cavity reservoir. Mar. Petrol. Geol. 167, 106978. https://doi.org/10.1016/j.marpetgeo.2024.106978.
Mingram, J., 1998. Laminated Eocene maar-lake sediments from Eckfeld (Eifel region, Germany) and their short-term periodicities. Paleogeogr. Paleoclimatol. Paleoecol. 140 (1), 289–305. https://doi.org/10.1016/S0031-0182(98)00021-2.
Moreira-Lima, B.E., De-Ros, L.F., 2019. Deposition, diagenetic and hydrothermal processes in the aptian pre-salt lacustrine carbonate reservoirs of the northern Campos Basin, Offshore Brazil. Sediment. Geol. 383, 55–81. https://doi.org/10.1016/j.sedgeo.2019.01.006.
Olivito, J.P.R., Souza, F.J., 2020. Depositional model of early cretaceous lacustrine carbonate reservoirs of the Coqueiros Formation — northern Campos Basin, southeastern Brazil. Mar. Petrol. Geol. 111, 414–439. https://doi.org/10.1016/j.marpetgeo.2019.07.013.
Pan, Y.S., Huang, Z.L., Li, T.J., et al., 2021. Pore structure characteristics and evaluation of lacustrine mixed fine-grained sedimentary rocks: A case study of the Lucaogou Formation in the Malang Sag, Santanghu Basin, Western China. J. Pet. Sci. Eng. 201, 108545. https://doi.org/10.1016/j.petrol.2021.108545.
Pang, X.J., Wang, G.W., Mountney, N.P., et al., 2023. Prediction of lamina structure and reservoir quality in shale using well logs: The Cretaceous Qingshankou Formation, Gulong Sag, Songliao Basin, China. Geoenergy Sci. Eng. 227, 211827. https://doi.org/10.1016/j.geoen.2023.211827.
Park, M.H., Fürsich, F.T., 2001. Cyclic nature of lamination in the Tithonian Solnhofen Plattenkalk of southern Germany and its palaeoclimatic implications. Int. J. Earth Sci. 90 (4), 847–854. https://doi.org/10.1007/s005310000191.
Peng, J., Yu, L.D., Xu, T.Y., et al., 2022. Logging identification of Milankovitch cycle and environmental response characteristics of lacustrine shale—A case study on Es^{4scs} in well Fanye 1, Dongying Sag, Jiyang Depression, Bohai Bay Basin. Oil Gas Geol 43 (4), 957–969. https://doi.org/10.11743/ogg20220417 (in Chinese).
Pollastro, R.M., 2007. Total petroleum system assessment of undiscovered resources in the giant Barnett shale continuous (unconventional) gas accumulation, Fort Worth Basin Texas. AAPG (Am. Assoc. Pet. Geol.) Bull. 91 (4), 551–578. https://doi.org/10.1306/06200606007.
Qiao, Y.P., Tan, X.C., Liu, Y., et al., 2020. Characteristics of high-frequency lake-level fluctuations in the saline lacustrine basin and its geological significance: A case

study from the upper member of the Paleogene Lower Ganchaigou Formation in the Yingxi area, Qaidam Basin. Acta Sedimentol. Sin. 38 (6), 1152–1165. https://doi.org/10.14027/j.issn.1000-0550.2019.113 (in Chinese).
Ran, Y., Wang, G.W., Lai, J., et al., 2016. Quantitative characterization of diagenetic. Acta Sedimentol. Sin. 34 (4), 694–706. https://doi.org/10.14027/j.cnki.cjxb.2016.04.010 (in Chinese).
Ren, H., Wang, X.G., 2021. Review of attention mechanism. J. Comput. Appl. 41 (s1), 1–6. https://doi.org/10.11772/j.issn.1001-9081.2020101634 (in Chinese).
Shi, J.Y., Jin, Z.J., Liu, Q.Y., et al., 2019. Cyclostratigraphy and astronomical tuning of the middle Eocene terrestrial successions in the Bohai Bay Basin, Eastern China. Global Planet. Change 174, 115–126. https://doi.org/10.1016/j.gloplacha.2019.01.001.
Shi, Z.S., Qiu, Z., Dong, D.Z., et al., 2018. Laminae characteristics of gas-bearing shale fine-grained sediment of the Silurian Longmaxi Formation of Well Wuxi 2 in Sichuan Basin, SW China. Petrol. Explor. Dev. 45 (2), 339–348. https://doi.org/10.11698/PED.2018.02.18 (in Chinese).
Song, G.Y., Liu, Z.G., Wang, Y.Q., et al., 2024. Lithofacies types, sedimentary cycles, and facies models of saline lacustrine hybrid sedimentary rocks: A case study of Neogene in Fengxi area, Qaidam Basin, NW China. Petrol. Explor. Dev. 51 (6), 1304–1316. https://doi.org/10.11698/PED.20230702 (in Chinese).
Song, H.Y., Yi, H.S., Fan, A.C., et al., 2010. Petrology and sedimentary environments of lacustrine carbonate rocks in the Xichagou Section, western Qaidam Basin. Geol. China 37 (1), 117–126 (in Chinese).
Song, M.S., Liu, H.M., Wang, Y., et al., 2020. Enrichment rules and exploration practices of Paleogene shale oil in Jiyang Depression, Bohai Bay Basin China. Petrol. Explor. Dev. 47 (2), 225–235. https://doi.org/10.11698/PED.2020.02.02 (in Chinese).
Song, T.R., Zhu, W.Y., Chen, Z.X., et al., 2023. A novel well-logging data generation model integrated with random forests and adaptive domain clustering algorithms. Geoenergy Sci. Eng. 231, 212381. https://doi.org/10.1016/j.geoen.2023.212381.
Song, Y., Zhou, L., Guo, X.G., et al., 2017. Characteristics and occurrence of lacustrine dolomitic tight-oil reservoir in the Middle Permian Lucaogou Formation, Jimusaer sag, southeastern Junggar Basin. Acta Petrol. Sin. 33 (4), 1159–1170 (in Chinese).
Sun, N.L., Chen, T.Y., Gao, J.B., et al., 2023. Lithofacies and reservoir characteristics of saline lacustrine fine-grained sedimentary rocks in the northern Dongpu Sag, Bohai Bay Basin: Implications for shale oil exploration. J. Asian Earth Sci. 252, 105686. https://doi.org/10.1016/j.jseaes.2023.105686.
Sun, X.J., Wang, P.X., 2005. How old is the Asian monsoon system? — palaeobotanical records from China. Paleogeogr. Paleoclimatol. Paleoecol. 222 (3), 181–222. https://doi.org/10.1016/j.palaeo.2005.03.005.
Sun, Y.Z., Pang, S.C., Zhang, Y.A., 2024. Innovative lithology identification enhancement via the recurrent transformer model with well logging data. Geoenergy Sci. Eng. 240, 213015. https://doi.org/10.1016/j.geoen.2024.213015.
Tian, M.Z., Zhu, C., Li, S.M., et al., 2023. Application of logging lithofacies identification technology of lacustrine carbonate rocks: A case study of Yingxi area, Qaidam Basin. China Pet. Explor. 28 (1), 135–143. https://doi.org/10.3969/j.issn.1672-7703.2023.01.012 (in Chinese).
Tribovillard, N., Algeo, T.J., Lyons, T., et al., 2006. Trace metals as paleoredox and paleoproductivity proxies: An update. Chem. Geol. 232 (1), 12–32. https://doi.org/10.1016/j.chemgeo.2006.02.012.
Vaswani, A., 2017. Attention is all you need. Adv. Neural Inf. Process. Syst. 30, 1–11.
Verbiest, N., Ramentol, E., Cornelis, C., et al., 2014. Preprocessing noisy imbalanced datasets using SMOTE enhanced with fuzzy rough prototype selection. Appl. Soft Comput. 22, 511–517. https://doi.org/10.1016/j.asoc.2014.05.023.
Vijayakumar, A., Alshardan, A., Alghamdi, A.G., et al., 2025. High resolution spatial modelling of soil parent material: An initial investigation study using Random Forest machine learning techniques. J. South Am. Earth Sci. 162, 105598. https://doi.org/10.1016/j.jsames.2025.105598.
Wang, B.J., Cai, M.J., Lin, C.M., et al., 2014. Characteristics and origin of lacustrine dolostone of the Paleogene Shahejie Formation in Tanggu area, Bohai Bay Basin. J. Palaegeogr. 16 (1), 65–76. https://doi.org/10.7605/gdlxb.2014.01.007 (in Chinese).
Wang, G.M., Zhong, J.H., 2004. A review and the prospects of the researches on sedimentary mechanism of lacustrine laminae. Acta Petrol. Mineral. 23 (1), 43–48. https://doi.org/10.3969/j.issn.1000-6524.2004.01.006 (in Chinese).
Wang, J.G., Zhang, D.W., Yang, S.Y., et al., 2020. Sedimentary characteristics and genesis of the salt lake with the upper member of the lower Ganchaigou Formation from Yingxi sag, Qaidam Basin. Mar. Petrol. Geol. 111, 135–155. https://doi.org/10.1016/j.marpetgeo.2019.08.006.
Wang, M., Guo, Z.Q., Jiao, C.X., et al., 2019. Exploration progress and geochemical features of lacustrine shale oils in China. J. Pet. Sci. Eng. 178, 975–986. https://doi.org/10.1016/j.petrol.2019.04.029.
Wang, M., Yang, J.L., Wang, X., et al., 2023. Identification of shale lithofacies by well logs based on Random Forest algorithm. Earth Sci 48 (1), 130–142. https://doi.org/10.3799/dqkx.2022.181 (in Chinese).
Wang, M., Zhu, J.J., Yu, G.H., et al., 2013. The shale lithofacies characteristics and logging analysis techniques in Luojia area. Well Logging Technol 37 (4), 426–431. https://doi.org/10.16489/j.issn.1004-1338.2013.04.020 (in Chinese).
Wang, W.R., Qu, L.B., Yue, D.L., et al., 2025. Integrated artificial intelligence approach for well-log fluid identification in dual-medium tight sandstone gas reservoirs. Front. Earth Sci. 13, 1591110. https://doi.org/10.3389/feart.2025.1591110.
Wang, Z.H., Zhu, X.M., Sun, Z.C., et al., 2015. Igneous lithology identification and lithofacies classification in the basin using logging data: Taking Junggar Basin as an example. Earth Sci. Front. 22 (3), 254–268. https://doi.org/10.13745/j.esf.2015.03.022 (in Chinese).
Wei, W., Algeo, T.J., Meyer, D., et al., 2025. Utility of the B/Ga salinity proxy in carbonate and marly sediments. Chem. Geol. 682, 122751. https://doi.org/10.1016/j.chemgeo.2025.122751.
Wen, Q.S., Zhou, T., Zhang, C.L., et al., 2023. Transformers in time series: A survey. Proc. Thirty-Second Inter. Joint Conf. Artif. Intell. 6778–6786.
Wilhelms, A., Larter, S.R., Leythaeuser, D., et al., 1989. Recognition and quantification of the effects of primary migration in a Jurassic clastic source-rock from the Norwegian continental shelf. Org. Geochem. 16 (1), 103–113. https://doi.org/10.1016/0146-6380(90)90030-4.
Wu, T., Dai, S.K., Cao, F., et al., 2017. Strike-slip fault system and it's controlling on the formation of "reservoir of fault controlling dissolution"—A case study of Toufutai area, Tahe Oil Fields. Adv. Geosci. 7 (5), 681–694. https://doi.org/10.12677/AG.2017.75069 (in Chinese).
Xin, B.X., Zhao, X.Z., Hao, F., et al., 2022. Laminae characteristics of lacustrine shales from the Paleogene Kongdian Formation in the Cangdong Sag, Bohai Bay Basin, China: Why do laminated shales have better reservoir physical properties? Int. J. Coal Geol. 260, 104056. https://doi.org/10.1016/j.coal.2022.104056.
Xiong, Y., Wu, K.Y., Tan, X.C., et al., 2018. Influence of lake-level fluctuation on the mixed saline lacustrine carbonate reservoir: A case study from the Upper Member of Paleogene Lower Ganchaigou Formation in the Yingxi area of Qaidam Basin. J. Palaegeogr. 20 (5), 855–868. https://doi.org/10.7605/gdlxb.2018.05.060 (in Chinese).
Yang, T., Cao, Y.C., Liu, K.Y., et al., 2020. Gravity-flow deposits caused by different initiation processes in a deep-lake system. AAPG Bull. 104 (7), 1463–1499. https://doi.org/10.1306/03172017081.
Yue, D.L., Li, W., Wang, W.R., et al., 2025. Advances and perspectives in intelligent characterization and modeling of clastic reservoir. J. Palaeogeogr. 27 (4), 903–923. https://doi.org/10.7605/gdlxb.2025.090 (in Chinese).
Zeng, W.R., Zhang, Z.H., Wang, B.R., et al., 2023. Formation mechanism of organic-rich mixed sedimentary rocks in saline lacustrine basin, Permian Lucaogou Formation, Jimsar Sag, Junggar Basin, Northwest China. Mar. Petrol. Geol. 156, 106452. https://doi.org/10.1016/j.marpetgeo.2023.106452.
Zhang, C., Guo, Y., Li, M., 2021. Review of development and application of artificial neural network models. Comput. Eng. Appl. 57 (11), 57–69. https://doi.org/10.3778/j.issn.1002-8331.2102-0256 (in Chinese).
Zhang, S.Q., Ren, S.X., Yue, D.L., et al., 2026. A shale oil 'sweet-spot' evaluation factor construction method based on quantitative analysis of the influence degree of geological controlling factors. Unconv. Resour. 13, 100404. https://doi.org/10.1016/j.uncres.2026.100404.
Zhang, X.H., 2000. Classification and origin of mixosedimentite. Geol. Sci. Technol. Inf. 19 (4), 31–34. https://doi.org/10.3969/j.issn.1000-7849.2000.04.006 (in Chinese).
Zhang, Y., Ye, Y.J., Qiu, J., et al., 2024. Study on quantitative interpretation of uranium spectral gamma-ray logging based on machine learning algorithm. Nucl. Eng. Technol. 56 (12), 4959–4965. https://doi.org/10.1016/j.net.2024.07.004.
Zhang, Y.S., Wu, K.Y., Jiang, Y.H., et al., 2018. Geological characteristics of deep carbonate hydrocarbon-bearing pool in the western Yingxiongling area in Qaidam Basin. Nat. Gas Geosci. 29 (3), 358–369. https://doi.org/10.11764/j.issn.1672-1926.2018.02.005 (in Chinese).
Zhang, Y.S., Fu, J.L., Wu, K.Y., et al., 2025. Lithofacies division and intelligent identification of the lacustrine mixed rocks in the Upper Xiaganchaigou Formation in Yingxi area of the Qaidam Basin, northwestern China. J. Palaeogeogr-English 14 (4), 100270. https://doi.org/10.1016/j.jop.2025.100270.
Zhao, W.Z., Zhu, R.K., Hu, S.Y., et al., 2020. Accumulation contribution differences between lacustrine organic-rich shales and mudstones and their significance in shale oil evaluation. Petrol. Explor. Dev. 47 (6), 1079–1089. https://doi.org/10.11698/PED.2020.06.02 (in Chinese).
Zhao, X., Wang, G.W., Li, Y.F., et al., 2025. Development characteristics and logging identification of natural fractures in shale oil reservoirs of the Yingxiongling area, Qaidam Basin. Nat. Gas Geosci. 36 (4), 713–733. https://doi.org/10.11764/j.issn.1672-1926.2024.09.009 (in Chinese).
Zhu, C., Liu, Z.G., Song, G.Y., et al., 2022. Sedimentary model, evolution and distribution of Paleogene lacustrine carbonate rocks in Yingxiongling structural belt, Qaidam Basin. Acta Petrol. Sin. 43 (11), 1558–1567. https://doi.org/10.7623/syxb202211004 (in Chinese).
Zhu, F., Li, C.X., Leng, J.Y., et al., 2023. Paleoenvironmental characteristics of lacustrine shale and its impact on organic matter enrichment in funing Formation of Subei Basin. Minerals 13 (11), 1439. https://doi.org/10.3390/min13111439.
Zhu, X., 2008. Sedimentary Petrology. Petroleum Industry Press, Beijing, pp. 81–88 (in Chinese).