Petroleum Science 23 (2026) 3037–3058

Contents lists available at ScienceDirect

Petroleum Science

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

Original Paper

Reinterpretation of volcanic and hydrothermal influence on organic matter accumulation and paleo-depositional environment in a lacustrine shale system: Insight from interpretable machine learning model

Enze Wang [a,b,c], Xiaoxiao Ma [a,b,c], Maowen Li [a,b,c,*], Yingxiao Fu [d], Menhui Qian [a,b,c], Tingting Cao [a,b,c], Zhiming Li [a,b,c], Yue Feng [e], Zhijun Jin [f], Tong Li [c]

[a] *State Key Laboratory of Shale Oil and Gas Enrichment Mechanisms and Efficient Development, SINOPEC, Beijing, 102206, China*
[b] *Key Laboratory of Shale Oil/Gas Exploration and Production Technology, SINOPEC, Beijing, 102206, China*
[c] *Petroleum Exploration and Production Research Institute, SINOPEC, Beijing, 102206, China*
[d] *School of Earth and Space Sciences, Peking University, Beijing, 100871, China*
[e] *Research Institute of Exploration and Development, PetroChina Changqing Oilfield Company, Xi'an, 710018, Shaanxi, China*
[f] *Institute of Energy, Peking University, Beijing, 100871, China*

A R T I C L E I N F O

Article history:
Received 12 August 2025
Received in revised form
13 October 2025
Accepted 6 February 2026
Available online 16 February 2026

Edited by Xi Zhang and Jie Hao

Keywords:
Lacustrine shale
Paleoenvironment reconstruction
Interpretable machine leaning
Volcanic and hydrothermal activity
Organic matter enrichment mechanisms
Ordos Basin

A B S T R A C T

The significant resource potential of lacustrine shale oil has spurred interest in understanding organic matter enrichment mechanisms. However, conventional analytical methods are often constrained by multicollinearity and statistical limitations, hindering accurate identification of enrichment mechanisms under complex depositional conditions. Interpretable machine learning models present a viable alternative to address this issue. The 7th member of the Yanchang Formation (T_3y_7) is one of the key targets for lacustrine shale oil exploration and development in China. Nevertheless, the role of volcanic activity in organic matter accumulation within this unit remains controversial, hindering a comprehensive understanding of organic enrichment mechanisms in lacustrine systems. To address this challenge, this study integrates geochemical data from YYa borehole in the southern Ordos Basin and previously published datasets to reconstruct the paleoenvironmental settings of various lithofacies assemblages in the T_3y_7 shale. By leveraging published literature and new experimental data, this study develops an interpretable machine learning model based on random forest (RF), extreme gradient boosting (XGBoost), and support vector machine (SVM) to quantitatively determine the dominant controls on organic matter accumulation. Using the T_3y_7 shale as a case study, the influence mechanisms of both volcanic and hydrothermal activities on organic enrichment in lacustrine shale systems are further explored. Two distinct lithofacies assemblages were identified in the T_3y_7 shale: laminated black shale and massive dark-gray mudstone. These assemblages are characterized by significant differences in paleoclimate, salinity, redox conditions, primary productivity, terrigenous influx, and the intensity of volcanic and hydrothermal activities. Results from the interpretable machine learning model indicate that organic matter enrichment in the T_3y_7 shale is primarily governed by the coupling of high primary productivity and favorable preservation conditions, and decoupling of volcanic and hydrothermal activities was a significant factor in the development of different lithofacies in the T_3y_7. Importantly, the influence of volcanic activity on organic matter enrichment in lacustrine systems is nonlinear: while moderate volcanic activity promotes organic accumulation, intense volcanic activity tends to exert a detrimental effect. In contrast, hydrothermal activity appears to play a more consistently positive role in the formation of organic-rich shale in the T_3y_7. These findings highlight the need to distinguish between volcanic and hydrothermal processes in future studies in order to more objectively evaluate their respective impacts on organic-rich shale development in lacustrine settings. This study provides a robust framework for understanding the mechanisms of organic matter enrichment in

* Corresponding author.
E-mail address: limw.syky@sinopec.com (M. Li).
Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2026.02.010

shale deposited under complex geological conditions, and deepens our insight into the formation of organic-rich shale in lacustrine sedimentary systems influenced by volcanic and hydrothermal activity.

1. Introduction

Organic-rich shales have garnered increasing attention from geologists due to their immense hydrocarbon resource potential. For instance, in 2023, shale (tight) oil production in North America accounted for 64% of total oil output, with a production of approximately 3.04 billion barrels (EIA, 2024). In China, annual production of lacustrine shale oil exceeded 4 million tons in 2023 (National Energy Administration of China, 2024). The considerable economic value of shale oil has driven intensive academic interest in the mechanisms governing the formation of organic-rich shales. Extensive research has been conducted to investigate the enrichment of organic matter (OM) in shales, with high primary productivity and favorable preservation conditions widely recognized as the two dominant mechanisms (Pedersen, 1990; Algeo and Li, 2020). However, the formation of organic-rich shales is inherently the result of complex interactions and coupling among multiple geological factors. Unfortunately, most current studies rely on traditional univariate statistical methods to examine the relationship between paleoenvironmental proxies and total organic carbon (TOC) content (Wang and Shi, 2019; Guan et al., 2021; Wu et al., 2021; Burnaz et al., 2022; Feng et al., 2023). While these approaches may yield valid insights under relatively simple geological settings, they face significant limitations when applied to complex systems influenced by multiple coexisting geological controls or major geological events. In such scenarios, univariate analyses often fail to identify the key drivers of OM enrichment, potentially leading to misinterpretations of the underlying mechanisms. As a result, a more refined and quantitative approach is urgently needed to clarify the controlling factors behind OM accumulation in shales deposited under complex sedimentary conditions.

In China, shale oil exploration currently focuses primarily on lacustrine sedimentary systems. Compared to marine environments, lacustrine systems are more sensitive to paleoenvironmental fluctuations (Katz and Lin, 2014; Guo et al., 2025) and generally exhibit faster sedimentation rates. Recent studies have demonstrated that lacustrine systems possess high-resolution records of episodic geological events, such as large igneous province (LIP) activity and oceanic anoxic events (OAEs) (Liu et al., 2021, 2022, 2024; Wang et al., 2023a). These short-term geological perturbations can exert a significant influence on the formation of organic-rich shales. Throughout history of the Earth, volcanic and hydrothermal activities have had profound impacts on paleoenvironmental conditions. Volcanic eruptions can affect the paleoclimate through the emission of gases and volcanic ash, which in turn can alter primary productivity and the redox state of the water column (O'Dowd et al., 2004; Duggen et al., 2007; Jones and Gislason, 2008; Madonia et al., 2013). Likewise, hydrothermal fluids are enriched in nutrients such as Si, N, P, and Fe; once introduced into lacustrine water columns, they can significantly modify the conditions for OM accumulation (Korzhinsky et al., 1994). Despite these insights, the quantitative coupling between episodic geological events, particularly volcanic and hydrothermal activities, and the formation of organic-rich shales in lacustrine systems remains poorly constrained.

The Ordos Basin is one of the most important regions for lacustrine shale oil exploration and development in China. Numerous studies have been conducted on paleoenvironmental reconstructions and OM enrichment mechanisms within the Triassic Yanchang Formation. However, two key issues remain unresolved. First, most existing research relies on univariate analytical approaches to interpret OM enrichment mechanisms (Wang et al., 2017; Yang et al., 2022; Li et al., 2024). Given that the Yanchang Formation was significantly influenced by volcanic and hydrothermal activities during its deposition (Qiu et al., 2015; He et al., 2016; Zhang et al., 2017, 2020), such simplified methods may not accurately reveal the dominant factors controlling OM accumulation under complex geological conditions. Second, while many scholars have emphasized the positive role of volcanic activity in promoting OM enrichment in the Yanchang Formation (Wang et al., 2017; Liu et al., 2022; Li et al., 2024), Zhang et al. (2020) argued that volcanic events could induce climate changes that are detrimental to OM accumulation. This discrepancy raises a critical scientific question: what accounts for the divergence in these interpretations? Moreover, although volcanic and hydrothermal activities often occur concurrently, whether they exert similar influences on OM accumulation within the Yanchang Formation remains unclear. Specifically, what are the respective roles of volcanic and hydrothermal processes in shaping OM enrichment? Therefore, a more refined investigation is required to disentangle the controlling factors of OM accumulation in the Triassic Yanchang Formation of the Ordos Basin and to differentiate and clarify the distinct impacts of volcanic and hydrothermal activities on the formation of organic-rich shales.

The integration of machine learning (ML) into geoscientific and petroleum engineering research has revolutionized analytical methodologies due to its powerful pattern recognition capabilities. Unlike traditional linear regression models constrained by predefined mathematical relationships, ML algorithms can autonomously capture complex, nonlinear interactions between predictor variables and target responses. This adaptive modeling paradigm has demonstrated superior predictive performance in handling high-dimensional, multivariate datasets in geoscience contexts (Zhu et al., 2020; Lee et al., 2022). Recent research emphasizes the importance of interpretable ML models to bridge the gap between computational results and geological understanding (Wang et al., 2025a). Interpretable ML enables the explicit revelation of both linear and nonlinear relationships embedded in the model. This unique capability not only compensates for the limitations of traditional methods in handling complex multivariate correlations but also provides novel insights for elucidating the intrinsic associations among multiple complex geological factors (Liao et al., 2023). Among various interpretability techniques, post-hoc explanatory tools have gained traction across disciplines for their ability to extract human-comprehensible insights from opaque models (Zhang et al., 2021). Methods such as permutation feature importance (PFI), shapley additive explanations (SHAP), and partial dependence plots (PDP) have been widely applied across scientific fields (Wang et al., 2025a). These tools enable geoscientists to quantitatively evaluate the relative importance of paleoenvironmental proxies in OM accumulation, particularly within sedimentary systems governed by multiple synergistic controls.

This study integrates new geochemical data from the YYa borehole in the southern Ordos Basin with previously published datasets (Wang et al., 2017; Zhang et al., 2017; Li et al., 2020; Yuan et al., 2020; Shi et al., 2022; Zheng et al., 2022; Yang et al., 2022) to reconstruct the paleoenvironmental conditions associated with various lithofacies assemblages of the 7th member of the Yanchang Formation (T_3y_7). An interpretable machine learning model is employed to identify the key factors controlling OM accumulation in the T_3y_7 shale. Furthermore, the roles of volcanic and hydrothermal activities in influencing OM enrichment in lacustrine shale systems are systematically evaluated using T_3y_7 as a representative case. The findings provide a robust framework for deciphering OM enrichment mechanisms in shales deposited under complex geological conditions and offer new insights into the formation of organic-rich shales within lacustrine sedimentary systems influenced by volcanic and hydrothermal processes.

2. Geological setting

The Ordos Basin, located in northwestern China, is one of the most significant petroliferous basins in China, covering an area of approximately 320,000 km. The basin is subdivided into six first-order tectonic units: the Yimeng Uplift, the Western Thrust Belt, the Tianhuan Sag, the Jinxi Fold Belt, the Weibei Uplift, and the Yishan Slope (Fig. 1(a); Niu et al., 2025). The lacustrine sedimentary succession within the Ordos Basin comprises a series of formations ranging in age from the Middle–Upper Triassic to the Quaternary (Zheng et al., 2022). During the Late Triassic deposition of the Yanchang Formation, the basin evolved into a large intracontinental lake basin, characterized by the accumulation of a thick (1000–1500 m) clastic succession representing fluvial, deltaic, and lacustrine facies (Fu et al., 2018). Stratigraphically, the Yanchang Formation is divided into ten members from top to bottom (Deng et al., 2018). Among them, the T_3y_7 represents the peak phase of lake development, during which the basin reached its greatest paleowater depth (Fig. 1(b); Wang et al., 2023b; Wang et al., 2024) and is characterized by widespread deposition of fine-grained sediments. It is noteworthy that the T_3y_7—currently in the peak hydrocarbon generation stage—also holds significant potential for future resource addition and production enhancement (Niu et al., 2025). This study therefore focuses on the T_3y_7 as the primary target of investigation.

3. Data and methods

3.1. Samples, experiments, and data

Samples were collected from the T_3y_7 in the YYa borehole, located in the southern Ordos Basin (sampling location shown in Fig. 1(a), 39 samples). Geochemical analyses included X-ray diffraction (XRD), total organic carbon (TOC), X-ray fluorescence (XRF), and inductively coupled plasma mass spectrometry (ICP-MS). XRD and TOC analyses were conducted at the SINOPEC Wuxi Institute of Petroleum Geology, while XRF and ICP-MS analyses were carried out at the Lanzhou Oil and Gas Resources Research Key Laboratory, Institute of Geology and Geophysics, Chinese Academy of Sciences. Every collected sample underwent all the experimental tests mentioned above, generating a total of 156 groups of data. For XRD analysis, an Ultima IV diffractometer was operated at 40 kV and 40 mA. Rock chips were crushed and ground to fine powders (<75 μm, 200 mesh), and the powders were mounted on glass slides for measurement. Diffraction patterns were recorded over a 2θ range of 5° to 45°. Notably, the traditional peak intensity comparison method was employed, not the Rietveld refinement method. Prior to TOC analysis, samples were also ground to 200 mesh, treated with hydrochloric acid to remove carbonates, rinsed with deionized water, and dried in a thermostatic oven at 80 °C for 4 h. TOC content was then measured using a Leco CS230 carbon/sulfur analyzer.

All samples were examined under a microscope before geochemical analysis to assess potential diagenetic alterations, mineralization, or secondary weathering. Samples were then ground using a contamination-free crusher, sieved to 200 mesh, and oven-dried at 80 °C for 3 h to remove moisture. Major element concentrations were measured using a 3080E3X X-ray fluorescence spectrometer (RIGAKU, Japan). For trace element analysis, samples were digested using a mixture of HF and HNO_3, followed by quantification via ICP–MS.

To ensure a sufficient sample size for developing interpretable machine learning models with robust performance and generalizable conclusions, we integrated our newly analyzed samples with 261 additional samples (in 20 boreholes and 4 outcrops, locations are shown in Fig. 1) from eight published studies on lacustrine shale systems in the T_3y_7 of the Ordos Basin. These studies provided datasets containing both TOC content and key major and trace element concentrations (Wang et al., 2017; Zhang et al., 2017; Li et al., 2020; Yuan et al., 2020; Pan et al., 2022; Shi et al., 2022; Zheng et al., 2022; Yang et al., 2022).

3.2. Paleoenvironment proxies

Based on our dataset, a total of 26 elemental geochemical proxies were selected to reconstruct the paleoenvironmental conditions. Because of volcanic activity can alter the distribution patterns and concentrations of rare earth elements in sediments (Hulya, 2016; Zhang et al., 2020). Therefore, the sum of rare earth elements ($\sum$REE), $(La/Yb)_N$, δCe_N, and δEu_N (N is normalized to the upper continental crust (UCC), Bau and Dulski, 1996; Ballard et al., 2002) were used to indicate volcanic activity (Hulya, 2016; Zhang et al., 2020). Hydrothermal activity was inferred using Al/(Al + Fe + Mn), (Fe + Mn)/Ti, and the cerium anomaly ($\delta Ce^* = 2Ce_n/(La_n + Pr_n)$, n is the PAAS standardization), which have been widely employed as hydrothermal indicators (Boström and Peterson, 1969; Adachi et al., 1986; Bau et al., 2014; He et al., 2016; Zhang et al., 2019).

Paleoclimatic and weathering conditions were assessed using the *C*-value, chemical index of alteration (CIA), and Sr/Cu ratio (Minyuk et al., 2007; Moradi et al., 2016; Wu et al., 2021; Lv et al., 2022). To reflect primary productivity, we utilized element/Al ratios including P/Al, Ba/Al, Cu/Al, Ni/Al, and Zn/Al (Tribovillard et al., 2006; Algeo and Ingall, 2007). Redox conditions were characterized using multiple proxies, including ratios of redox-sensitive elements (e.g., V/Cr, V/(V + Ni), Ni/Co, and U/Th; Jones and Manning, 1994; Wignall and Twitchett, 1996; Tribovillard et al., 2006), enrichment factors of redox-sensitive elements (Mo_{EF} and U_{EF}; Algeo and Li, 2020), and other indicators such as the Corg/P ratio (Algeo and Ingall, 2007). Since Al and Ti are relatively immobile during sedimentary processes and resistant to diagenetic alteration, their concentrations, along with the Ti/Al ratio, were also employed (Wedepohl, 1971; Tribovillard et al., 2006). Additionally, due to the differential solubility of barite ($BaSO_4$) and celestite ($SrSO_4$) under varying salinity conditions, the Sr/Ba ratio was used as a salinity proxy (Wei and Algeo, 2020).

3.3. Data processing and machine learning model

3.3.1. Data processing

The data preprocessing and model development workflow proposed in this study is illustrated in Fig. 2. Prior to dataset construction, raw data were systematically preprocessed.

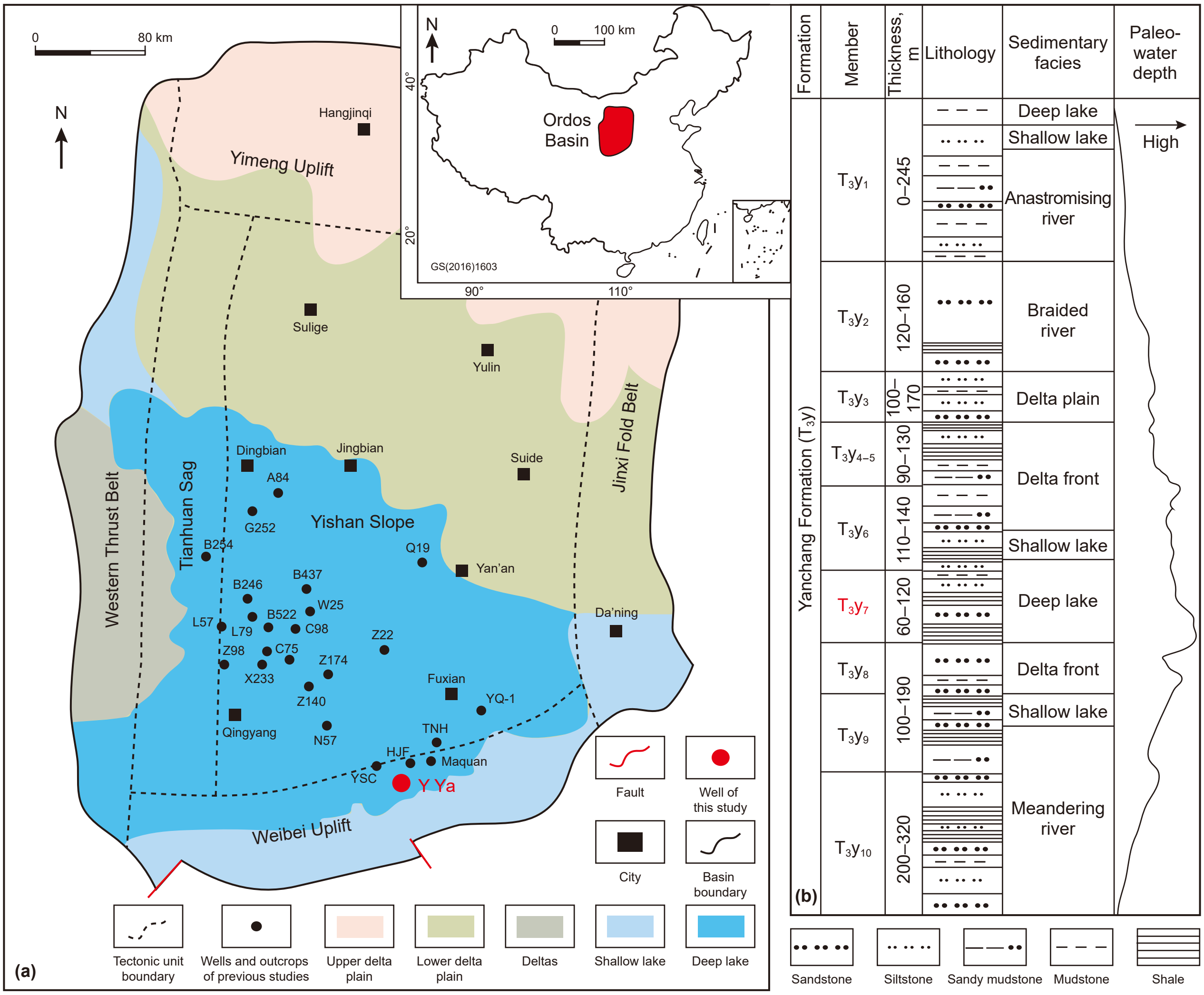

Fig. 1. Comprehensive geological map of the study area. (**a**) The location of three borehole of this study (From Chen et al., 2017); (**b**) Strata column of the Yanchang Formation in the Erdos Basin (Modified from Qiao et al., 2021).

Specifically, the dataset comprising 300 samples was divided into features and the target variable. The features included the 26 elemental geochemical proxies described previously, while the target variable was TOC content. Outliers within the feature set were identified using the interquartile range (IQR) method and subsequently treated as missing values.

To address these missing values, the MissForest algorithm proposed by Stekhoven and Bühlmann (2012) was employed. MissForest is an iterative imputation method based on random forest models. Unlike conventional imputation techniques, it captures complex interactions and nonlinear relationships among variables. Initially, missing values are imputed using a simple heuristic. In each iteration, a random forest regressor is trained using observed values and previously imputed estimates as predictors. The trained model is then used to update the missing values of the target variable. This process is repeated until a convergence criterion is met—typically, either a predefined maximum number of iterations or the stabilization of imputed values (Stekhoven and Bühlmann, 2012; Wang et al., 2025a). The generalization performance of the imputation model, specifically its ability to accurately predict missing values, was evaluated using the out-of-bag (OOB) score and the coefficient of determination from five-fold cross-validation ($R^2{}_{cv}$), both of which can be calculated without the need for an external test set (Liao et al., 2023). Proxies with excessively low OOB and $R^2{}_{cv}$ values indicate that the imputation of missing data was unsatisfactory. Such proxies will be excluded from subsequent analysis. The initial imputation settings and iteration termination criteria in this study followed the framework proposed by Liao et al. (2023) and were implemented in Python (version 3.10). The features were initially sorted by the number of missing values, and then the mean value was used for imputation. Next, for each feature, MissForest fits a RF on the observed part and apply the fitted model to predict and replace the missing part. γ is the stopping criterion and is calculated as the quotient of the square of the difference between the data before and after iteration and the data after iteration, the iteration process is stopped as soon as γ increases for the first time. Following the preprocessing procedures, a clean and complete dataset was obtained for subsequent model development.

3.3.2. Model development and evaluation

Prior to model development, the dataset was randomly partitioned into training and testing subsets. The training set, comprising 80% of the data, was used to develop ML models, while the remaining 20% was reserved as a test set to evaluate the generalization performance of the final model. To mitigate the

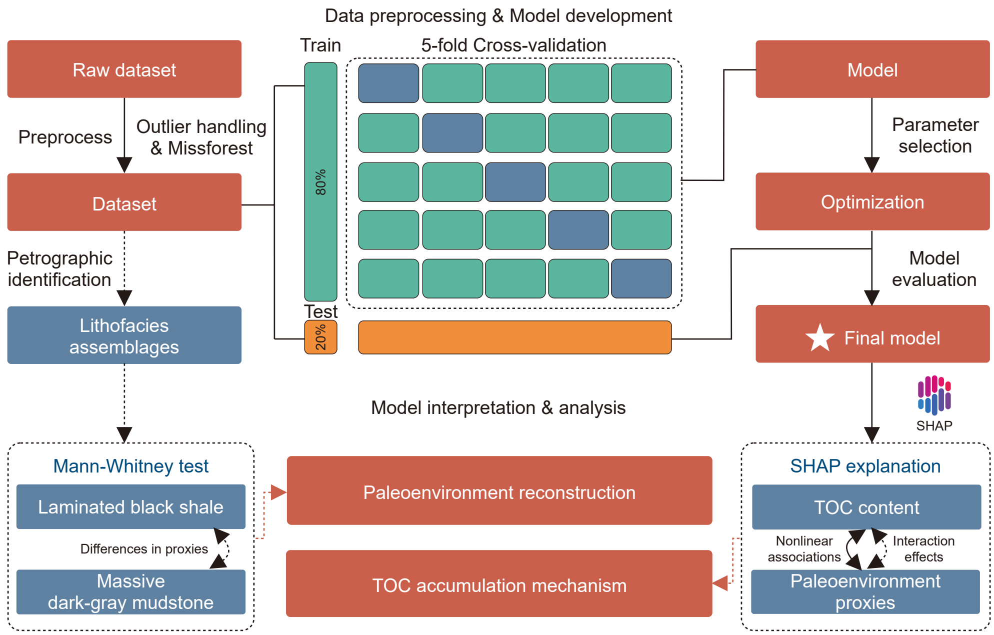

Fig. 2. Schematic of the workflow in this work.

impact of scale differences among features on model performance, Z-score normalization was applied to the training set after the data split. The mean and standard deviation calculated from the training data were then used to normalize the test set, thereby avoiding data leakage during model training (James et al., 2023; Wang et al., 2025a).

To identify the optimal predictive model for TOC content, three widely adopted ML algorithms were implemented and compared: random forest (RF), extreme gradient boosting (XGBoost), and support vector machine (SVM). These three algorithms (RF, XGBoost, SVM) were selected primarily due to their proven robustness in handling non-linear data and reducing overfitting, which are critical characteristics for the complex feature relationships present in the dataset of this study. The detailed hyperparameter settings for each model are summarized in Table 1. All model training and evaluation procedures were conducted using the "scikit-learn" library in Python.

3.3.2.1. Machine learning models. RF is a robust ensemble learning technique based on the bagging algorithm, which is widely applicable to both regression and classification tasks. It constructs an ensemble of decision trees and aggregates their predictions to generate a final output. RF employs bootstrap sampling, whereby both features and samples are randomly selected with replacement (Breiman, 2001). This randomness helps mitigate overfitting and enhances the capability of the model to handle relatively small datasets effectively (Fan et al., 2022; Palansooriya et al., 2022; Zhong et al., 2022). XGBoost, a highly efficient and scalable implementation of gradient boosting decision trees, has been widely adopted in various domains (Chen et al., 2019; Liu et al., 2023). Unlike RF, which builds trees independently, XGBoost employs a sequential boosting strategy, where each new tree is trained to correct the errors of the previous ones. It modifies the sample weights at each iteration, assigning greater emphasis to those instances with higher prediction errors, thereby improving model performance on difficult cases (Chen and Guestrin, 2016; Otchere et al., 2021). SVM is a non-probabilistic, binary linear classifier designed to assign new data points to one of two categories (Cortes and Vapnik, 1995). It maps data into a high-dimensional space and identifies the optimal hyperplane that maximizes the margin between classes. In regression tasks, SVM seeks to construct a hyperplane that fits the data while maximizing the distance (margin) between the predicted and actual values, especially focusing on minimizing deviations beyond a predefined threshold (Drucker et al., 1996).

To optimize the hyperparameters of these ML models, we adopted the differential evolution (DE) algorithm, a population-based global optimization strategy (Bilal et al., 2020). DE iteratively generates new candidate solutions through differential mutation and crossover operations. It is particularly well-suited for continuous parameter spaces and is recognized for its simplicity, strong global search capability, and consistent performance across various optimization problems. Compared to more conventional hyperparameter tuning methods, DE avoids exhaustive parameter evaluation, significantly improving efficiency in high-dimensional search spaces. It also reduces the risk of local optima via stochastic mechanisms, a limitation that often affects conventional methods and compromises geoscience ML model performance.

Table 1
Hyperparameter setting information during model development.

Target model		Hyperparameter setting
ML model	RF	n_estimators (100-300), max_depth(3-50)
	XGBoost	n_estimators (100-300), max_depth(3-50), learning_rate(0.05-0.5)
	SVM	C(0.1-10), gamma(0.1-10)
Optimization method	DE	size_pop(40), max_iter(50)

3.3.2.2. Model evaluation. During model development and evaluation, standard regression metrics were employed, specifically the coefficient of determination (R^2, Eq. (1)) and root mean squared error (RMSE, Eq. (2)). These metrics are formally defined as follows.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{1}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i}(y_i - \hat{y}_i)^2} \tag{2}$$

where $\hat{y}_i$ are the predicted data, y_i are the actual values, $\bar{y}$ is the actual data mean, n is the number of points for each dataset.

In addition, a 5-fold cross-validation strategy was applied to the training dataset to enhance the generalization capability of the models (Gjelsvik et al., 2023; Wang et al., 2023c). By systematically evaluating the average predictive performance, expressed as cross-validated R^2 (R^2_{CV}) and RMSE (RMSE_{CV}), across the five folds, we were able to incorporate optimization algorithms into the model selection process. This approach facilitated the comprehensive exploration of model configurations and enabled the robust identification of the optimal model for the given regression task.

3.3.3. Model interpretation

3.3.3.1. Mann–Whitney U test. The non-parametric Mann–Whitney U test was employed to rigorously assess differences in specific proxies between lithofacies groups. This statistical method was chosen for its appropriateness in comparing two independent samples without the assumption of normal data distribution. A significance level (α) of 0.05 was adopted, indicating a 5% probability of incorrectly concluding that a statistically significant difference exists when, in fact, it does not. This threshold is widely accepted across scientific disciplines as a benchmark for statistical significance, ensuring the robustness and reliability of analytical conclusions (Lu et al., 2025). All statistical tests were conducted using the Scipy.stats package Python 3.10.

3.3.3.2. Shapley additive explanations method. To interpret the contributions of individual features to model predictions, the shapley additive explanations (SHAP) method was utilized. As a game-theoretic approach, SHAP quantifies the marginal contribution of each input feature to a model's output by calculating the mean absolute SHAP value across all training instances (Shapley, 1952; Lundberg and Lee, 2017). A positive (or negative) SHAP value for a given feature in an instance indicates that the feature contributes positively (or negatively) to the predicted output. The magnitude of the SHAP value reflects the strength of this contribution (Mangalathu et al., 2022; Nazar et al., 2023; Nadege et al., 2024). SHAP provides a global interpretation of model behavior by aggregating local feature attributions while maintaining fidelity to the original model at the instance level (Lundberg et al., 2020; Tamoto et al., 2023). The SHAP value can be formally defined as follows:

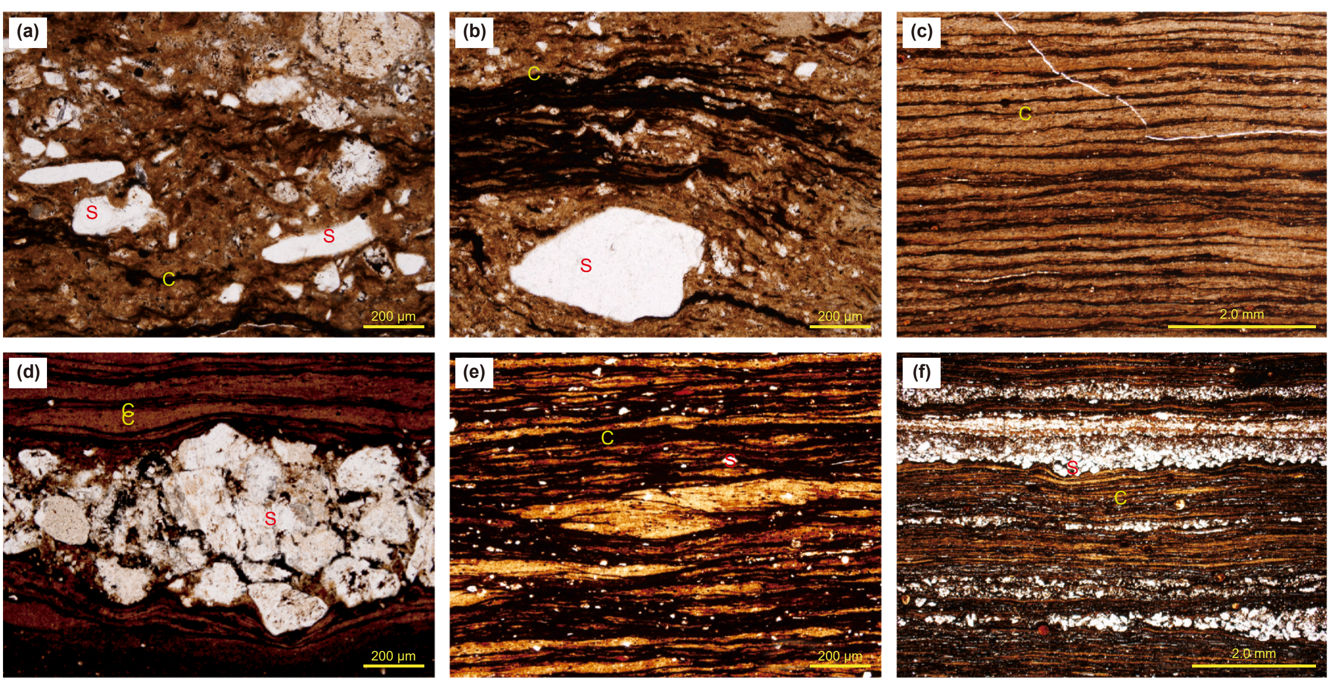

Fig. 3. Sedimentary textures of the Mbr 7 of the Yanchang Formation in YYa borehole. (**a**) 211.80 m, dark-gray, massive texture; (**b**) 211.06 m, dark-gray mudstone, massive texture; (**c**) 229.60 m, black shale, laminated texture; (**d**) 228.44 m, black shale, lenticular quartz; (**e**) 229.50 m, black shale, laminated texture; (**f**) 229.60 m, black shale, laminated texture. Note: S: siliceous minerals; C: clay minerals.

Table 2
Mineral composition, TOC content, sedimentary texture of black shale and dark-gray mudstone of the T_3y_7 shale from YYa borehole of the Ordos Basin.

Lithofacies assemblages	Clay mineral, %	Quartz, %	Potassium feldspar, %	Plagioclase, %	Calcite, %	Dolomite, %	Pyrite, %	TOC content, %	Sedimentary texture
Black shale	28.8–57.3 (40.6)	13.7–29.1 (21.8)	0.4–5.0 (1.8)	1.4–11.5 (3.8)	0.7–13.1 (3.5)	0.2–7.2 (2.3)	13.1–38.3 (20.6)	5.87–25.25 (13.44)	Laminated
Dark-gray mudstone	15.3–59.3 (41.0)	25.8–61.7 (34.9)	0.1–16.8 (2.1)	0.3–13.2 (5.9)	0.3–25.6 (3.7)	0.4–16.0 (4.4)	0.3–8.7 (3.5)	0.44–4.29 (1.76)	Massive

Note: Minimum–maximum (average).

$$\phi_j(\mathrm{val}) = \sum_{S\subseteq\{1,\ldots,p\}\setminus\{j\}} \frac{|S|!(p-|S|-1)!}{p!}(\mathrm{val}(S\cup(j)) - \mathrm{val}(S)) \tag{3}$$

where S is a subset of the features utilized in the model, x is the vector of feature values for the instance being explained and p is the number of features. val(S) is the prediction for feature values in set S, marginalized over the features not included in set S:

$$\mathrm{val}_x(S) = \int \widehat{f}(x_1,\ldots,x_p)\mathrm{d}\mathbb{P}_{x\notin S} - E_x(\widehat{f}(X)) \tag{4}$$

According to Lundberg and Lee (2017), SHAP values are an additive feature attribution method. SHAP specifies the explanation as:

$$g(z') = \phi_0 + \sum\nolimits_{j=1}^{M}\phi_j z'_j \tag{5}$$

where g is the explanation model, $z' \in \{0,1\}^M$ is the coalition vector, M is the maximum coalition size and $\phi_j \in \mathbb{R}$ is the feature attribution for a feature j, the SHAP values.

In addition to main effects, SHAP also accounts for pairwise interaction effects between variables through SHAP interaction values (Kim and Lee, 2023). These interaction values are computed by integrating contributions from feature pairs and help uncover complex dependencies in the model (Xiao et al., 2021). The SHAP interaction index is formally defined as follows:

$$\phi_{ij} = \sum\nolimits_{S\subseteq\{1,\ldots,p\}\setminus\{i,j\}} \frac{|S|!(p-|S|-2)!}{2(p-1)!}\delta_{ij}(s) \tag{6}$$

$$\delta_{ij}(s) = f(S\cup\{i,j\}) - f(S\cup\{i\}) - f(S\cup\{j\}) + f(S) \tag{7}$$

where ϕ_{ij} represents the SHAP interaction value between feature i and feature j, $f(S\cup\{i,j\})$ represents the model results with both feature i and j. In this study, we applied the TreeExplainer in SHAP package (Python 3.10) to calculate both SHAP values and SHAP interaction values (Kim and Lee, 2023).

4. Results

4.1. Lithofacies and paleoenvironment proxies

Lithofacies represent the fundamental units for describing the geological characteristics of sedimentary deposits, and distinct lithofacies often reflect contrasting depositional environments. Coupled with variations in hydrocarbon generation and diagenetic conditions, each lithofacies exhibits unique geological attributes and resource potential (Dapples et al., 1948; Ma et al., 2022). For the Yanchang Formation in the Ordos Basin, previous studies have identified two dominant lithofacies assemblages, commonly referred to as dark-gray mudstone and black shale. These two lithofacies assemblages differ significantly in terms of mineral composition, TOC content, kerogen type, and sedimentary texture (Zhao et al., 2020; Wang et al., 2025b).

Specifically, compared to dark-gray mudstone, black shale is generally characterized by higher contents of brittle minerals and TOC content, along with superior kerogen quality. Their sedimentary textures also differ considerably (Zhao et al., 2020; Wang et al., 2025b). In the YYa borehole of the study area, the observed distribution of lithofacies assemblages is consistent with previous findings (Zhao et al., 2020), with both dark-gray mudstone and black shale being clearly identifiable. Their geological characteristics are illustrated in Fig. 3 and summarized in Table 2. A consistent pattern in sedimentary textures is observed: samples

Table 3
The TOC content and paleoenvironment proxies of the laminated black shale and massive dark-gray mudstone of the T_3y_7 shale from YYa borehole of the Ordos Basin.

Lithofacies	TOC content, %	ΣREE	$(La/Yb)_N$	δCe_N	δEu_N	Al/(Al + Fe + Mn)	(Fe + Mn)/Ti	δCe*	CIA	C-value	Sr/Cu	P/Al	Ba/Al
Laminated black shale	5.87–25.25 (13.44)	112.71–200.41 (155.82)	0.78–6.06 (1.59)	0.42–0.46 (0.44)	0.39–0.62 (0.49)	0.26–0.93 (0.60)	5.32–54.77 (20.49)	0.85–0.94 (0.89)	70.20–96.53 (79.41)	0.18–3.72 (1.25)	0.47–32.69 (6.70)	0.0019–0.1186 (0.0258)	0.0043–0.0177 (0.0080)
Massive dark-gray mudstone	0.44–4.29 (1.76)	116.49–247.67 (176.76)	0.07–2.18 (1.01)	0.36–0.51 (0.44)	0.11–0.57 (0.45)	0.26–0.87 (0.70)	6.43–36.45 (11.10)	0.72–1.03 (0.90)	64.78–82.21 (74.67)	0.27–1.07 (0.52)	3.06–90.23 (16.03)	0.0013–0.0571 (0.0092)	0.0031–0.0142 (0.0065)
Cu/Al	Ni/Al	Zn/Al	V/Cr	V/(V + Ni)	U/Th	Ni/Co	Mo_{EF}	U_{EF}	Corg/P	Al (ppm)	Ti (ppm)	Ti/Al	Sr/Ba
0.000012–0.00356 (0.00106)	0.000033–0.000716 (0.000332)	0.000441–0.00207 (0.00101)	1.61–5.63 (3.43)	0.69–0.92 (0.84)	0.27–12.62 (4.35)	0.21–2.42 (1.35)	0.45–207.03 (59.34)	1.15–61.44 (19.56)	22.26–2401.47 (597.28)	41029.41–110647.06 (80056.40)	1111.55–4061.31 (2951.40)	0.0129–0.0536 (0.0385)	0.0498–0.6213 (0.3545)
0.000037–0.00114 (0.000345)	0.000040–0.000776 (0.000295)	0.000447–0.00148 (0.000922)	0.86–4.51 (1.74)	0.55–0.94 (0.77)	0.19–3.46 (0.47)	0.27–6.59 (2.30)	0.18–59.89 (5.38)	0.49–13.35 (2.09)	16.93–1096.97 (128.38)	30335.29–105293.34 (85621.09)	792.27–5096.34 (3812.99)	0.0102–0.0796 (0.0449)	0.267–1.110 (0.506)

Note: Minimum–maximum (average).

Table 4
Imputation performance after the MissForest process.

Evaluation parameters	ΣREE	$(La/Yb)_N$	δCe_N	δEu_N	Al/(Al + Fe + Mn)	(Fe + Mn)/Ti	δCe*	CIA	*C*-value	Sr/Cu	P/Al	Ba/Al	Cu/Al
R^2_{train}	0.94	0.92	1.00	0.97	0.99	0.99	1.00	0.94	0.97	0.95	0.96	0.92	0.99
R^2_{CV}	0.50	0.48	0.98	0.74	0.93	0.93	0.99	0.55	0.76	0.63	0.71	0.38	0.89
OOB score	0.58	0.47	0.98	0.79	0.94	0.93	0.99	0.57	0.80	0.66	0.72	0.45	0.91
Ni/Al	Zn/Al	V/Cr	V(V + Ni)	U/Th	Ni/Co	Mo_{EF}	U_{EF}	Corg/P	Al	Ti	Ti/Al	Sr/Ba	
0.97	0.97	0.98	0.96	0.99	0.95	0.99	0.99	0.95	0.98	0.99	0.96	0.94	
0.83	0.80	0.86	0.70	0.90	0.49	0.86	0.93	0.66	0.84	0.91	0.65	0.53	
0.82	0.77	0.87	0.73	0.92	0.57	0.89	0.94	0.65	0.84	0.92	0.70	0.54	

classified as black shale predominantly exhibit laminated textures, whereas those assigned to dark-gray mudstone display massive textures (Fig. 3). Additionally, the TOC content of black shale ranges from 5.87% to 25.25%, with an average of 13.44%, while that of dark-gray mudstone spans from 0.44% to 4.29%, averaging 1.76%. From the perspective of mineral composition, the framework grain content of black shale (35.3%–60.9%, average 47.9%) is generally higher than that of dark-gray mudstone (31.0%–63.3%, average 43.6%). Notably, the elevated framework grain content in black shale is primarily attributed to its higher pyrite content (13.1%–38.3%, average 20.6%), rather than to quartz or feldspar abundance. Based on these geological characteristics, two distinct lithofacies assemblages, laminated black shale and massive dark-gray mudstone, can be identified in the YYa borehole. The

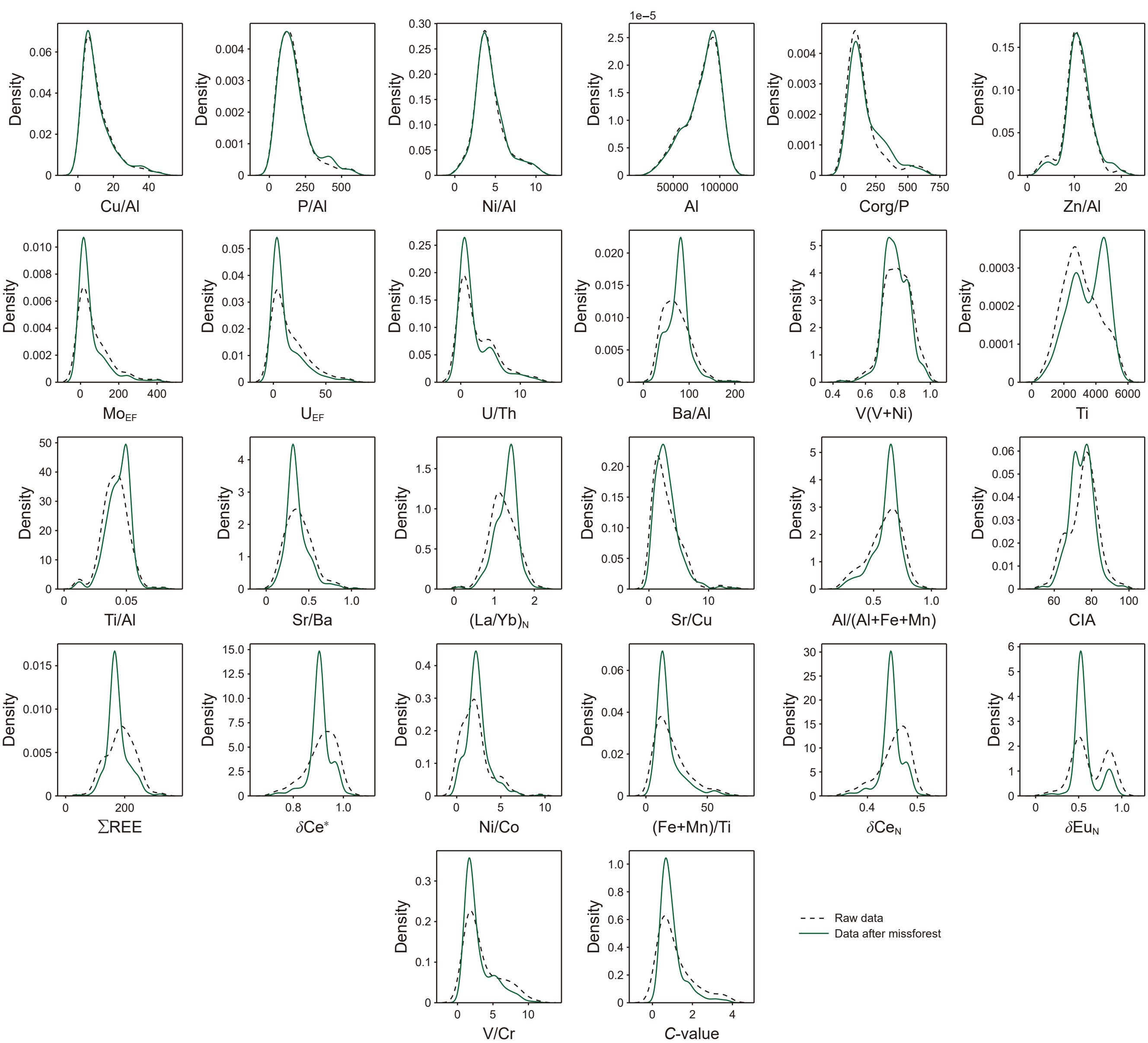

Fig. 4. Comparison of data distribution between the raw dataset and the data after the MissForest process.

distributions of paleoenvironmental proxies for laminated black shale and massive dark-gray mudstone are presented in Table 3. The marked differences in proxy values between the two lithofacies assemblages further indicate distinct depositional environments during the time of sedimentation. A detailed paleoenvironmental reconstruction based on these proxies is provided in Section 5.1.

4.2. *Establishment of the interpretable model for organic matter enrichment mechanism*

4.2.1. *Data preprocessing result*

As noted in Section 3.3, to systematically investigate the organic matter enrichment mechanisms in the T_3y_7 of the Ordos Basin, we performed data interpolation on previously published datasets based on the database established in this study. The R^2cv and OOB scores for the interpolated proxies are summarized in Table 4. The results indicate that, except for the Ba/Al ratio, which exhibited poor interpolation performance, the interpolation for the remaining proxies was generally reliable. Most proxies achieved OOB scores greater than 0.7, suggesting that the imputation results are trustworthy. To ensure the robustness of subsequent interpretations, the Ba/Al ratio was excluded from further discussion. The distribution patterns of the raw and interpolated data are presented in Fig. 4. Overall, the distributions remained largely consistent after interpolation, confirming that the procedure did not significantly distort the underlying data characteristics (Table 5).

4.2.2. *Evaluation of the model performance*

Following the development and tuning of machine learning models, the optimal hyperparameter configurations for each algorithm were determined, as listed in Table 6 along with their corresponding performance metrics. Among the evaluated models, the RF-based model demonstrated superior performance. The performance indicators of the optimal RF model are shown in Fig. 5. For this model, the coefficient of determination (R^2) for the training and testing sets reached 0.97 and 0.82, respectively, while the corresponding RMSE values were 0.81 and 3.58. These results highlight the RF model's strong predictive capability and generalization performance.

4.3. *Model interpretation result*

To interpret the optimal TOC prediction model, SHAP analysis was conducted to evaluate the contribution of each feature to the predictions of model. As shown in Fig. 6, the SHAP summary plot illustrates both the magnitude and direction of each geochemical proxy's impact on TOC content. Features are ranked along the y-axis in descending order of their overall SHAP importance. For a specific sample, a positive (or negative) SHAP value of a proxy indicates that the proxy has a positive (or negative) contribution to TOC. The magnitude of the SHAP value reflects the strength of this contribution. The top ten most influential proxies are V/Cr, Cu/Al, U/Th, Corg/P, U_{EF}, Mo_{EF}, Ba/Al, V/(V + Ni), Sr/Cu, and (Fe + Mn)/Ti. Among these, all but Sr/Cu exhibit relatively clear positive contributions to TOC content. From a paleoenvironmental perspective, redox-sensitive indicators are the most influential in predicting TOC content, accounting for 64.4% of the total model contribution. Most redox-related proxies display a positive association, suggesting that more reducing conditions are generally favorable for TOC accumulation. In addition, productivity-related proxies contribute 19.3% to the overall prediction performance and also show a positive relationship with TOC, indicating that higher primary productivity tends to enhance organic matter enrichment.

Table 5
Paleoenvironment proxies after data interpolation.

Lithofacies	TOC content, %	ΣREE	$(La/Yb)_N$	δCe_N	δEu_N	Al/(Al + Fe + Mn)	(Fe + Mn)/Ti	δCe^*	CIA	C-value	Sr/Cu	P/Al	Ba/Al
Laminated black shale	5–39.39 (14.52)	45.1–281.8 (169.08)	0.61–2.00 (1.30)	0.36–0.50 (0.44)	0.32–0.97 (0.58)	0.25–0.93 (0.56)	5.32–68.69 (22.45)	0.74–1.01 (0.90)	60.85–97.05 (77.34)	0.18–3.74 (1.24)	0.17–9.46 (1.97)	18.86–583.94 (210.45)	13.00–198.24 (77.99)
Massive dark-gray mudstone	0.03–4.95 (2.19)	88–323.5 (183.89)	0.07–2.18 (1.28)	0.36–0.51 (0.45)	0.11–0.88 (0.56)	0.26–0.87 (0.68)	6.43–36.45 (11.89)	0.72–1.03 (0.92)	52.98–83.04 (70.21)	0.27–1.07 (0.61)	1.56–14.16 (4.32)	12.82–570.56 (118.18)	31.09–142.37 (74.70)
Cu/Al	Ni/Al	Zn/Al	V/Cr	V/(V + Ni)	U/Th	Ni/Co	Mo_{EF}	U_{EF}	Corg/P	Al (ppm)	Ti (ppm)	Ti/Al	Sr/Ba
1.24–47.06 (15.40)	0.33–10.61 (5.06)	2.68–20.69 (11.90)	0.73–10.99 (3.84)	0.46–0.97 (0.82)	0.26–12.73 (3.73)	0.21–5.52 (1.94)	0.45–413.03 (85.21)	0.97–74.06 (18.24)	22.26–621.45 (236.24)	26600–110647.06 (74856.43)	1111.55–5000 (3113.18)	0.0129–0.0705 (0.0421)	0.0498–1.036 (0.351)
0.37–11.35 (4.97)	0.40–7.76 (3.46)	4.47–14.80 (10.19)	0.86–4.51 (1.63)	0.55–0.94 (0.75)	0.19–6.54 (0.55)	0.27–6.78 (2.68)	0.18–59.89 (10.41)	0.49–13.35 (2.74)	7.2–622.74 (87.16)	30335.29–113000 (90707.96)	792.27–5513.56 (4042.69)	0.0102–0.0796 (0.0451)	0.15–0.86 (0.37)

Note: Minimum–maximum (average).

Table 6
ML model evaluation results.

Models	Parameters					Performance	
	N_estimators	Max_depth	Learning_rate	C	Gamma	R^2	RMSE
RF	124	15	/	/	/	0.82	3.45
XGBoost	153	4	0.24	/	/	0.75	4.16
SVM	/	/	/	9	0.13	0.78	3.95

Although the contributions of proxies representing paleoclimate, hydrothermal activity, terrigenous input, and salinity are relatively minor, proxies within each of these categories exhibit internally consistent trends in their influence on TOC content. In contrast, proxies reflecting volcanic activity show only negligible effects on TOC content, with $(La/Yb)_N$ being the only one showing a weak positive trend. This implies that volcanic activity may not be conducive to TOC enrichment.

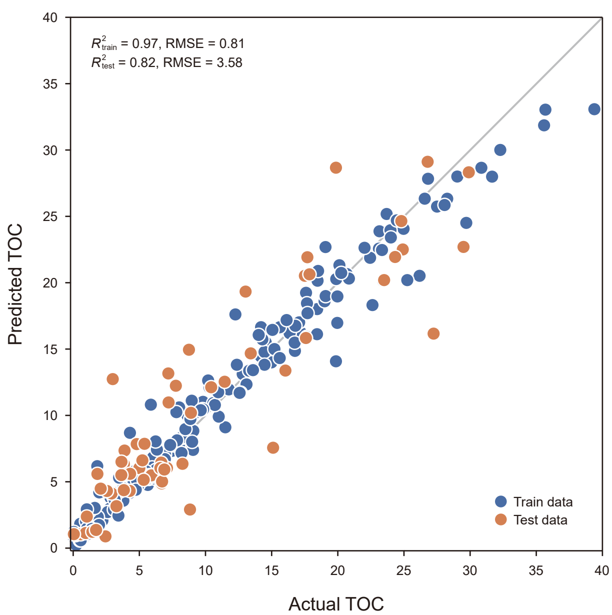

Fig. 5. Optimal model predictive performance of TOC content.

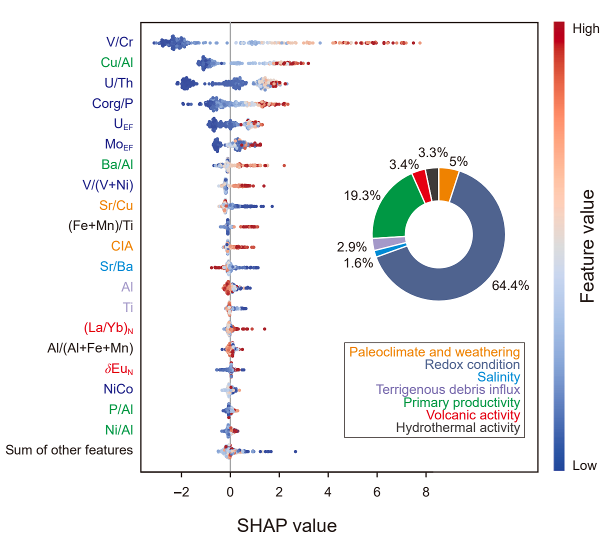

Fig. 6. SHAP values of different proxies for the TOC content prediction model.

Based on the SHAP summary plot results, partial dependence analysis was conducted for representative proxies within each category to further explore how paleoenvironmental factors influence TOC content patterns. The selection of representative proxies was primarily guided by their SHAP-derived importance. As shown in Fig. 7, V/Cr, Cu/Al, (Fe + Mn)/Ti, and $(La/Yb)_N$ show positive contributions to TOC content, while Al, Sr/Cu, and salinity exert negative influences. Among these, V/Cr and Al demonstrate approximately linear relationships with TOC content, whereas Sr/Cu, Cu/Al, (Fe + Mn)/Ti, and $(La/Yb)_N$ exhibit nonlinear response patterns. Specifically, the positive contribution of Sr/Cu diminishes when values are below 2 and becomes negligible above that threshold. Cu/Al shows an increasing positive effect on TOC when values are below 20, stabilizing thereafter. The impact of (Fe + Mn)/Ti on TOC is minor when values are below 30, but becomes significantly positive above this threshold. Similarly, $(La/Yb)_N$ exhibits a minimal effect when below 1.5, yet begins to positively influence TOC once it exceeds that value. These nonlinear trends suggest complex relationships between paleoenvironmental conditions and organic matter accumulation.

5. Discussion

5.1. Differences of paleoenvironment proxies in laminated black shale and massive dark-gray mudstone

As previously mentioned, both existing studies and our dataset indicate the presence of two predominant lithofacies assemblages within the T_3y_7 shale of the Ordos Basin: laminated black shale and massive dark-gray mudstone (Zhao et al., 2020; Wang et al., 2025b). In most cases, these two lithofacies assemblages can be effectively distinguished based on their TOC content. Accordingly, this study adopts TOC content as the primary criterion for lithofacies identification within previously published datasets. The boundary value for distinguishing laminated black shale from massive dark-gray mudstone typically ranges from 5% to 6% TOC, as reported in the literature (Zhang et al., 2021; Liu et al., 2022; Wang et al., 2025b). Based on our data and these prior studies, a TOC threshold of 5% was selected in this study to delineate the two lithofacies assemblages and facilitate consistent classification across the integrated dataset. Following lithofacies identification, we conducted a comprehensive statistical analysis of paleoenvironmental proxies for the T_3y_7 shale based on both our own and previously published data (Table 3), providing a foundation for paleoenvironmental reconstruction.

5.1.1. Paleoclimate and weathering

Given that different elements exhibit distinct accumulation behaviors under varying climatic and weathering conditions, the CIA, *C*-value, and Sr/Cu ratio are widely recognized as effective proxies for assessing paleoclimate and chemical weathering intensity (Minyuk et al., 2007; Moradi et al., 2016; Wu et al., 2021; Lv et al., 2022). The distributions of CIA, *C*-value, and Sr/Cu ratio for different lithofacies assemblages are illustrated in Fig. 8. Statistical

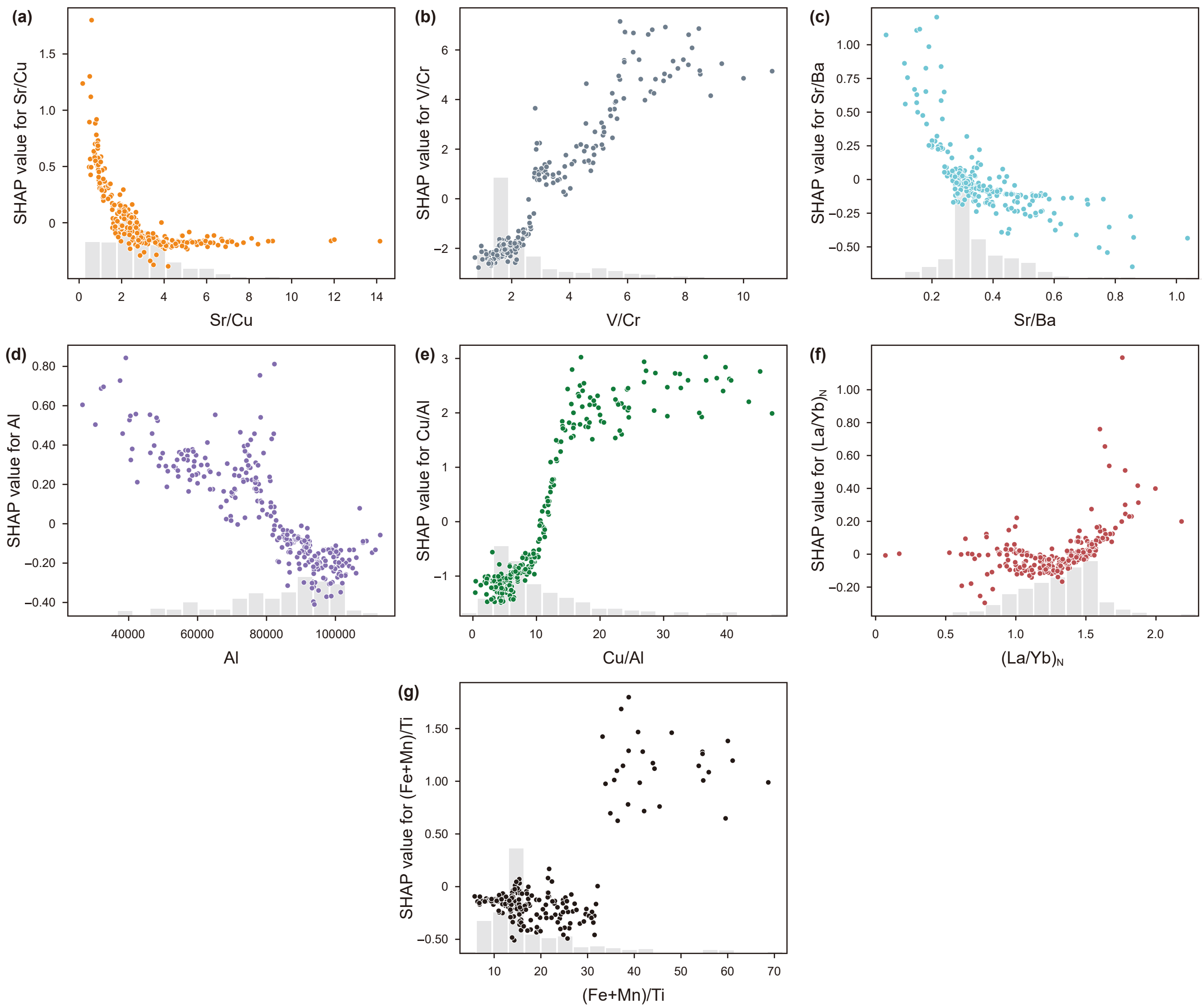

Fig. 7. SHAP partial dependence of the typical proxies for the TOC content prediction model.

tests reveal that the probability values (*P*-values) for all three proxies are below 0.001, indicating strong statistical significance.

The results show that laminated black shale exhibits markedly higher CIA and *C*-value values (average CIA 77.4, average *C*-value 1.24) than massive dark-gray mudstone (average CIA 70.2, average *C*-value 0.61). Conversely, the Sr/Cu ratio is significantly lower in laminated black shale (average 1.97) than in massive dark-gray mudstone (average 4.32). These findings suggest that the laminated black shale was deposited under more humid climatic conditions with stronger chemical weathering intensity compared to the massive dark-gray mudstone. This conclusion is further supported by the mineralogical composition of the samples. As shown in Table 2, laminated black shale contains substantially lower proportions of unstable minerals such as feldspar relative to massive dark-gray mudstone, indirectly corroborating the interpretation of a more humid and strongly weathered depositional environment. Collectively, the CIA and *C*-value data indicate that the laminated black shale was primarily deposited in a humid climate with moderate to strong chemical weathering, whereas the massive dark-gray mudstone formed under semi-humid conditions with moderate weathering intensity.

5.1.2. Redox condition and salinity

Due to the differing accumulation patterns of redox-sensitive elements under varying redox conditions, this study employed V/Cr, V/(V + Ni), Ni/Co, U/Th, Mo_{EF}, U_{EF}, and Corg/P as proxies for redox conditions (Jones and Manning, 1994; Wignall and Twitchett, 1996; Tribovillard et al., 2006). The distributions of these proxies for different lithofacies assemblages within the T_3y_7 shale are illustrated in Fig. 9. All seven proxies yielded *P*-values less than 0.001, indicating strong statistical significance. Except for the Ni/Co ratio, the proxy values in laminated black shale are significantly higher than those in massive dark-gray mudstone, suggesting that the depositional environment of laminated black shale was more reducing. As emphasized by Algeo and Liu (2020), redox studies should incorporate multiple geochemical proxies to ensure robust interpretations. The relatively weak discrimination power of the Ni/Co ratio in this context may suggest that this

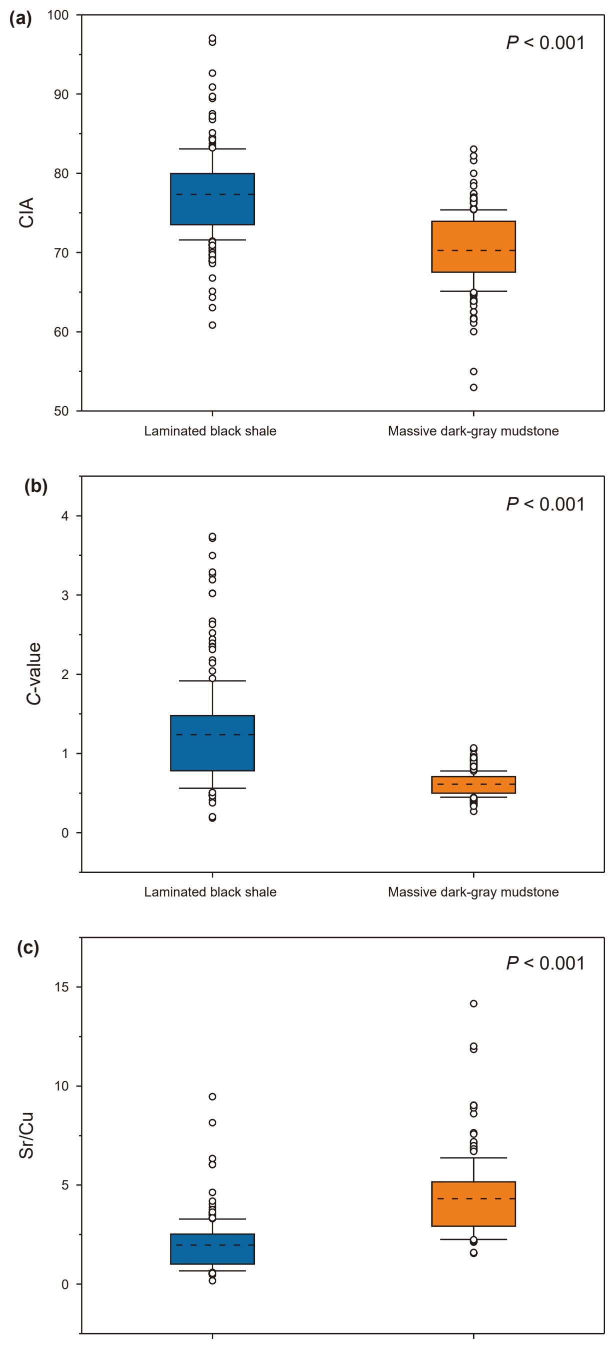

Fig. 8. CIA, *C*-value, and Sr/Cu ratio of the various lithofacies assemblages in the T_3y_7 shale of the Ordos Basin.

particular proxy is not suitable for redox assessment in the study area. Based on the combined interpretations of these redox proxies and previously established criteria, we infer that the laminated black shale was predominantly deposited under dysoxic to anoxic conditions, while the massive dark-gray mudstone mainly accumulated under oxic to dysoxic conditions.

Given the differential solubilities of $BaSO_4$ and $SrSO_4$ in waters of varying salinity, the Sr/Ba ratio was employed as a proxy for paleo-salinity (Wei and Algeo, 2020). The Sr/Ba ratio of the various lithofacies assemblages in the T_3y_7 shale is shown in Fig. 9(h). The associated *P*-value is 0.044, which is below the 0.05 threshold, indicating statistical significance. However, the results suggest no substantial salinity differences between the laminated black shale and the massive dark-gray mudstone. According to the Sr/Ba ratio, both lithofacies assemblages were deposited in freshwater to brackish environments. Volcanic and hydrothermal activity can introduce significant quantities of metal complexes into the water column through volcanic ash and hydrothermal fluids, thereby potentially increasing water salinity (Censi et al., 2010; Ding et al., 2019; Yang et al., 2019). Previous studies have proposed that such volcanic and hydrothermal processes contributed positively to OM enrichment in the T_3y_7 shale (Wang et al., 2017; Zhang et al., 2020; Liu et al., 2022; Li et al., 2024). Based on this hypothesis, it could be expected that the laminated black shale would have formed in a more saline environment. However, our results indicate that there was no significant salinity difference between the laminated black shale and the massive dark-gray mudstone during deposition. This apparent contradiction may be attributed to the more humid climate prevailing during the deposition of the laminated black shale, which could have diluted the salinity of the lake water despite the influence of hydrothermal or volcanic input.

5.1.3. Primary productivity and terrigenous debris influx

Ni, Cu, and Zn can act as organometallic ligands and essential micronutrients (Tribovillard et al., 2006), while elements such as P play a crucial role in promoting plankton proliferation (Algeo and Ingall, 2007; Feng et al., 2023). Consequently, these elements are commonly used as geochemical proxies for primary productivity. In contrast, Al is typically used to indicate terrigenous detrital input, as it is rarely affected by diagenetic processes (Tribovillard et al., 2006). Based on these considerations, this study employs P/Al, Ba/Al, Cu/Al, Ni/Al, and Zn/Al ratios to evaluate primary productivity in the water column (Tribovillard et al., 2006; Algeo and Ingall, 2007). The distribution of these primary productivity-related proxies is shown in Fig. 10. Due to the low OOB score (<0.5) for the interpolated Ba/Al data, Ba/Al ratio was not considered a reliable proxy in this study. The remaining four proxies all yield *P*-values below 0.001, indicating statistically significant differences. Overall, the laminated black shale exhibits significantly higher values of productivity proxies compared to the massive dark-gray mudstone, implying enhanced planktonic productivity during the deposition of the laminated black shale. This interpretation is further supported by previous findings, which show that the kerogen type in laminated black shale is more favorable for hydrocarbon generation than that in massive dark-gray mudstone (Zhao et al., 2020), lending additional credence to our conclusion.

As discussed earlier, Al and Ti contents, along with the Ti/Al ratio, are commonly used to infer terrigenous input and sedimentation rates. The distributions of these three proxies in different lithofacies assemblages of the T_3y_7 shale are presented in Fig. 11. All three proxies have *P*-values less than 0.01, indicating strong statistical significance. The data reveal that both Al and Ti concentrations are higher in the massive dark–gray mudstone than in the laminated black shale, suggesting a higher influx of terrigenous material and a faster sedimentation rate in the former. This conclusion is also corroborated by the sedimentary textures observed in different lithofacies assemblages. Compared to massive textures, laminated textures generally indicate weaker hydrodynamic conditions during deposition. Therefore, the dominance of laminated textures in laminated black shale supports the inference of lower terrigenous input and reduced sedimentation rates relative to the massive dark-gray mudstone.

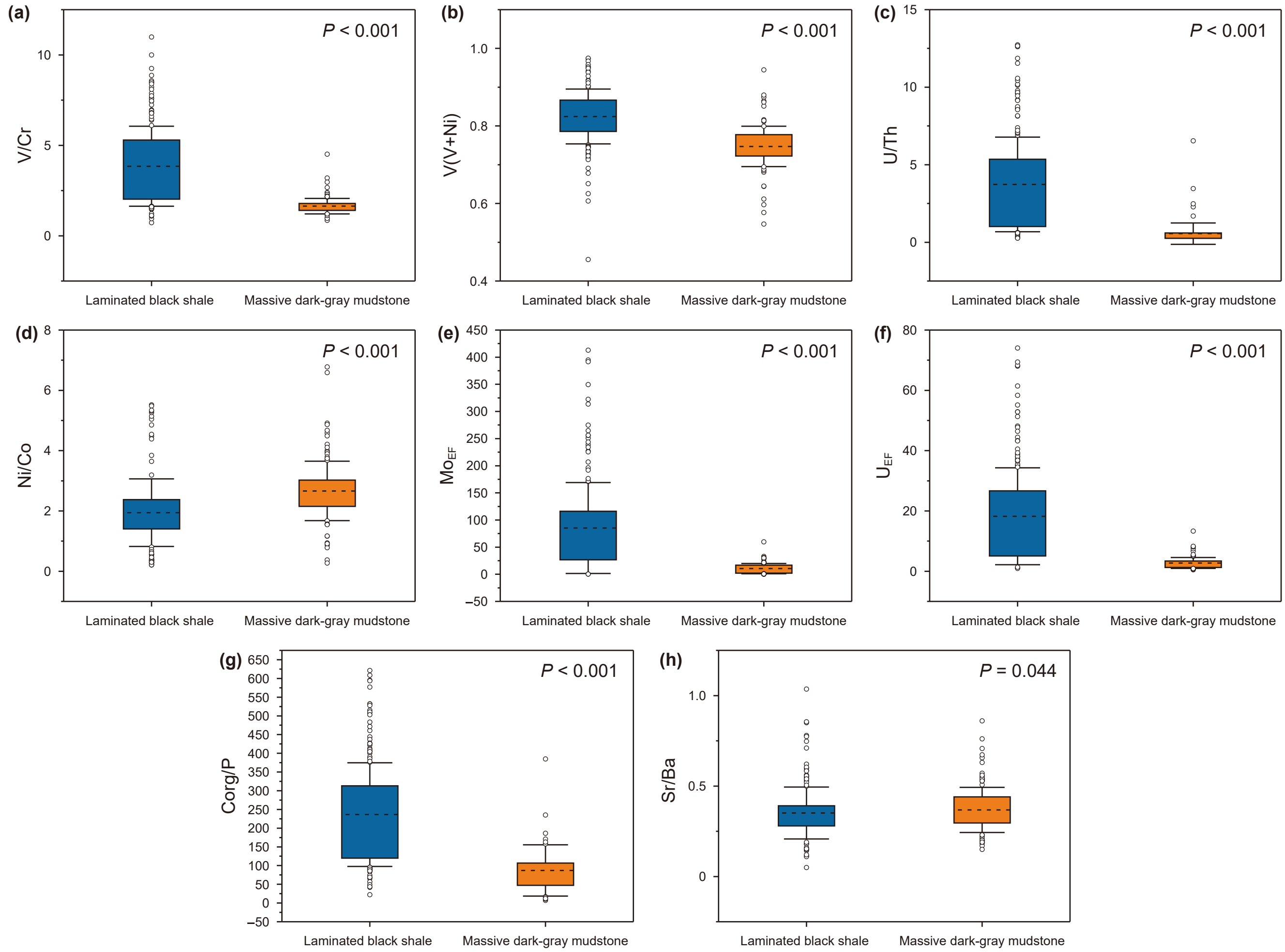

Fig. 9. Redox condition and salinity proxies of various lithofacies assemblages in the T_3y_7 shale of the Ordos Basin.

5.1.4. Volcanic and hydrothermal activities

Volcanic and hydrothermal activities can be recorded and reflected by the elemental characteristics of sedimentary deposits (Adachi et al., 1986; Hulya, 2016). For instance, volcanic activity can lead to variations in $\sum$REE, $(La/Yb)_N$, δCe_N, and δEu_N (Bau and Dulski, 1996; Ballard et al., 2002), while hydrothermal activity is commonly assessed using proxies such as Al/(Al + Fe + Mn), (Fe + Mn)/Ti, and δCe^* (Boström and Peterson, 1969; Adachi et al., 1986; Bau et al., 2014; He et al., 2016; Zhang et al., 2019). The distributions of these proxies in various lithofacies assemblages of the T_3y_7 shale in the Ordos Basin are shown in Figs. 12 and 13.

Among the four proxies indicating volcanic activity, two ($(La/Yb)_N$ and δEu_N) exhibit notably high *P*-values, suggesting a lack of statistical significance. The remaining two proxies, $\sum$REE and δCe_N, exhibit *P*-values below 0.05, indicating statistically significant differences. Based on these proxies, both lithofacies assemblages experienced some degree of volcanic influence during deposition; however, the intensity of volcanic activity appears to have been slightly weaker during the formation of the laminated black shale compared to the massive dark-gray mudstone.

In contrast, all three proxies of hydrothermal activity yield *P*-values below 0.05, confirming statistical significance. These proxies consistently suggest that the laminated black shale was more strongly influenced by hydrothermal activity than the massive dark-gray mudstone. Interestingly, our results present a somewhat counterintuitive finding. Volcanic and hydrothermal activities are generally considered closely associated (Zhang et al., 2020); however, our dataset implies a decoupling between them. Additionally, our results suggest that volcanic activity may have exerted a negative impact on OM accumulation, a finding that appears inconsistent with previous studies, which have emphasized the positive role of volcanic and hydrothermal processes in OM enrichment (Wang et al., 2017; Liu et al., 2021, 2022). This anomalous observation will be further examined and discussed in detail in Section 5.3.

5.2. OM accumulation mechanisms and sedimentary models of various lithofacies assemblages

In the previous sections, a comprehensive analysis of paleoenvironmental proxies across various lithofacies assemblages of the T_3y_7 shale in the Ordos Basin was conducted. Herein, for overcoming the drawbacks of traditional methods such as linear regression in revealing the organic matter enrichment mechanisms, we integrate these proxies with interpretable machine learning models to quantitatively unravel the mechanisms

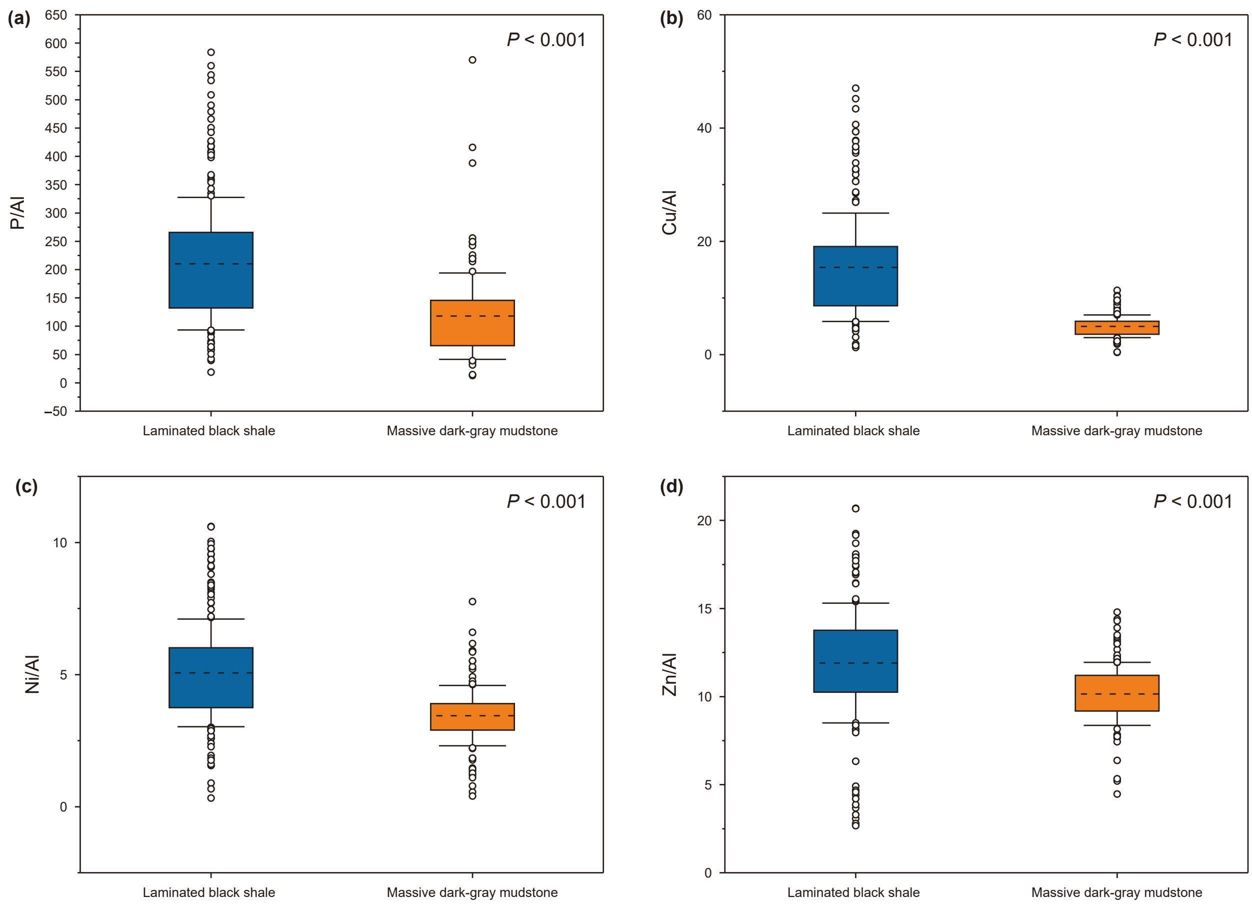

Fig. 10. Primary productivity proxies of various lithofacies assemblages in the T_3y_7 shale of the Ordos Basin.

underlying OM enrichment in the T_3y_7 shale and to construct sedimentary models that explain the genesis of different lithofacies assemblages.

The SHAP values derived from the interpretable machine learning model are presented in Fig. 6. All proxies are ranked along the y-axis in descending order of SHAP importance. Based on this ranking, we focus on the top 20 proxies to investigate their influence on OM enrichment. The results indicate that the V/Cr ratio, representing redox conditions, and the Cu/Al ratio, indicative of primary productivity, are the two most influential factors. This highlights that redox conditions and primary productivity are the primary controls on OM accumulation in the T_3y_7 shale. Compared to these two dominant factors, terrigenous debris influx, paleoclimate, salinity, volcanic activity, and hydrothermal activity exhibit significantly lower contributions to the model. Additionally, the SHAP summary plots provide insights into how each feature affects the model output. Among the redox proxies, all except the Ni/Co ratio show positive contributions, indicating that favorable preservation conditions under reducing environments play a critical role in OM accumulation. Regarding primary productivity, the Ba/Al ratio was excluded due to its low OOB score after data interpolation. Among the remaining proxies, most (except for Zn/Al ratio) contribute positively to the model, suggesting that elevated primary productivity promotes OM accumulation.

In contrast, Al and Ti contents demonstrate a clearly negative impact on the model output. This suggests that increased influx of terrigenous material, which typically accelerates sedimentation rates, may dilute the effect of primary productivity. While some studies have suggested a positive relationship between terrigenous influx and primary productivity (Wang et al., 2022), such an interpretation is not applicable in this case. The T_3y_7 shale is predominantly composed of Type II and Type I kerogen, reflecting the major contribution of planktonic organism (Zhao et al., 2020; Wang et al., 2025b). In contrast, enhanced terrigenous input typically results in Type III kerogen, as seen in the Jurassic Ziliujing Formation of the Sichuan Basin (Wang et al., 2023a). From the perspective of paleoclimate, the *C*-value shows an ambiguous effect due to its relatively low SHAP importance. However, the CIA exhibits a clearly positive impact, while the Sr/Cu ratio has a negative influence. Both ranks relatively high among the proxies (9th and 11th, respectively), suggesting that a humid climate and intense chemical weathering are conducive to OM enrichment.

The Sr/Ba ratio exhibits a negative influence on the model, implying a detrimental effect of salinity on TOC content. This finding contrasts with many previous studies on lacustrine shales, which have argued that increased salinity under freshwater conditions can promote water column stratification due to density differences, thereby facilitating the development of reducing conditions at the bottom waters (Tang et al., 2020; Jiang et al., 2022). However, since the T_3y_7 shale was mainly deposited under freshwater to brackish conditions, variations in salinity were unlikely to induce significant stratification. Additionally, under humid climatic conditions, the paleo-water depth of the depositional system likely increased, which would have further strengthen reducing environment of bottom water. This

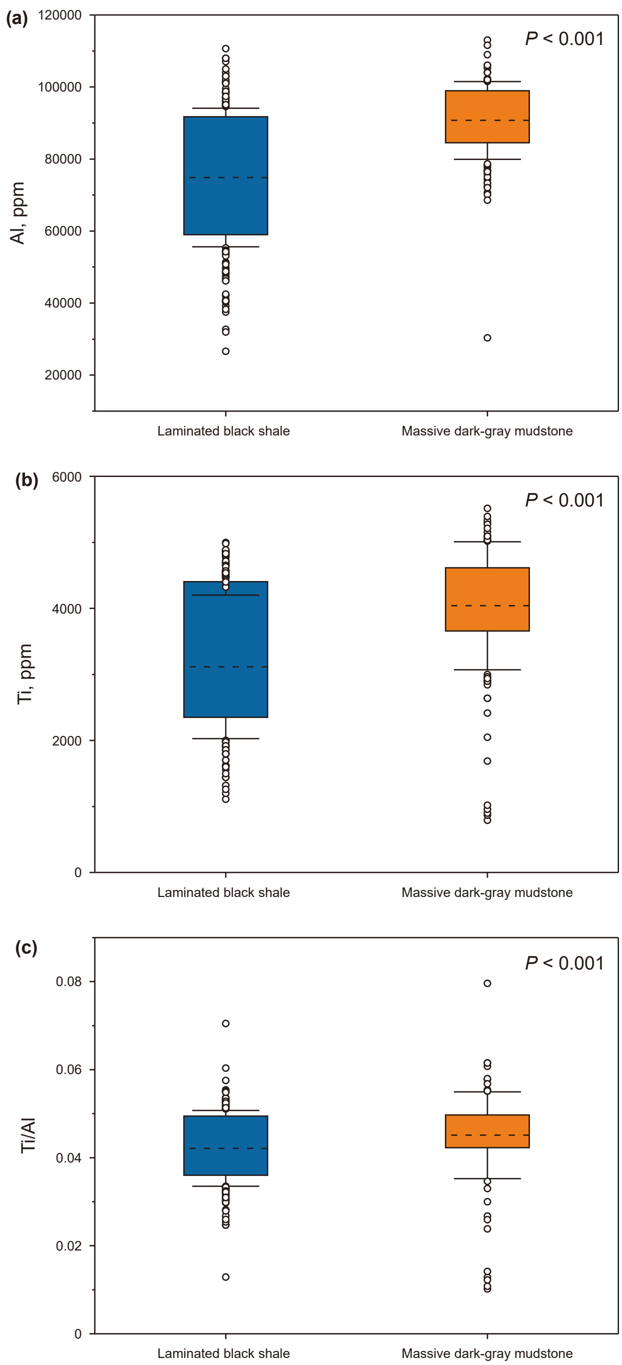

Fig. 11. Terrigenous debris influx proxies of various lithofacies assemblages in the T_3y_7 shale of the Ordos Basin.

interpretation is supported by the effects of CIA and Sr/Cu ratio observed in the model. Therefore, salinity did not exert a positive influence on TOC content in this study. Previous studies have suggested that volcanic activity positively affected OM accumulation in the T_3y_7 shale (Wang et al., 2017; Liu et al., 2022). Interestingly, our model shows that $\sum$REE exhibits a negative influence, while $(La/Yb)_N$ presents a predominantly positive effect, both indicating that volcanic activity may have had a negative impact on OM enrichment. In contrast, δEu_N shows ambiguous trends and ranks low in SHAP importance. Taken together, these proxies suggest that volcanic activity was not favorable for OM accumulation, a conclusion that diverges from some earlier studies (Wang et al., 2017; Liu et al., 2022) but aligns with the findings of Zhang et al. (2020). The discrepancies among these studies will be further explored in Section 5.3. Finally, among the proxies for hydrothermal activity, only the (Fe + Mn)/Ti ratio falls within the range of relevant SHAP importance. This ratio exerts a positive contribution to the model, suggesting that hydrothermal activity may have played a facilitating role in OM accumulation. The contrasting roles of hydrothermal and volcanic activities in OM enrichment will also be discussed in detail in Section 5.3.

In summary, by leveraging an interpretable machine learning model, we quantitatively assessed the contributions of various paleoenvironmental proxies to OM enrichment. This approach overcomes the limitations of single-factor analysis by simultaneously considering all proxies as independent variables, thereby yielding more objective and comprehensive results.

Based on the variations in paleoenvironmental proxies among the various lithofacies assemblages, we reconstructed conceptual models to illustrate their depositional processes (Fig. 14). By integrating the controlling factors of OM enrichment, we further analyzed the formation mechanisms of different lithofacies assemblages. The laminated black shale was deposited in a humid climate characterized by moderate to strong chemical weathering. Such humid conditions likely enhanced paleo-water depth. Although these facies experienced some influence from volcanic activity, the overall intensity of volcanism was relatively weak. In contrast to the massive dark-gray mudstone, the laminated black shale was associated with stronger hydrothermal activity. The combination of relatively intense hydrothermal input and increased paleo-water depth facilitated the establishment of reducing conditions in the bottom waters of the lake (Zhang et al., 2010). Moreover, hydrothermal fluids supplied abundant nutrient elements, which promoted plankton proliferation and enhanced the primary productivity of the water column (Korzhinsky et al., 1994; McKibben et al., 1990; Dover, 2001). It is worth noting that, due to the humid climate, hydrothermal activity did not significantly affect the salinity of the water column. Compared to the massive dark-gray mudstone, the laminated black shale also experienced lower terrigenous debris influx, which favored the development of its laminated texture. Collectively, these conditions created an environment highly favorable for OM enrichment and ultimately led to the formation of laminated black shale.

The massive dark-gray mudstone, on the other hand, was deposited under semi-humid conditions with moderate weathering intensity. The relatively arid paleoclimate likely resulted in reduced paleo-water depth, making it difficult to develop a persistent reducing environment at the lake bottom. As a result, the depositional setting was predominantly oxic to dysoxic. Compared to the laminated black shale, the influence of hydrothermal fluids was weaker, while volcanic activity was more intense. However, elevated volcanic activity did not lead to enhanced primary productivity in the water column. During the deposition of the massive dark-gray mudstone, both terrigenous debris influx and sedimentation rates were higher. The strong influx of clastic material likely diluted the effects of primary productivity, suppressed OM accumulation, and contributed to the development of a massive texture.

5.3. Reinterpretation of the influences of volcanic and hydrothermal activities on OM accumulation

As discussed earlier, volcanic activity appears to have been more intense during the deposition of the massive dark-gray

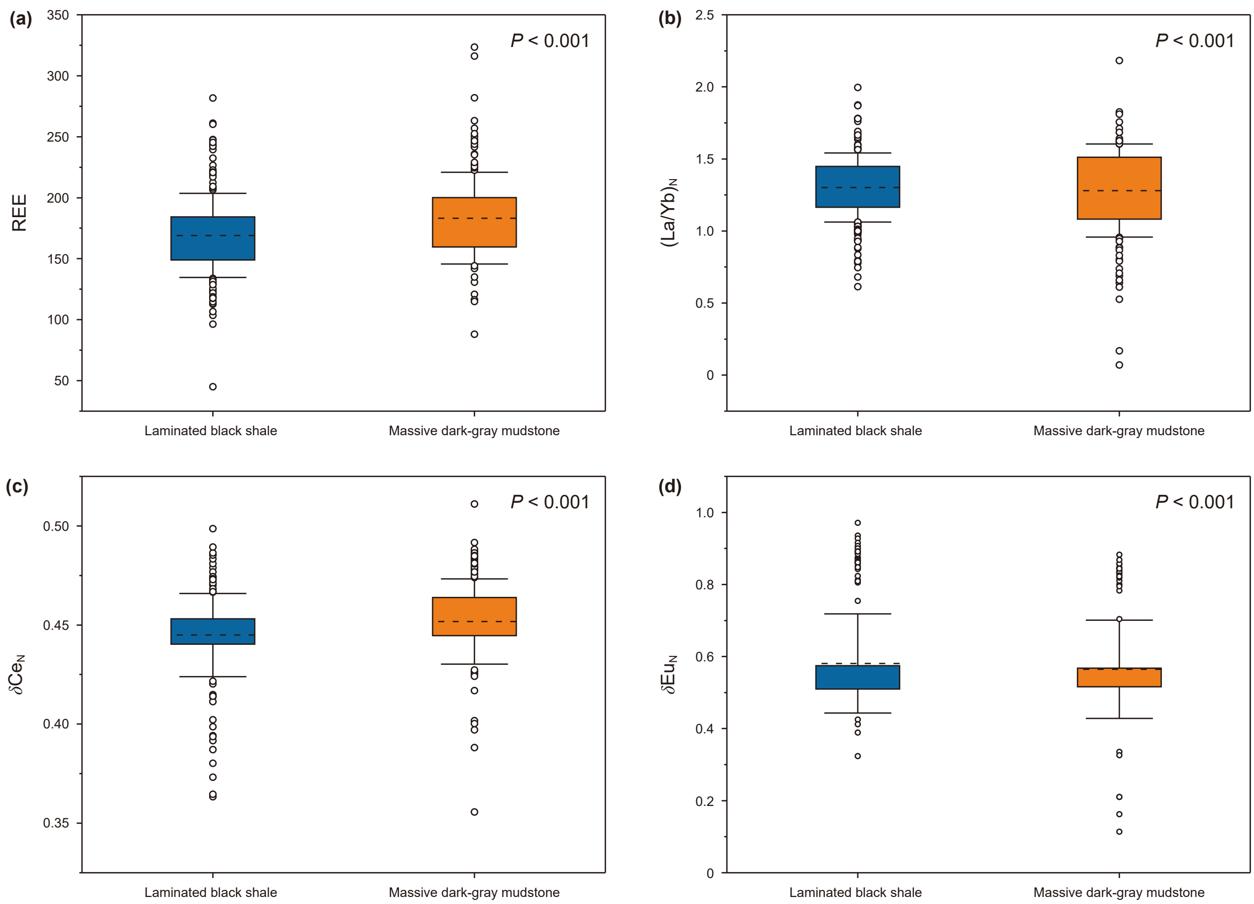

Fig. 12. Volcanic activity proxies of various lithofacies assemblages in the T_3y_7 shale of the Ordos Basin.

mudstone, whereas the laminated black shale was more notably affected by hydrothermal alteration. This observation contrasts with previous studies that emphasized a positive influence of volcanic activity on OM enrichment in the Yanchang Formation (Wang et al., 2017; Liu et al., 2021, 2022). In contrast, Zhang et al. (2020) proposed that volcanic eruptions could induce climatic shifts, toward drier and cooler conditions, which are unfavorable for OM accumulation. The contrasting paleoclimatic conditions reconstructed for the laminated black shale and the massive dark-gray mudstone in this study support the conclusions of Zhang et al. (2020), suggesting that intensified volcanic activity can indeed lead to climatic deterioration. However, we argue that the influence of volcanic activity on the T_3y_7 shale is multifaceted. First, we acknowledge that volcanic activity may positively contribute to OM enrichment in certain contexts. Previous studies have documented the presence of multiple tuff and volcanic ash interlayers within the T_3y_7 (Zhang et al., 2020; Liu et al., 2021), and this interpretation is also supported by geochemical proxies in our dataset. These tuff and ash layers are largely the result of volcanic eruptions that released fine ash into the depositional system. Volcanic ash can introduce abundant nutrients that promote plankton proliferation and, consequently, increase primary productivity in the water column. However, it is important to note that such productivity enhancement may occur with a hysteresis period. In addition, volcanic ash can improve water column reducibility by altering redox conditions (Liu et al., 2022). From this perspective, the positive role of volcanic activity in promoting OM enrichment, as highlighted in previous studies, is reasonable. Nevertheless, as Zhang et al. (2020) emphasized, the intensity of volcanic activity must be considered when evaluating its impact on OM accumulation. Beyond its climatic implications, volcanic activity also plays a role in driving tectonic movements. Specifically, intense eruptions are often accompanied by significant tectonic reorganization, which can alter the tectonic framework of the affected basin. Moreover, because volcanic ash can be transported across vast distances through the atmosphere, even eruptions occurring far from the depositional center may influence OM accumulation in lacustrine environments.

Integrating these insights with our study results, we propose a conceptual model to elucidate the dual impact of volcanic activity on lacustrine shale systems. When volcanic activity is weak or the lake is located far from the eruption center, the basin is primarily influenced by airborne volcanic ash, while tectonic activity remains relatively stable. Under these conditions, the influx of volcanic ash can enhance primary productivity and promote a reducing environment in the water column, thereby favoring OM enrichment. However, when volcanic activity intensifies or the lake lies in close proximity to the eruption center, both the paleoclimate and tectonic pattern may be significantly disturbed. For instance, elevated terrigenous input associated with stronger tectonic activity can dilute productivity signals, while aridification may reduce paleo-water depth, impeding the development of reducing environment of bottom water. These combined effects may offset the positive contributions of volcanic ash, ultimately creating conditions less favorable for OM accumulation.

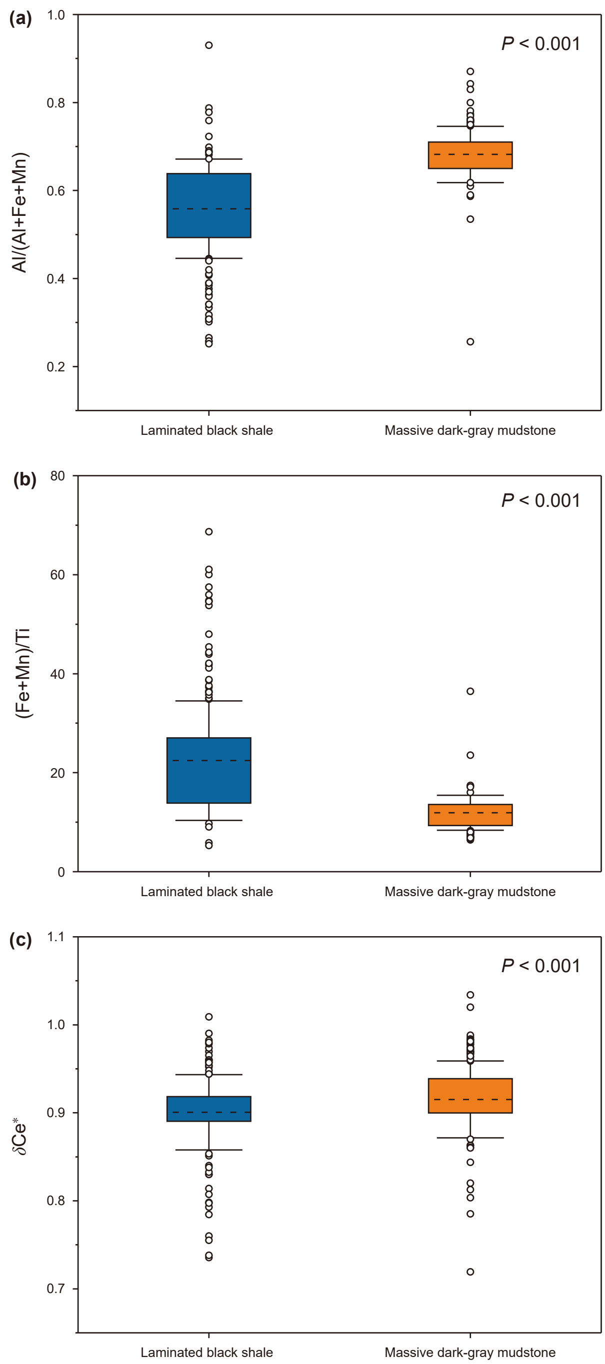

Fig. 13. Hydrothermal activity proxies of various lithofacies assemblages in the T_3y_7 shale of the Ordos Basin.

Taking the T_3y_7 shale as an example, volcanic activity exerted a discernible influence during the deposition of the laminated black shale; however, its intensity was significantly lower than that associated with the massive dark-gray mudstone (Fig. 12). As a result, the laminated facies are characterized by finely laminated textures and elevated TOC content. The high abundance of pyrite further suggests additional sulfate (SO_4^{2-}) input, likely sourced from volcanic activity, in agreement with previous findings (Liu et al., 2021). In contrast, during the deposition of the massive dark-gray mudstone, geochemical proxies indicative of volcanic activity reflects either a higher eruption intensity or a shorter distance from the volcanic center. Consequently, this interval is associated with increased terrigenous debris influx, a drier paleoclimate, and massive sedimentary textures. These conditions collectively led to substantially lower TOC contents in the massive facies. Based on the above discussion, a key conclusion can be drawn: the influence of volcanic activity on OM enrichment in lacustrine shale systems is inherently complex. This influence appears to be governed by both the intensity of volcanic activity and the proximity of the basin to the eruption center. Moderate volcanic activity may enhance OM accumulation by delivering nutrients and promoting reducing conditions, whereas intense volcanic events tend to exert a detrimental effect. SHAP interaction values further reveal that increases in proxies representing volcanic activity tend to suppress the positive contributions of redox conditions and primary productivity to OM enrichment (Fig. 15(a) and (c)). Specifically, for samples with relatively intense volcanic activity (characterized by red dot in the figure), paleoproductivity and redox factors show a negative correlation with their corresponding SHAP interaction values, respectively. This observation provides additional support for our conclusion. However, it should be noted that in addition to the intensity of volcanic activity, the type of volcanism must also be considered. For example, a study of the Cretaceous Shahezi Formation in the Songliao Basin found distinct differences in how subaqueous and subaerial volcanic activities influence organic matter enrichment (Xie et al., 2024).

In addition to volcanic activity, hydrothermal fluids can also significantly influence the redox conditions and primary productivity within the water column (Korzhinsky et al., 1994; McKibben et al., 1990; Dover, 2001). Although volcanic and hydrothermal activities are often closely related, hydrothermal fluids can still reach the lake environment through fault systems or other pathways, even in the absence of substantial volcanic activity. This process does not necessarily require intense tectonic disturbances. As shown in the spatial distribution of geochemical proxies (Fig. 13), hydrothermal activity was more pronounced during the deposition of the laminated black shale. Therefore, in addition to moderate volcanic activity that contributed volcanic ash, hydrothermal fluids likely delivered abundant nutrients to the lake, enhancing both primary productivity and the development of reducing conditions.

In contrast, the massive dark-gray mudstone exhibits weaker hydrothermal activity, despite stronger volcanic signals during its deposition. This discrepancy suggests that the volcanic activity during this period may not have directly affected the lake system via magmatic input. Consistently, SHAP interaction values indicate that an increase in the proxy representing hydrothermal activity enhances the positive influence of redox and productivity-related proxies on OM enrichment (Fig. 15(b) and (d)). For samples with relatively intense hydrothermal activity (indicated by red dots in Fig. 15), primary productivity and redox factors show a positive correlation with their corresponding SHAP interaction values, respectively. This finding further supports the interpretation that OM enrichment was more favorable in the laminated black shale facies.

In summary, volcanic activity of varying intensities and types exerts differentiated effects on OM content in lacustrine shale systems. Meanwhile, the role of hydrothermal activity in OM accumulation should not be overlooked. Future studies should consider volcanic and hydrothermal activities as distinct factors in order to more objectively elucidate their respective impacts on the development of organic-rich shale in lacustrine systems.

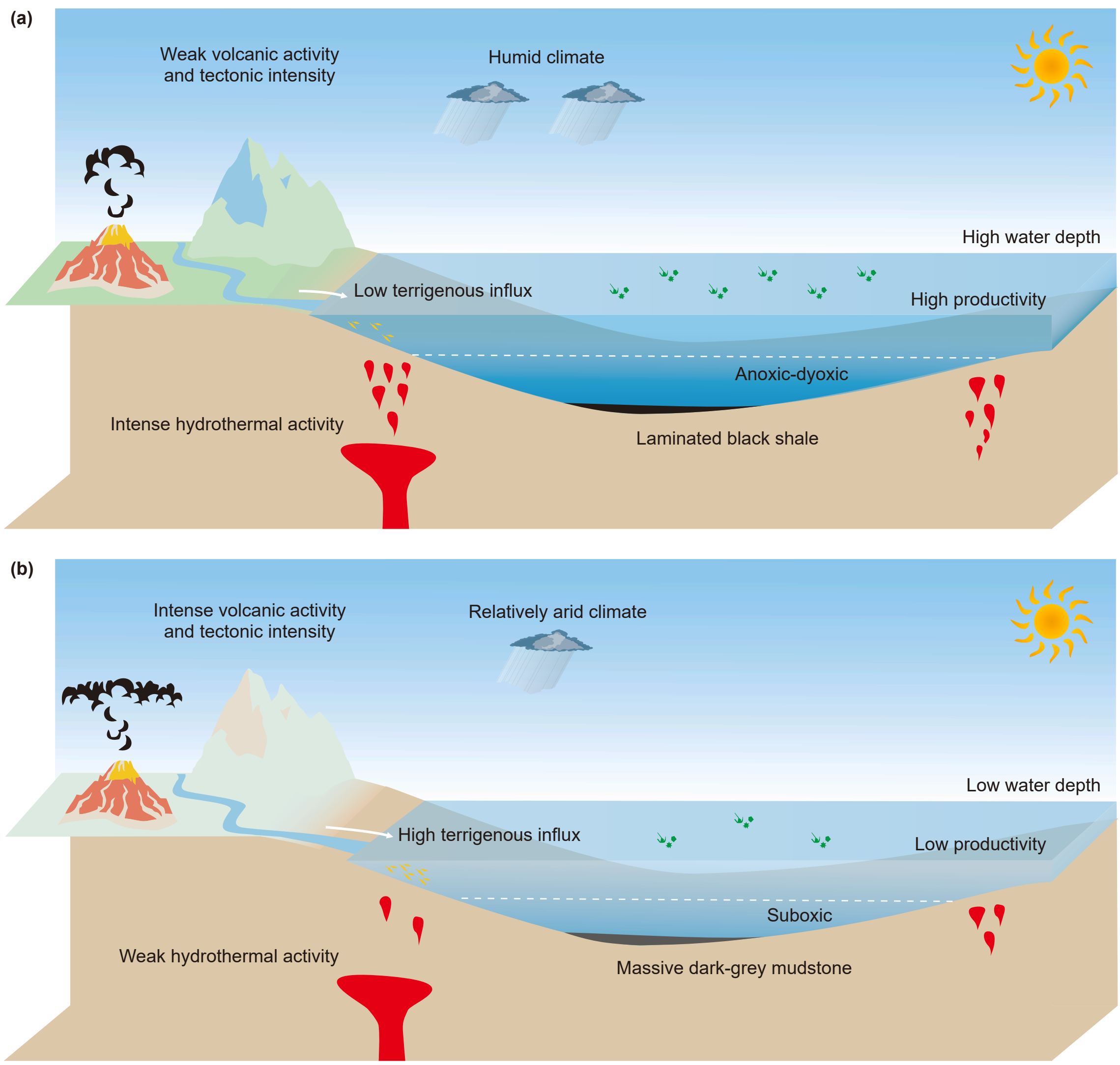

Fig. 14. Conceptual diagram of organic matter enrichment in various lithofacies assemblages in the T_3y_7 shale of the Ordos Basin.

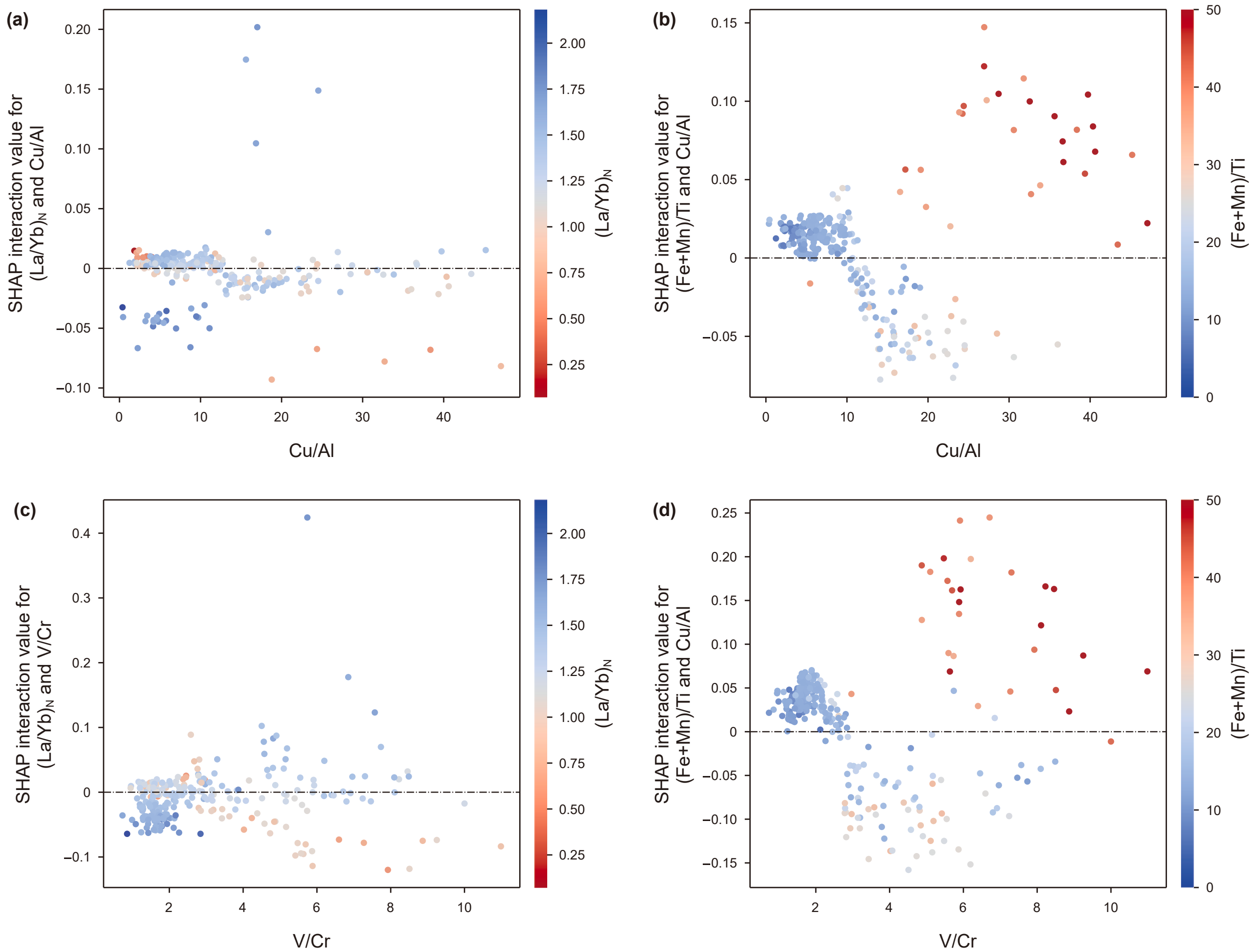

Fig. 15. Differential interaction effects for volcanic and hydrothermal proxies on primary productivity and redox conditions.

6. Conclusions

Based on newly obtained geochemical data from YYa borehole in the southern Ordos Basin and previously published datasets, this study reconstructs the paleoenvironmental conditions associated with various lithofacies assemblages within the T_3y_7. An interpretable machine learning model was developed to identify the dominant controls on OM accumulation, and the distinct influences of volcanic and hydrothermal activity on OM enrichment were systematically examined using the T_3y_7 shale as a case study. The major conclusions are as follows:

(1) Two lithofacies assemblages are identified in the T_3y_7 shale: laminated black shale and massive dark-gray mudstone. These assemblages were deposited under markedly different conditions, including variations in paleoclimate, salinity, redox conditions, primary productivity, terrigenous debris influx, and both volcanic and hydrothermal activities. These environmental differences led to distinct mineral compositions, sedimentary textures, and organic geochemical characteristics in each lithofacies assemblage.

(2) The interpretable machine learning model enables quantitative assessment of the influence of multiple paleoenvironmental proxies on OM enrichment. This approach overcomes the limitations of single-factor analyses by considering all proxies simultaneously as input variables, resulting in more objective and robust conclusions. The model suggests that OM accumulation in the T_3y_7 shale is primarily controlled by the coupling of high primary productivity and effective preservation conditions. Additionally, a humid paleoclimate, moderate volcanic activity, active hydrothermal input, and relatively low terrigenous influx and sedimentation rates are all favorable for OM enrichment.

(3) The influence of volcanic activity on OM accumulation in lacustrine shale systems is non-linear. The decoupling of volcanic and hydrothermal activities significantly influenced the formation of the distinct lithofacies within the T_3y_7. While moderate volcanic activity promotes OM enrichment, intense volcanic activity tends to have a detrimental effect. Moreover, hydrothermal activity plays a significant and independent role in facilitating OM accumulation and should be considered separately from volcanic activity in future research. This distinction is crucial for accurately understanding the mechanisms governing organic matter enrichment in lacustrine shale systems.

CRediT authorship contribution statement

Enze Wang: Writing – original draft, Investigation, Conceptualization. **Xiaoxiao Ma:** Writing – review & editing, Resources, Data curation. **Maowen Li:** Writing – review & editing, Conceptualization. **Yingxiao Fu:** Writing – original draft, Visualization, Methodology. **Menhui Qian:** Resources, Data curation. **Tingting Cao:** Resources, Data curation. **Zhiming Li:** Resources, Data curation. **Yue Feng:** Writing – original draft, Visualization. **Zhijun Jin:** Supervision, Project administration. **Tong Li:** Writing – review & editing, Investigation.

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.

Acknowledgments

This study is supported by the National Natural Science Foundation of China (Projects 42090022, U24B6002, and 42502151).

References

Adachi, M., Yamamoto, K., Sugisaki, R., 1986. Hydrothermal chert and associated siliceous rocks from the northern Pacific their geological significance as indication of ocean ridge activity. Sediment. Geol. 47, 125–148. https://doi.org/10.1016/0037-0738(86)90075-8.
Algeo, T.J., Ingall, E., 2007. Sedimentary corg: P ratios, paleocean ventilation, and Phanerozoic atmospheric pO_2. Palaeogeogr. Palaeoclimatol. Palaeoecol. 256 (3), 130–155. https://doi.org/10.1016/j.palaeo.2007.02.029.
Algeo, T.J., Li, C., 2020. Redox classification and calibration of redox thresholds in sedimentary systems. Geochem. Cosmochim. Acta 287, 8–26. https://doi.org/10.1016/j.gca.2020.03.003.
Algeo, T.J., Liu, J.S., 2020. A re-assessment of elemental proxies for paleoredox analysis. Chem. Geol. 540, 119549. https://doi.org/10.1016/j.chemgeo.2020.119549.
Ballard, J.R., Palin, M.J., Campbell, I.H., 2002. Relative oxidation states of magmas inferred from Ce (IV)/Ce(III) in zircon: Application to porphyry copper deposits of northern Chile. Contrib. Mineral. Petrol. 144 (3), 347–364. https://doi.org/10.1007/s00410-002-0402-5.
Bau, M., Dulski, P., 1996. Distribution of yttrium and rare-earth elements in the penge and kuruman iron-formations, transvaal supergroup, South Africa. Precambr. Res. 79 (1–2), 37–55. https://doi.org/10.1016/0301-9268(95)00087-9.
Bau, M., Schmidt, K., Koschinsky, A., Hein, J., Kuhn, T., Usui, A., 2014. Discriminating between different genetic types of marine ferro-manganese crusts and nodules based on rare earth elements and yttrium. Chem. Geol. 381, 1–9. https://doi.org/10.1016/j.chemgeo.2014.05.004.
Bilal, Pant, M., Zaheer, H., Garcia-Hernandez, L., Abraham, A., 2020. Differential Evolution: a review of more than two decades of research. Eng. Appl. Artif. Intell. 90, 103479.
Boström, K., Peterson, M.N.A., 1969. The Origin of aluminum-poor Ferromanganoan sediments in areas of high heat flow on the east Pacific Rise, 7, pp. 427–447. https://doi.org/10.1016/0025-3227(69)90016-4.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.
Burnaz, L., Kalmkov, A., Grohmann, S., Kalmykov, G., Littke, R., 2022. Geochemistry and organic petrology of organic-rich shales of the Upper Jurassic–Lower Cretaceous Bazhenov Horizon in the Frolov Region, West Siberian Basin: Implications for the reconstruction of the organic matter origin and paleoredox conditions. Mar. Petrol. Geol. 143, 105809. https://doi.org/10.1016/j.marpetgeo.2022.105809.
Censi, P., Randazzo, L.A., Zuddas, P., et al., 2010. Trace element behaviour in seawater during Etna's pyroclastic activity in 2001: Concurrent effects of nutrients and formation of alteration minerals. J. Volcanol. Geoth. Res. 193, 106–116. https://doi.org/10.1016/j.jvolgeores.2010.03.010.
Chen, T., Guestrin, C., 2016. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Presented at the KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, San Francisco California USA, pp. 785–794. https://doi.org/10.1145/2939672.2939785.
Chen, Z.H., Guo, Q.L., Jiang, C.Q., Liu, X.J., Reyes, J., Mort, A., Jia, Z.K., 2017. Source rock characteristics and Rock-Eval-based hydrocarbon generation kinetic models of the lacustrine Chang-7 Shale of Triassic Yanchang Formation, Ordos Basin. China. Int. J. Coal Geol. 182, 52–65. https://doi.org/10.1016/j.coal.2017.08.017.
Chen, M., Liu, Q., Chen, S., Liu, Y., Zhang, C.H., Liu, R., 2019. XGBoost-based algorithm interpretation and application on post-fault transient stability status prediction of power system. IEEE Access 7, 13149–13158. https://doi.org/10.1109/ACCESS.2019.2891516.
Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20, 273–297. https://doi.org/10.1007/BF00994018.
Dapples, E.C., Krumbein, W.C., Sloss, L.L., 1948. Tectonic control of lithofacies associations. AAPG Bull. 32, 1924–1947.
Deng, S., Lu, Y.Z., Luo, Z., Fan, R., Li, X., Zhao, Y., Ma, X.Y., Zhu, R.K., Cui, J.W., 2018. Subdivision and age of the Yanchang Formation and the Middle/Upper Triassic boundary in Ordos Basin, North China. Sci. China Earth Sci. 61 (10), 1419–1439. https://doi.org/10.1007/s11430-017-9215-3.
Ding, X., Qu, J., Iminc, A., Zha, M., Su, Y., Jiang, Z., Jiang, H., 2019. Organic matter origin and accumulation in tuffaceous shale of the lower Permian Lucaogou Formation, Jimsar Sag. J. Pet. Sci. Eng. 179, 696–706. https://doi.org/10.1016/j.petrol.2019.05.004.
Dover, V.C.L., 2001. Biogeography and ecological setting of Indian ocean hydrothermal vents. Science 294, 818–823. https://doi.org/10.1126/science.1064574.
Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A., Vapnik, V., 1996. Support vector regression machines. In: Proceedings of the 9th international conference on neural information processing systems. MIT Press, Cambridge, MA, USA, pp. 155–161. NIPS'96.
Duggen, S., Croot, P., Schacht, U., Hoffmann, L., 2007. Subduction zone volcanic ash can fertilize the surface ocean and stimulate phytoplankton growth: Evidence from biogeochemical experiments and satellite data. Geophys. Res. Lett. 34, 1–5. https://doi.org/10.1029/2006GL027522.
EIA, 2024. How much shale (tight) oil is produced in the united states? https://www.eia.gov/tools/faqs/faq.php?id=847&t=6.
Fan, G.F., Zhang, L.Z., Yu, M., Hong, W.C., Dong, S.Q., 2022. Applications of random forest in multivariable response surface for short-term load forecasting. Int. J. Electr. Power Energy Syst. 139, 108073.
Feng, Y., Xiao, X.M., Gao, P., Wang, E.Z., Hu, D.F., Liu, R.B., Li, G., Lu, C.G., 2023. Restoration of sedimentary environment and geochemical features of deep marine Longmaxi shale and its significance for shale gas: A case study of the Dingshan area in the Sichuan Basin, South China. Mar. Petrol. Geol. 151, 106186. https://doi.org/10.1016/j.marpetgeo.2023.106186.
Fu, J.H., Li, S.X., Xu, L.M., et al., 2018. Paleo-sedimentary environmental restoration and its significance of Chang 7 Member of Triassic Yanchang Formation in Ordos Basin, NW China. Petrol. Explor. Dev. 45 (6), 998–1008. https://doi.org/10.1016/S1876-3804(18)30104-6.
Gjelsvik, E.L., Fossen, M., Tøndel, K., 2023. Current overview and way forward for the use of machine learning in the field of petroleum gas hydrates. Fuel 334, 126696. https://doi.org/10.1016/j.fuel.2022.126696.
Guan, M., Wu, S., Hou, L., Jiang, X., Ba, D., Hua, G., 2021. Paleoenvironment and chemostratigraphy heterogenity of the Cretaceous organic-rich shales. Adv. Geo-Energy Res. 5 (4), 444–455. https://doi.org/10.46690/ager.2021.04.09.
Guo, X.S., Wang, E.Z., Ma, X.X., Li, M.W., Qian, M.H., Cao, T.T., Li, Z.M., Chu, J.J., 2025. Lacustrine shale oil systems in China: Advances in characterization methods and resource accumulation models. Earth Sci. Rev. 270, 105256. https://doi.org/10.1016/j.earscirev.2024.105256.
He, C., Ji, L., Wu, Y., Su, A., Zhang, M., 2016. Characteristics of hydrothermal sedimentation process in the Yanchang Formation, south Ordos Basin, China: Evidence from element geochemistry. Sediment. Geol. 345, 33–41. https://doi.org/10.1016/j.sedgeo.2016.09.001.
Hulya, K., 2016. Major, trace and rare earth element (REE) characteristics of tuffs in the Yenice-saraycık area (Demirci, Manisa), Western Anatolia, Turkey. J. Geochem. Explor. 168, 169–176. https://doi.org/10.1016/j.gexplo.2016.06.014.
James, G., Witten, D., Hastie, T., Tibshirani, R., Taylor, J., 2023. An introduction to statistical learning: With applications in python, springer texts in statistics. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-031-38747-0.
Jiang, F.J., Chen, D., Zhu, C.X., Ning, K.C., Ma, L., Xu, T.W., Qin, R., Li, B.S., Chen, Y.Y., Huo, L.N., Xu, Z., 2022. Mechanisms for the anisotropic enrichment of organic matter in saline lake basin: A case study of the Early Eocene Dongpu Depression, eastern China. J. Petrol. Sci. Eng. 210, 110035. https://doi.org/10.1016/j.petrol.2021.110035.
Jones, M.T., Gislason, S.R., 2008. Rapid releases of metal salts and nutrients following the deposition of volcanic ash into aqueous environments. Geochem. Cosmochim. Acta 72, 3661–3680. https://doi.org/10.1016/j.gca.2008.05.030.
Jones, B., Manning, D.A., 1994. Comparison of geochemical indices used for the interpretation of palaeoredox conditions in ancient mudstones. Chem. Geol. 111 (1–4), 111–129. https://doi.org/10.1016/0009-2541(94)90085-X.
Katz, B., Lin, F., 2014. Lacustrine basin unconventional resource plays: key differences. Mar. Petrol. Geol. 56, 255–265. https://doi.org/10.1016/j.marpetgeo.2014.02.013.
Kim, S., Lee, S., 2023. Nonlinear relationships and interaction effects of an urban environment on crime incidence: Application of urban big data and an interpretable machine learning method. Sustain. Cities Soc. 91, 104419. https://doi.org/10.1016/j.scs.2023.104419.
Korzhinsky, M.A., Tkachenko, S.I., Shmlovich, K.I., 1994. Discovery of a pure-rhenium mineral at Kudriavy volcano. Nature 369, 51–52. https://doi.org/10.1038/369051a0.
Lee, J., Lumley, D.E., Lim, U.Y., 2022. Improving total organic carbon estimation for unconventional shale reservoirs using Shapley value regression and deep machine learning methods. Am. Assoc. Petrol. Geol. Bull. 106 (11), 2297–2314. https://doi.org/10.1306/02072221021.
Li, D., Shi, Q., Mi, N., Xu, Y., Wang, X., Tao, W., 2020. The type, origin and preservation of organic matter of the fine-grain sediments in Triassic Yanhe Profile, Ordos Basin, and their relation to paleoenvironment condition. J. Petrol. Sci. Eng. 188, 106875. https://doi.org/10.1016/j.petrol.2019.106875.

Li, X.C., Hou, Y.G., Chen, Z.H., Duan, Z.N., Zeng, Y., He, S., Chu, D.L., 2024. Volcanism and turbidity current events as drivers of the evolution of benthic redox conditions: organic matter enrichment in the Chang 7 Member, Upper Triassic Yanchang Formation, Ordos Basin, North China. Mar. Petrol. Geol. 170, 107087. https://doi.org/10.1016/j.marpetgeo.2023.107087.

Liao, Z., Lu, J., Xie, K., Wang, Y., Yuan, Y., 2023. Prediction of photochemical properties of dissolved organic matter using machine learning. Environ. Sci. Technol. 57, 17971–17980. https://doi.org/10.1021/acs.est.3c05664.

Liu, Q.Y., Li, P., Jin, Z.J., Liang, X.P., Zhu, D.Y., Wu, X.Q., Meng, Q.Q., Liu, J.Y., Fu, Q., Zhao, J.H., 2021. Preservation of organic matter in shale linked to bacterial sulfate reduction (BSR) and volcanic activity under marine and lacustrine depositional environments. Mar. Petrol. Geol. 127, 104950. https://doi.org/10.1016/j.marpetgeo.2021.104950.

Liu, Q.Y., Li, P., Jin, Z.J., Liang, X.P., Zhu, D.Y., Wu, X.Q., Zhao, J.H., 2022. Effect of volcanic activity on the formation of marine shale and freshwater lacustrine shale. Geochimica 51 (5), 556–569. https://doi.org/10.19700/j.0379-1726.2022.05.005 (in Chinese).

Liu, B., Rostamian, A., Kheirollahi, M., Mirseyed, S.F., Mohammadian, E., Golsanami, N., Liu, K., Ostadhassan, M., 2023. NMR log response prediction from conventional petrophysical logs with XGBoost-PSO framework. Geoenergy Sci. Eng. 224, 211561. https://doi.org/10.1016/j.geoen.2023.211561.

Liu, Q.Y., Li, O., Jiang, L., Jin, Z.J., Liang, X.P., Zhu, D.Y., Pang, Q., Zhang, R., Liu, J.Y., 2024. Distinctive volcanic ash–rich lacustrine shale deposition related to chemical weathering intensity during the Late Triassic: Evidence from lithium contents and isotopes. Sci. Adv. 10, eadi6594. https://doi.org/10.1126/sciadv.adi6594.

Lu, M., Duan, G., Zhang, T., Liu, N., Song, Y., Zhang, Z., Qiao, J., Wang, Z., Fang, Z., Luo, Q., 2025. Influences of paleoclimatic changes on organic matter enrichment mechanisms in freshwater and saline lacustrine oil shales in China: A machine learning approach. Earth Sci. Rev. 262, 105061. https://doi.org/10.1016/j.earscirev.2024.105061.

Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. In: Proceedings of the 31st international conference on neural information processing systems, NIPS'17. Curran Associates Inc., Red Hook, NY, USA, pp. 4768–4777.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I., 2020. From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. 2, 56–67. https://doi.org/10.1038/s42256-019-0138-9.

Lv, D.W., Wang, L.J., Isbell, J.L., Lu, C.Y., Li, P.P., Wang, Y.J., Zhang, Z.H., 2022. Records of chemical weathering and volcanism linked to paleoclimate transition during the Late Paleozoic Icehouse. Global Planet. Change 217, 103934. https://doi.org/10.1016/j.gloplacha.2022.103934.

Ma, Y.S., Cai, X.Y., Zhao, P.R., Hu, Z.Q., Liu, H.M., Gao, B., Wang, W.Q., Li, Z.M., Zhang, Z.L., 2022. Geological characteristics and exploration practices of continental shale oil in China. Acta Geol. Sin. 96 (1), 155–171. https://doi.org/10.19762/j.cnki.dizhixuebao.2022278 (In Chinese).

Madonia, P., Cangemi, M., Bellomo, S., D'Alessandro, W., 2013. Influence of volcanic activity on the quality of water collected in roof water catchment systems at stromboli island (Italy). J. Geochem. Explor. 131, 28–36. https://doi.org/10.1016/j.gexplo.2012.08.018.

Mangalathu, S., Karthikeyan, K., Feng, D.-C., Jeon, J.-S., 2022. Machine-learning interpretability techniques for seismic performance assessment of infrastructure systems. Eng. Struct. 250, 112883. https://doi.org/10.1016/j.engstruct.2021.112883.

McKibben, M.A., Williams, A.E., Hall, G.E.M., 1990. Solubility and transport of platinumgroup elements and Au in saline hydrothermal fluids; constraints from geothermal brine data. Econ. Geol. 85 (8), 1926–1934. https://doi.org/10.2113/gsecongeo.85.8.1926.

Minyuk, P.S., Brigham-Grette, J., Melles, M., Borkhodoev, V.Y., Glushkova, O.Y., 2007. Inorganic geochemistry of El'gygytgyn Lake sediments (northeastern Russia) as an indicator of paleoclimatic change for the last 250 kyr. J. Paleolimnol. 37, 123–133. https://doi.org/10.1007/s10933-006-9027-4.

Moradi, A.V., Sarı, A., Akkaya, P., 2016. Geochemistry of the Miocene oil shale (Hançili Formation) in the Çankırı-Çorum Basin, Central Turkey: Implications for Paleoclimate conditions, source–area weathering, provenance and tectonic setting. Sediment. Geol. 341, 289–303. https://doi.org/10.1016/j.sedgeo.2016.05.002.

Nadege, M.N., Jiang, S., Mwakipunda, G.C., Kouassi, A.K.F., Harold, P.K., Hugues Roland, K.Y., 2024. Brittleness index prediction using modified random forest based on particle swarm optimization of Upper Ordovician Wufeng to Lower Silurian Longmaxi shale gas reservoir in the Weiyuan Shale Gas Field, Sichuan Basin, China. Geoenergy Sci. Eng. 233, 212518. https://doi.org/10.1016/j.geoen.2024.212518.

National Energy Administration of China, 2024. Ten significant achievements in national oil and gas exploration and development of 2023 in China. https://www.nea.gov.cn/2024-01/09/c_1310759352.htm.

Nazar, S., Yang, J., Wang, X.-E., Khan, K., Amin, M.N., Javed, M.F., Althoey, F., Ali, M., 2023. Estimation of strength, rheological parameters, and impact of raw constituents of alkali-activated mortar using machine learning and SHapely Additive exPlanations (SHAP). Constr. Build. Mater. 377, 131014. https://doi.org/10.1016/j.conbuildmat.2022.131014.

Niu, X.B., Lyu, C.F., Feng, S.B., et al., 2025. Lamina combination characteristics and differential shale oil enrichment mechanisms of continental organic-rich shale: A case study of Triassic Yanchang Formation Chang 73 sub-member, Ordos Basin, NW China. Petrol. Explor. Dev. 52 (2), 316–329. https://doi.org/10.1016/S1876-3804(25)60569-6.

O'Dowd, C.D., Facchini, M.C., Cavalli, F., Ceburnis, D., Mircea, M., Decesari, S., Fuzzi, S., Yoon, Y., Putaud, J., 2004. Biogenically driven organic contribution to marine aerosol. Nature 431, 676–680. https://doi.org/10.1038/nature02959.

Otchere, D.A., Ganat, T.O.A., Gholami, R., Lawal, M., 2021. A novel custom ensemble learning model for an improved reservoir permeability and water saturation prediction. J. Nat. Gas Sci. Eng. 91, 103962. https://doi.org/10.1016/j.jngse.2021.103962.

Palansooriya, K.N., Li, J., Dissanayake, P.D., Suvarna, M., Li, L., Yuan, X., Sarkar, B., Tsang, D.C.W., Rinklebe, J., Wang, X., Ok, Y.S., 2022. Prediction of soil heavy metal immobilization by biochar using machine learning. Environ. Sci. Technol. 56, 4187–4198. https://doi.org/10.1021/acs.est.1c08302.

Pan, S., Liao, Y., Jiang, B., Wan, Z., Wang, F., 2022. Impact of natural weathering on source rocks: Organic and inorganic geochemical evidence from the Triassic Chang 7 outcrop profile in Tongchuan of the Southern Ordos Basin (China). Int. J. Coal Geol. 263, 104119. https://doi.org/10.1016/j.coal.2022.104119.

Pedersen, T.F., 1990. Anoxia vs. productivity: What controls the formation of organic carbon rich sediments and sedimentary rocks? Am. Assoc. Petrol. Geol. 74 (4), 454–466. https://doi.org/10.1306/0C9B2397-1710-11D7-8645000102C1865D.

Qiao, J.Q., Baniasad, A., Zieger, L., Zhang, C., Luo, Q., Littke, R., 2021. Paleodepositional environment, origin and characteristics of organic matter of the Triassic Chang 7 Member of the Yanchang Formation throughout the mid-western part of the Ordos Basin, China. Int. J. Coal Geol. 237, 103636. https://doi.org/10.1016/j.coal.2021.103636.

Qiu, X., Liu, C., Mao, G., Deng, Y., Wang, F., Wang, J., 2015. Major, trace and platinumgroup element geochemistry of the Upper Triassic nonmarine hot shales in the Ordos Basin, Central China. Appl. Geochem. 53, 42–52. https://doi.org/10.1016/j.apgeochem.2014.11.028.

Shapley, L.S., 1952. A value for N-Person games. RAND Corporation.

Shi, J., Zou, Y., Cai, Y.L., Zhan, Z.W., Sun, J.N., Liang, T., 2022. Organic matter enrichment of the Chang 7 Member in the Ordos Basin: Insights from chemometrics and element geochemistry. Mar. Petrol. Geol. 135, 105404. https://doi.org/10.1016/j.marpetgeo.2022.105404.

Stekhoven, D.J., Bühlmann, P., 2012. MissForest—Non-parametric missing value imputation for mixed-type data. Bioinformatics 28, 112–118. https://doi.org/10.1093/bioinformatics/btr597.

Tamoto, H., Gioria, R.D.S., Carneiro, C.D.C., 2023. Prediction of nuclear magnetic resonance porosity well-logs in a carbonate reservoir using supervised machine learning models. J. Pet. Sci. Eng. 220, 111169. https://doi.org/10.1016/j.petrol.2022.111169.

Tang, L., Song, Y., Pang, X.Q., Jiang, Z.X., Guo, Y.C., Zhang, H.G., Pan, Z.H., Jiang, H., 2020. Effects of paleo sedimentary environment in saline lacustrine basin on organic matter accumulation and preservation: A case study from the Dongpu Depression, Bohai Bay Basin China. J. Petrol. Sci. Eng. 185, 106669. https://doi.org/10.1016/j.petrol.2020.106669.

Tribovillard, N., Algeo, T.J., Lyons, T., Riboulleau, A., 2006. Trace metals as paleoredox and paleoproductivity proxies: An update. Chem. Geol. 232, 12–32. https://doi.org/10.1016/j.chemgeo.2006.02.012.

Wang, C., Shi, G., 2019. Redox condition and organic carbon accumulation mechanism in the Cryogenian Nanhua Basin, South China: Insights from iron chemistry and sulfur, carbon, oxygen isotopes of the Datangpo Formation. Adv. Geo-Energy Res. 3 (1), 67–75. https://doi.org/10.26804/ager.2019.01.05.

Wang, C., Wang, Q.X., Chen, G.J., He, L., Xu, Y., Chen, L.Y., Chen, D.F., 2017. Petrographic and geochemical characteristics of the lacustrine black shales from the Upper Triassic Yanchang Formation of the Ordos Basin, China: Implications for the organic matter accumulation. Mar. Petrol. Geol. 86, 52–65. https://doi.org/10.1016/j.marpetgeo.2017.05.038.

Wang, E.Z., Guo, T.L., Li, M.W., Xiong, L., Dong, X.X., Zhang, N.X., Wang, T., 2022. Depositional environment variation and organic matter accumulation mechanism of marine–continental transitional shale in the Upper Permian Longtan Formation, Sichuan Basin, SW China. ACS Earth Space Chem. 6 (9), 2199–2214. https://doi.org/10.1021/acsearthspacechem.2c00133.

Wang, E.Z., Feng, Y., Guo, T.L., Li, M.W., Xiong, L., Lash, G.G., Dong, X.X., Wang, T., Ouyang, J.S., 2023a. Sedimentary differentiation triggered by the Toarcian Oceanic Anoxic Event and formation of lacustrine shale oil reservoirs: Organic matter accumulation and pore system evolution of the Early Jurassic sedimentary succession, Sichuan Basin, China. J. Asian. Earth Sci. 256, 105825. https://doi.org/10.1016/j.jseaes.2023.105825.

Wang, X.B., Dong, L., Li, T., Wang, Z.Y., Zhang, C.H., Jin, Z.J., Fu, J.H., Zhu, R.K., Ling, K., Wang, T.Y., 2023b. Late diagenetic alteration of carbonate evidenced by carbonate-deficient siliciclastic laminae. Chem. Geol. 622, 121394. https://doi.org/10.1016/j.chemgeo.2023.121394.

Wang, Q., Chen, D., Li, M., Li, S., Wang, F., Yang, Z., Zhang, W., Chen, S., Yao, D., 2023c. A novel method for petroleum and natural gas resource potential evaluation and prediction by support vector machines (SVM). Appl. Energy 351, 121836. https://doi.org/10.1016/j.apenergy.2023.121836.

Wang, E.Z., Li, M.W., Ma, X.X., Qian, M.H., Cao, T.T., Li, Z.M., Yang, W.W., Jin, Z.J., 2024. Diahopane and diasterane as the proxies for paleoenvironment, hydrocarbon generation condition, and shale oil accumulation. Chem. Geol. 670, 122447. https://doi.org/10.1016/j.chemgeo.2023.121394.

Wang, E.Z., Fu, Y.X., Guo, T.L., Li, M.W., 2025a. A new approach for predicting oil mobilities and unveiling their controlling factors in a lacustrine shale system: Insights from interpretable machine learning model. Fuel 379, 132958. https://doi.org/10.1016/j.fuel.2024.132958.

Wang, E.Z., Li, M.W., Ma, X.X., Qian, M.H., Cao, T.T., Li, Z.M., Li, S., Jin, Z.J., 2025b. Can clay-rich reservoirs in predominantly-freshwater lacustrine shale systems serve as primary exploration targets in low-medium maturity? A case study of the Triassic Yanchang Formation of the Ordos Basin. Pet. Sci. 22 (10), 3925–3938. https://doi.org/10.1016/j.petsci.2025.06.020.
Wedepohl, K.H., 1971. Environmental influences on the chemical composition of shales and clays. Phyh. Chem. Earth 8 (71), 307–333. https://doi.org/10.1016/0079-1946(71)90020-6.
Wei, W., Algeo, T.J., 2020. Elemental proxies for paleosalinity analysis of ancient shales and mudrocks. Geochem. Cosmochim. Acta 287, 341–366. https://doi.org/10.1016/j.gca.2019.12.008.
Wignall, P.B., Twitchett, R.J., 1996. Oceanic anoxia and the end Permian mass extinction. Science 272 (5265), 1155–1158. https://doi.org/10.1126/science.272.5265.1155.
Wu, Z.Y., Zhao, X.Z., Wang, E.Z., Pu, X.G., Lash, G., Han, W.Z., Zhang, W., Feng, Y., 2021. Sedimentary environment and organic enrichment mechanisms of lacustrine shale: A case study of the Paleogene Shahejie Formation, Qikou Sag, Bohai Bay Basin. Palaeogeogr. Palaeoclimatol. Palaeoecol. 573, 110404. https://doi.org/10.1016/j.palaeo.2021.110404.
Xiao, L., Lo, S., Liu, J., Zhou, J., Li, Q., 2021. Nonlinear and synergistic effects of TOD on urban vibrancy: Applying local explanations for gradient boosting decision tree. Sustain. Cities Soc. 72, 103063. https://doi.org/10.1016/j.scs.2021.103063.
Xie, H.Y., Jiang, Z.X., Wang, L., Xue, X.Y., 2024. Organic matter enrichment model of fine-grained rocks in volcanic rift lacustrine basin: A case study of lower submember of second member of Lower Cretaceous Shahezi Formation in Lishu rift depression of Songliao Basin, NE China. Petrol. Explor. Dev. 51 (5), 1232–1246. https://doi.org/10.1016/S1876-3804(25)60537-4.
Yang, Y., Qiu, L., Wan, M., Jia, X., Cao, Y., Lei, D., Qu, C., 2019. Depositional model for a salinized lacustrine basin: The permian Lucaogou Formation, Jimsar Sag, Junggar Basin, NW China. J. Asian Earth Sci. 178, 81–95. https://doi.org/10.1016/j.jseaes.2019.03.008.
Yang, Y.Y., Liu, Y.Q., Zhou, D.W., Jiao, X., Cao, Q., Meng, Z.Y., Zhao, M.R., 2022. Lithotypes, organic matter and paleoenvironment characteristics in the Chang 73 submember of the Triassic Yanchang Formation, Ordos Basin, China: Implications for organic matter accumulation and favourable target lithotype. J. Petrol. Sci. Eng. 216, 110691. https://doi.org/10.1016/j.petrol.2022.110691.
Yuan, W., Liu, G., Zhou, X., Xu, L., Li, C., 2020. Palaeoproductivity and organic matter accumulation during the deposition of the Chang 7 organic-rich shale of the Upper Triassic Yanchang Formation, Ordos Basin, China. Geol. J. 55 (4), 3139–3156. https://doi.org/10.1002/gj.3578.
Zhang, W., Yang, H., Xie, L., Yang, Y., 2010. Lake-bottom hydrothermal activity and their influences on the high-quality source rock development: A case from Chang 7 source rocks in Ordos Basin, China. Petrol. Explor. Dev. 37 (4), 424–429. https://doi.org/10.1016/S1876-3804(10)60057-1.
Zhang, W., Yang, W., Xie, L., 2017. Controls on organic matter accumulation in the Triassic Chang 7 lacustrine shale of the Ordos Basin, central China. Int. J. Coal Geol. 183, 38–51. https://doi.org/10.1016/j.coal.2017.09.015.
Zhang, L.C., Xiao, D.S., Lu, S.F., Jiang, S., Lu, S.D., 2019. Effect of sedimentary environment on the formation of organic-rich marine shale: Insights from major/trace elements and shale composition. Int. J. Coal Geol. 204, 34–50. https://doi.org/10.1016/j.coal.2018.12.001.
Zhang, K., Liu, R., Liu, Z.J., Li, B.L., Han, J.B., Zhao, K.G., 2020. Influence of volcanic and hydrothermal activity on organic matter enrichment in the Upper Triassic Yanchang Formation, southern Ordos Basin, Central China. Mar. Petrol. Geol. 112, 104059. https://doi.org/10.1016/j.marpetgeo.2019.104059.
Zhang, Y., Tiňo, P., Leonardis, A., Tang, K., 2021. A Survey on neural network interpretability. IEEE Trans. Emerg. Top. Comput. Intell. 5 (5), 726–742. https://doi.org/10.1109/TETCI.2021.3100641.
Zhao, W.Z., Zhu, R.K., Hu, S.Y., Hou, L.H., Wu, S.T., 2020. Accumulation contribution differences between lacustrine organic-rich shales and mudstones and their significance in shale oil evaluation. Petrol. Explor. Dev. 47 (6), 1160–1171. https://doi.org/10.1016/S1876-3804(20)60224-1.
Zheng, R., Zeng, W., Li, Z., Chen, X., Man, K., Zhang, Z., Wang, G., Shi, S., 2022. Differential enrichment mechanisms of organic matter in the Chang 7 Member mudstone and shale in Ordos Basin, China: Constraints from organic geochemistry and element geochemistry. Palaeogeogr. Palaeoclimatol. Palaeoecol. 601, 111126. https://doi.org/10.1016/j.palaeo.2022.111126.
Zhong, S., Zhang, Y., Zhang, H., 2022. Machine learning-assisted QSAR models on contaminant reactivity toward four oxidants: Combining small data sets and knowledge transfer. Environ. Sci. Technol. 56, 681–692. https://doi.org/10.1021/acs.est.1c05307.
Zhu, L.Q., Zhang, C., Zhang, C.M., Zhang, Z.S., Zhou, X.Q., Liu, W.N., Zhu, B.Y., 2020. A new and reliable dual model- and data-driven TOC prediction concept: A TOC logging evaluation method using multiple overlapping methods integrated with semi-supervised deep learning. J. Petrol. Sci. Eng. 188, 106944. https://doi.org/10.1016/j.petrol.2019.106944.