Petroleum Science 22 (2025) 5047–5067

Contents lists available at ScienceDirect

Petroleum Science

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

Original Paper

A stacking ensemble approach for pore pressure prediction in real-time during drilling based on mud log data

Dong-Yang Zhang [a], Tian-Shou Ma [a,*], Yang Liu [b], De-Cheng Zhang [c]

[a] State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Southwest Petroleum University, Chengdu, 610500, Sichuan, China
[b] School of Mechatronic Engineering, Southwest Petroleum University, Chengdu, 610500, Sichuan, China
[c] School of Civil and Transportation Engineering, Hebei University of Technology, Tianjin, 300401, China

ARTICLE INFO

Article history:
Received 14 April 2025
Received in revised form
10 September 2025
Accepted 10 September 2025
Available online 17 September 2025

Edited by Jia-Jia Fei

Keywords:
Pore pressure
Stacking ensemble approach
Machine learning
Real-time
Mud log

ABSTRACT

The precise and real-time prediction of pore pressure is critical for optimizing drilling efficiency and mitigating the potential risks associated with drilling operations. In order to surmount the limitations of empirical methods and to reduce reliance on logging-while-drilling data, this study proposed a stacking ensemble approach that utilized only conventional mud log data. The workflow involves the preliminary processing of data using isolation forest and wavelet thresholding techniques to effectively eliminate outliers and noise. The Eaton index was estimated using a Bayesian inversion algorithm, and pore pressure was estimated by integrating the d_c exponent and Eaton method. A feature selection strategy combining data distribution characteristics and regression-based importance ranking was used to optimize the input parameters. Subsequently, a stacking approach was developed for predicting pore pressure, and the corresponding base learner, meta learner, and hyperparameters were optimized. Finally, the validity of the optimized model was substantiated through field data from three test wells (X1, X5, X3) under different drilling scenarios. The results indicated that the global optimal Eaton index of the block was 0.2449, the maximum percentage error of pore pressure estimation was 3.66%, and the corresponding average error was 1.59% compared to the wireline formation test data. The optimal combination of input features was determined to be R-ROP, WOB, TG, PT, MW, PFI, TVD, H, BR, BT, and SPP. The optimal basic learners were identified as BPNN, CNN, LSTM, and LightGBM, while the optimal meta learners were XGBoost. Prediction accuracy is improved when offset wells are densely distributed, spatially balanced, and proximal to the target well; conversely, sparse or distant offset wells result in reduced prediction performance. The mean absolute percentage errors for test offset well X1, X5, and X3 were 0.4353%, 0.4646%, and 0.6856%, respectively, with the corresponding R^2 values of 0.9362, 0.9078, and 0.8950, respectively. Consequently, this approach has the capacity to accurately and in real-time predict pore pressure using solely conventional mud log data. This capability enables timely adjustments to drilling parameters, thereby enhancing operational efficiency and mitigating drilling risks.

1. Introduction

Pore pressure refers to the fluid pressure within the pore spaces of subsurface formations, and is primarily controlled by the sedimentary environment, diagenetic processes, and the hydrodynamic evolution of the basin over geologic time (Azadpour et al., 2015; Zhang et al., 2024; Liu et al., 2024). As a key parameter in subsurface characterization, accurate prediction of pore pressure is critical for optimizing drilling fluid density, maintaining wellbore stability, and designing effective fracture stimulation strategies. In addition, reliable pore pressure estimation is essential to prevent major drilling hazards such as blowouts and wellbore leakages, which pose serious operational and environmental risks (Mitchell and Miska, 2011; Zhang, 2011; Peng et al., 2021; Zhang et al., 2023).

Pore pressure prediction methods for drilling operations can generally be divided into three types: pre-drilling prediction, while-drilling prediction, and post-drilling evaluation. Pre-drilling prediction primarily relies on seismic interval velocities to

* Corresponding author.
E-mail addresses: matianshou@126.com, matianshou@swpu.edu.cn (T.-S. Ma).
Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2025.09.020

estimate formation pore pressure, thereby assisting in well planning and mud weight design (Fillippone, 1982; Sayers et al., 2002). However, due to the limited resolution of seismic data, this approach often fails to capture localized pressure anomalies—especially within reservoir interiors or thin interbedded formations. Furthermore, its accuracy is constrained by seismic data quality, introducing significant uncertainty into the predictions (Ugwu, 2015). Post-drilling evaluation aims to refine pre-drilling predictions using a range of datasets acquired from completed wells, with logging-based approaches being the most commonly used methods (Eaton, 1975; Bowers, 1995; Lu et al., 2022). Although well logging interpretation provides a relatively accurate pressure estimate under various geological conditions, it is inherently limited to previously drilled wells. During the drilling phase, these methods often rely on logging-while-drilling (LWD) data, which may not always be available in real time due to operational constraints. Prediction while drilling enables real-time monitoring of pore pressure and drilling conditions, thus supporting field decision-making and mitigating operational risks. Common techniques in this category include the d_c exponent, sigma-logarithmic, and mechanical specific energy methods, which utilize surface drilling parameters such as rate of penetration (ROP) and rotary speed (RPM) to estimate pore pressure (Teale, 1965; Rehm and McClendon, 1971; Shajari and Najibi, 2012; Roy et al., 2010). However, these empirical methods often struggle to dynamically adapt to pressure variations, especially in geologically complex or heterogeneous formations. In addition, frequent changes in bottomhole assembly configurations, continuous advances in drilling technologies, and variations in rig equipment further compromise the robustness and accuracy of these static models (Chen et al., 2022).

In recent years, machine learning (ML) techniques have been widely applied in various areas of petroleum engineering, including rock mechanical property prediction, in-situ stress estimation, reservoir characterization, wellbore stability evaluation, and rockburst prediction (Otchere et al., 2021; Ma et al., 2022, 2024; Phan et al., 2022; Zhang et al., 2022; AlBahrani and Morita, 2022; Zhao et al., 2024; Fissha et al., 2025). In the context of pore pressure prediction, Ahmed et al. (2019) used artificial neural network (ANN) to estimate pore pressure by integrating mud log data such as weight on bit (WOB), RPM, and ROP with conventional logging data. Zhang et al. (2022) used decision tree (DT), support vector regression (SVR), and ANN to estimate pore pressure using logging data. Delavar and Ramezanzadeh (2023) developed least squares support vector machine (LSSVM), ANN, and random forest (RF) models to predict pore pressure in carbonate reservoirs by incorporating both conventional logging and mud log data, which significantly improved predictive performance. Collectively, these studies demonstrate that ML-based methods can outperform traditional empirical methods in predictive accuracy. However, a notable limitation of most existing approaches is their heavy reliance on logging data or combinations of logging and mud log data, which limits their application to wells where such logging data is readily available. Only a limited number of studies have investigated the feasibility of predicting pore pressure using only mud log data. For example, Rashidi and Asadi (2018) developed a neural network model using drilling efficiency and mechanical specific energy, both derived from mud log data, as input features to estimate pore pressure. Similarly, Ahmed et al. (2021) used SVR and functional networks to predict pore pressure using surface drilling parameters such as standpipe pressure (SPP), ROP, RPM, WOB, torque (T), and flow rate (FR). However, these models typically rely on point-wise prediction strategies that fail to characterize the continuous depth-dependent trend of pore pressure. Consequently, they are limited in their ability to generate complete and continuous pore pressure profiles, thus limiting their practical utility in real-time drilling operations.

Despite the advances attained by prevailing ML-based models, their practical implementation is encumbered by several limitations: (1) the majority depend on LWD or comprehensive logging data, which are not universally accessible during real-time operations; (2) only a few models utilize solely mud log data, and those that do frequently employ point-wise prediction strategies, lacking the capacity to capture continuous pore pressure profiles; and (3) the impact of data distribution discrepancies between wells is predominantly overlooked, thereby diminishing model robustness when applied to new wells. In order to surmount the limitations of empirical methods and to reduce reliance on LWD data, this study proposed a stacking ensemble approach that utilized only conventional mud log data. The overall workflow of the proposed methodology is illustrated in Fig. 1. Firstly, the mud log data from the working block was collected, and the outliers and noise in the data were processed. Secondly, the global optimal Eaton index of the working block was inverted based on the measured pore pressure values obtained from the wireline formation test (WFT), and the pore pressure was estimated according to the inversion results. Thirdly, the input features were optimized based on the data distribution characteristics and feature importance, and the input parameter combination was determined. Subsequently, the base learners and meta-learners of the stacking model were selected, and the hyperparameters were optimized. Finally, the pore pressure of the test wells under three different scenarios was predicted to verify the prediction accuracy and generalization ability of the model.

This comprehensive methodology enables the accurate prediction of continuous pore pressure profiles in real time, which is validated across multiple drilling scenarios. Thus, the following key innovations are the most significant contributions of this study: (1) The implementation of a stacking ensemble model that integrates multiple learners has been demonstrated to improve predictive performance. (2) The reliability of the predicted pore pressure values is increased by means of a Bayesian inversion technique, which is used to derive the optimal Eaton index for the entire working block. (3) A feature selection strategy was developed that accounts for inter-well data distribution differences with the objective of mitigating generalization issues. (4) The employment of Bayesian optimization in the refinement of the hyperparameters of the ensemble model has been demonstrated to enhance both predictive accuracy and robustness.

2. Mud log data analysis

This study is conducted in a block located within the H depression of the ZJ Basin, South China Sea. The target formation consists mainly of medium- to coarse-grained sandstones deposited in a braided river delta to Tanba frontal facies. These sandstones exhibit good sorting and high compositional maturity, with porosities ranging from 10% to 18% and permeabilities between 10 and 100 mD, characterizing them as high-quality reservoir rocks. An effective caprock seal is provided by the overlying regional mudstone. The lithological assemblage and petrophysical properties of the interbedded sandstone-mudstone sequence exert strong control over the reservoir pore structure, providing a solid geological foundation for the development of an empirical pore pressure prediction model. These favorable geological characteristics enhance the model's ability to learn and generalize pore pressure variation trends with depth, ultimately improving prediction accuracy and reliability.

The dataset for this study consists of mud log data from six wells drilled on the study block (X1, X2, X3, X4, X5 and X6). Their

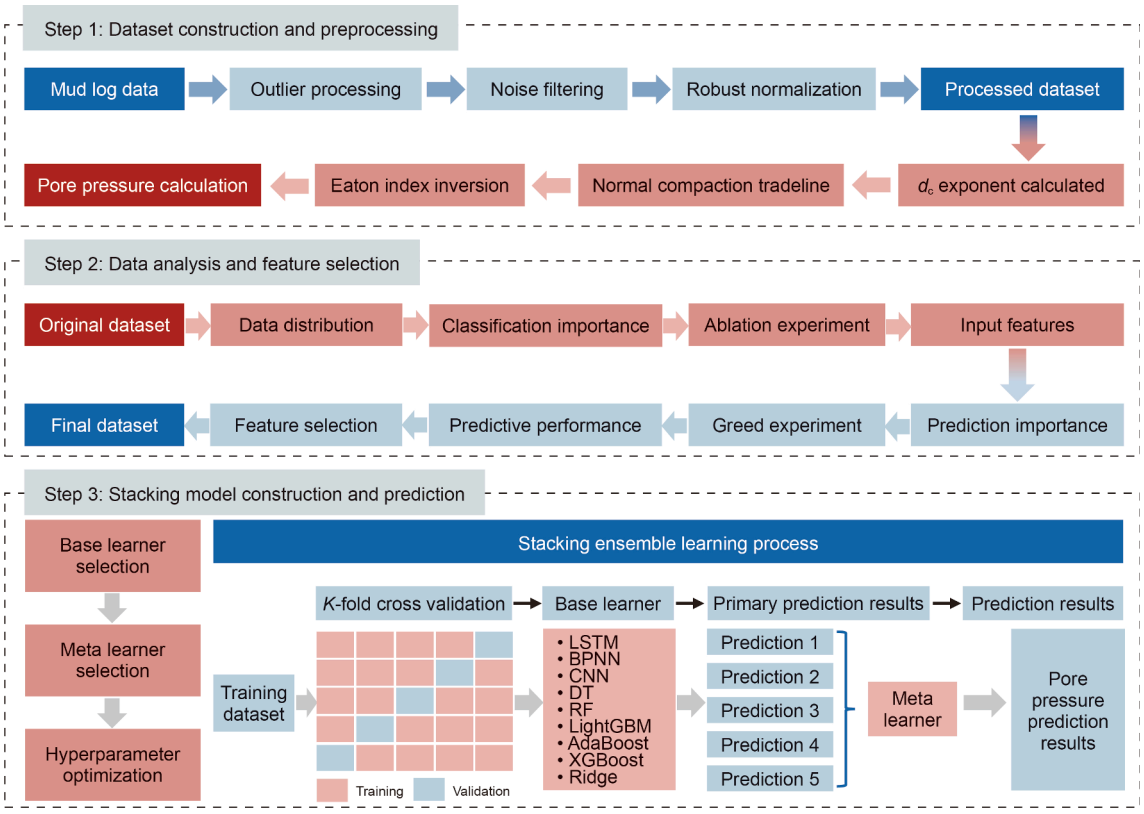


Fig. 1. Schematic diagram of pore pressure prediction approach using stacking model.

spatial distribution is shown in Fig. 2, where X3 is a vertical well and the remaining wells are directional wells. Due to formation variations, geological conditions can change significantly at the same depth, resulting in differences in drilling parameters and conditions between wells. This presents a significant challenge to pore pressure prediction from mud log data. In addition, the Well X3 is geographically distant from several other wells, further increasing the difference in data distribution. This exacerbates the difficulty of using offset data to predict formation pore pressure in Well X3 and highlights the need for robust and adaptive prediction models.

Table 1 presents a comprehensive description and statistical summary of the features collected from the six wells. The data exhibit considerable variability across different features. For instance, the WOB ranges from 0.93 to 186.30 kN, while the R-ROP varies from 0.31 to 142.86 min/m. These wide ranges highlight substantial fluctuations in drilling conditions, both within individual wells and across different wells. Additionally, the raw dataset contains outliers and noise, which are primarily attributed to sensor inaccuracies and unexpected environmental disturbances during data acquisition. If left unprocessed, such low-quality data can negatively impact the training of machine learning models, resulting in increased prediction errors. To mitigate these issues and enhance model performance, this study implements a rigorous data preprocessing workflow prior to model training. This includes outlier detection, noise reduction, and feature normalization, ensuring that the input data are of high quality and suitable for robust model learning.

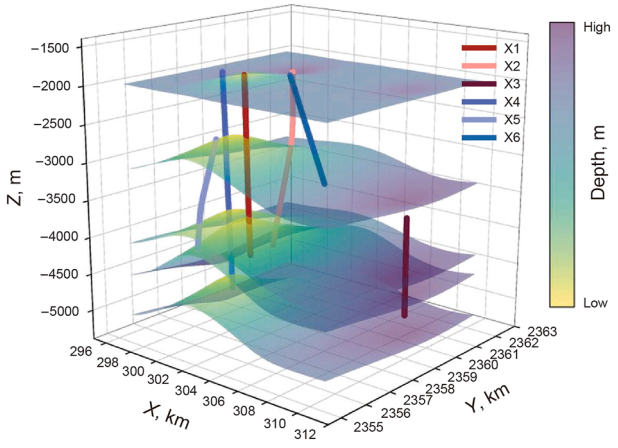


Fig. 2. The spatial location of the well trajectory.

3. Methodology

3.1. Data preprocessing

3.1.1. Outlier and noise processing

Data preprocessing is a critical step in cleaning, modifying, and organizing raw data prior to training ML models, commonly referred to as data preparation (Alasadi and Bhaya, 2017). Mud log data frequently contain a considerable number of outliers and noise due to sensor malfunctions and human mistakes, which can adversely affect the performance of ML models. To tackle this problem, this study used the Isolation Forest (IForest) algorithm for outlier detection and elimination. After identifying and excluding outliers, the linear interpolation method is adopted to fill in the missing values, guaranteeing the continuity and integrity of the dataset. To further mitigate the influence of noise on model

Table 1
Features descriptions for the data set.

Feature name	Description	Unit	Min	Mean	Max	Std
MD	Measured depth	m	1950.00	3496.15	4978.00	768.26
TVD	True vertical depth	m	1903.84	3386.68	4978.00	743.94
HKL	Hook load	kN	616.52	1287.78	2163.20	419.11
WOB	Weight on bit	kN	0.93	118.57	186.30	23.28
RPM	Revolutions per minute	r/min	30.00	110.04	162.00	27.52
SPP	Standpipe pressure	MPa	3.72	18.46	26.62	3.59
T	Torque	kN·m	4.21	18.54	32.00	4.48
R-ROP	Reciprocal of ROP	min/m	0.31	4.35	142.86	4.81
PFI	Pump flow in rate	L/min	643.00	3409.75	4399.00	994.14
PFO	Pump flow out rate	L/min	0	1104.90	1716.15	461.55
BT	Bit time	h	0	53.68	195.27	39.41
BR	Bit run	m	0	1512.04	3246.95	933.32
PT	Pump time	h	0.11	89.94	304.43	61.58
MW	Mud weight	g/cm^3	0.89	1.25	1.40	0.06
HH	Hook height	m	−0.15	14.36	31.09	8.32
TG	Total gas	ppm	0	9719.53	330558.00	20495.61

training, this study employs the wavelet threshold denoising technique, which effectively filters and reduces the noise in the data. This preprocessing strategy improves data quality and ensures the accuracy and robustness of the ML model training process.

The IForest is a decision tree-based outlier detection approach, where outliers are identified by randomly selecting features and partitioning the data. Unlike traditional anomaly detection methods based on distance or density, the IForest does not rely on direct measures such as distance or density. Instead, it characterizes the "sparsity distance" of data points, or how quickly they become isolated from the data set (Alsini et al., 2021). The IForest consists of multiple isolation trees, and the outlier score for each tree is determined by calculating the path length of data points within the tree. Since outliers are more likely to be isolated near the root node, they tend to have shorter path lengths and therefore higher outlier scores. As shown in Fig. 3, the structure of an isolation tree (ITree) is similar to a binary search tree, allowing the use of binary search tree algorithms to derive the average path length for a given data point (Liu et al., 2008).

The average depth of branches traversed for a given data point is then converted to an anomaly score using Eq. (1).

$$S(x,n)=2^{-\frac{E[h(x)]}{c(n)}} \tag{1}$$

where $S(x, n)$ is the anomaly score of x, in the range (0, 1). The closer the value is to 1, the more likely it's an outlier. The closer the value is to 0, the more likely it's to be a normal point; x is the input point; n is the total number of samples; $E(h(x))$ is the mean value of the depths a single data point (x) reaches in all trees; $c(n)$ is the normalizing factor defined as the average depth in an unsuccessful search in a Binary Search Tree (BST):

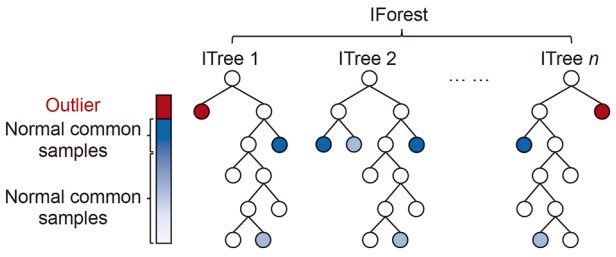


Fig. 3. Distribution of abnormal scores in IForest.

$$c(n)=2H(n-1)-\frac{2(n-1)}{n} \tag{2}$$

where H is the harmonic number; n is the number of points used in the construction of trees.

After dealing with the outliers in the data, further noise filtering is still necessary to improve the data quality. In this research, a wavelet threshold denoising approach is used to denoise the mud log data. The specific principle and procedure can be found in Zhang et al. (2025). In this study, after numerous attempts, we finally determined the model parameters used in the IForest and wavelet threshold denoising methods. Specifically, the number of trees in the IForest was set to 50, and the proportion of outliers was 0.02. For wavelet threshold denoising, the "db8" wavelet was used, and the final threshold was 0.1. Taking X4 as an example, Fig. 4 presents the comparison effect between the data after outlier processing and noise filtering and the original data. It can be observed that the processed mud log data effectively eliminates the outliers, and the burr phenomenon of the data is significantly smoothed, with the overall fluctuation tending to stabilize. Such processing significantly improves the data quality and provides more reliable base data for training the subsequent machine learning model.

3.1.2. Data normalization

In the domains of machine learning and data analytics, data normalization constitutes a crucial preprocessing stage that can notably enhance the convergence speed of the model and its predictive performance (Mining, 2006). The magnitudes of diverse mud recording parameters exhibit a wide variance. If not normalized, the machine learning model might be biased toward the parameters of larger magnitudes during training, resulting in biased prediction results.

Common data normalization approaches encompass max-min normalization methods predicated on the maximum value, along with normalization methods founded on the mean and standard deviation (Evans, 2006). The max-min normalization method is applicable when the data distribution is relatively stable and there are no outliers or the outliers exert minimal influence. Its essence lies in scaling the data to a fixed range (Sharma, 2022). Nevertheless, this method is highly sensitive to the maximum value of the data. If outliers exist in the data, it might lead to the compression of the dynamic range of the normal data, thereby influencing model performance. On the contrary, the normalization method is suitable for scenarios where the data conforms to a

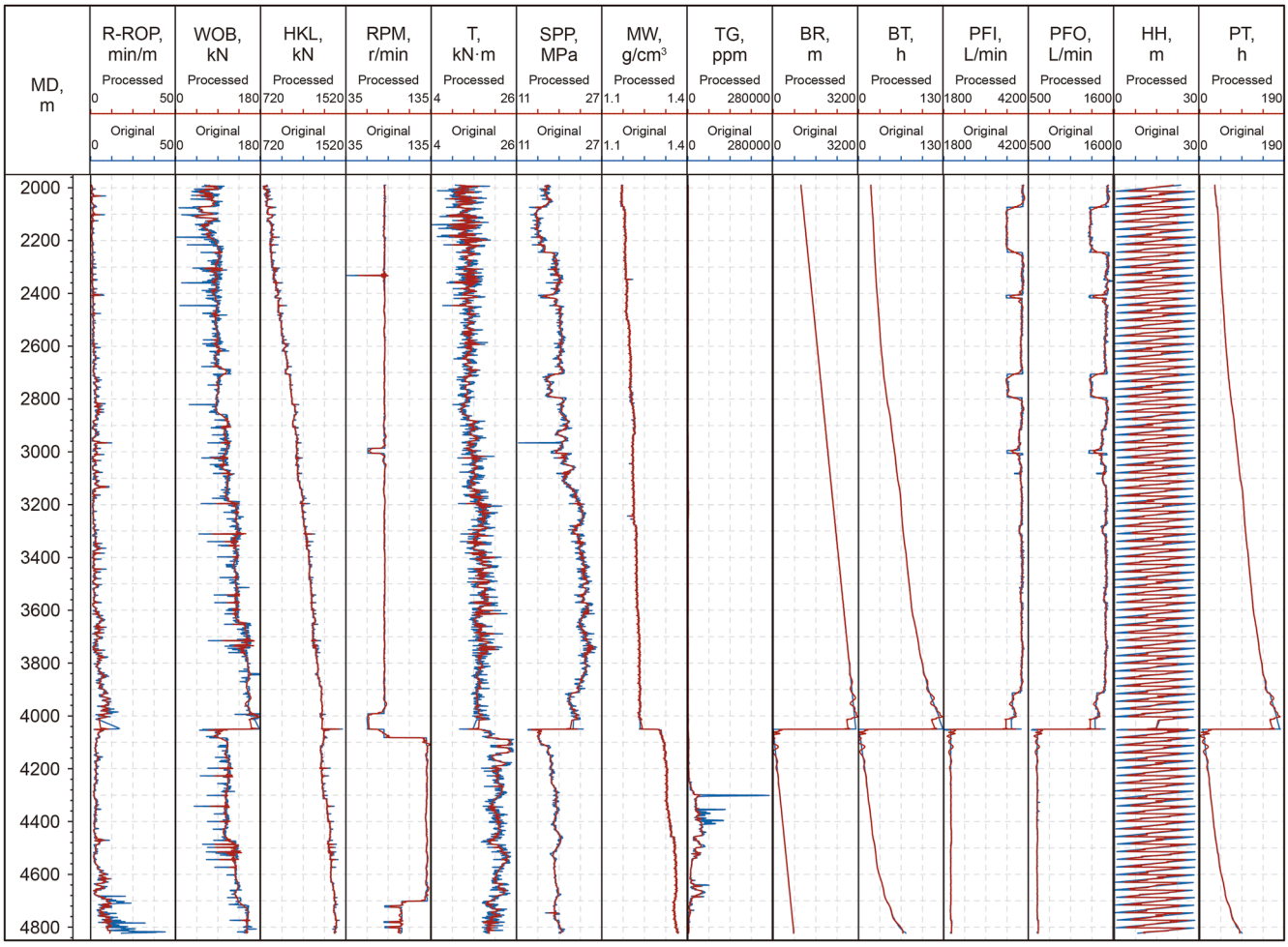


Fig. 4. Results of data comparison before and after processing.

normal distribution and is normalized by subtracting the mean and dividing by the standard deviation. However, this method is prone to extreme values, which can cause the normalization results to be skewed (Gal and Rubinfeld, 2019). Despite the previous steps having dealt with the data for outliers, it is impossible to guarantee that all outliers have been effectively eliminated. To further guarantee the quality of the normalized data, the robust scaler method was selected for data normalization in this study.

Robust scaler is a normalization approach that is less susceptible to outliers, scaling based on the median and interquartile range of the data instead of the mean and standard deviation or maximum and minimum values of conventional methods. This confers a distinct advantage to robust scaler when handling circumstances that involve a large number of outliers or where the data distribution is highly skewed. Contrary to other normalization methods, robust scaler does not enforce the data to be compressed within a fixed range but rather adapts to the distribution of the data itself (Rousseeuw and Croux, 1993). For each feature x, the transformation formula of the robust scaler is:

$$x_{\text{scaled}} = \frac{x - Q_2}{IQR} \tag{3}$$

where x_{scaled} is the normalized value; x is the original value; Q_2 is the median of the feature; IQR is the interquartile range of the feature.

3.2. Pore pressure estimation

Under normal pore pressure circumstances, the ROP typically exhibits a decreasing tendency as the well depth increases. Nevertheless, when drilling into the pressure transition zone, the ROP accelerates on account of the reduction in pore pressure difference and the increase in rock porosity. The d_c exponent method is founded on this variance in variation and is employed to identify regions with abnormally high pressure. The method computes the d_c exponent from key parameters such as ROP, WOB, and RPM and corrects for mud density to predict the pore pressure gradient. Eq. (4) provides the formula for calculating the d_c exponent (Rehm and McClendon, 1971):

$$d_c = \frac{\log\left(\frac{3.282}{NT}\right)}{\log\left(\frac{0.0684W}{D_b}\right)} \times \frac{\rho_n}{\rho_d} \tag{4}$$

where d_c is d_c exponent; N is rotational speed, r/min; T is time of drilling, min/m; W is weight on bit, kN; D_b is bit diameter, mm; ρ_n is normal mud density, g/cm^3; ρ_d is actual drilling fluid density, g/cm^3.

After calculating the d_c exponent, it is also necessary to establish its normal compaction trend line and thus estimate the pore

pressure by Eaton's method, which can be expressed as (Eaton, 1975):

$$\rho_p = \rho_o - (\rho_o - \rho_n)\left(\frac{d_c}{d_{cn}}\right)^m \tag{5}$$

where ρ_p is pore pressure equivalent density, g/cm^3; ρ_o is overburden pressure equivalent density, g/cm^3; ρ_n is hydrostatic pressure equivalent density, g/cm^3; d_{cn} is the d_c exponent corresponding to the normal compaction line at a given depth; m is the Eaton index.

Pore pressures were measured at 13 points through the WFT. To acquire the optimal Eaton index for the block, which minimizes the error between the calculated and measured pore pressures, a Bayesian optimization method was employed for the inversion of the Eaton index. The process of Bayesian optimization for the inversion of the Eaton index is depicted in Fig. 5. This process can be simply described as: set the initial m to calculate the pore pressure, obtain the loss value (*Loss*) between the calculated value and the actual value, substitute it into another m for calculation, and repeat the iteration until the loss value drops below the threshold or the maximum number of iterations is reached. At this point, the corresponding m is the optimum Eaton index for the block. The formula for *Loss* can be expressed as follows:

$$Loss = \sum_i \left(\hat{\rho}_p^{(i)} - \rho_p^{(i)}(m)\right)^2 \tag{6}$$

where i is the currently selected measurement point; $\hat{\rho}_p^{(i)}$ is the measured pore pressure equivalent density of the point i, g/cm^3; $\rho_p^{(i)}(m)$ is the calculated pore pressure equivalent density at the point i with current Eaton index, g/cm^3.

It should be noted that the pore pressure labels adopted in this study are calculated based on the Eaton method and calibrated using 13 WFT measured points from 6 wells. Due to the inevitable limitation of the measured data, this method provides a feasible means for pressure estimation, but it also inevitably introduces the assumptions of deterministic physical models, thus giving the labels a certain "synthetic" nature. Therefore, the model in this study is essentially more of a fit to the Eaton method rather than a modelling completely dependent on measured data.

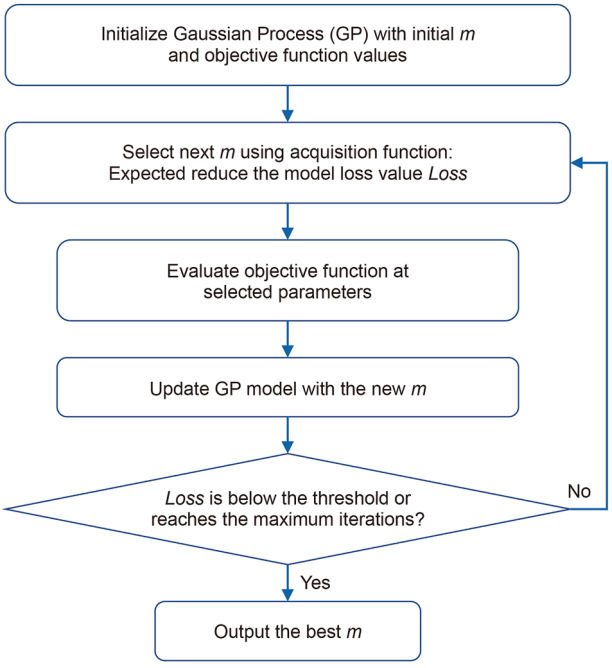


Fig. 5. Flowchart of Eaton index inversion by Bayesian algorithm (modified from Mughees et al., 2024).

3.3. Feature selection

3.3.1. Feature selection based on data distribution

In pore pressure prediction studies, feature selection is a critical aspect that affects model performance. Existing research primarily relies on the correlation between input features and target variables for feature selection, and input features are typically determined by computing the correlation coefficients between mud log data and pore pressure. However, the generally low correlation between mud log data and pore pressure limits the effectiveness of correlation-based feature selection methods to some extent. Additionally, as depicted in Fig. 6, there is a remarkable disparity in the distribution of mud log data among different wells. This distribution difference results in the poor applicability of the highly correlated features screened in the training set to the target wells, which affects the generalization ability and predictive accuracy of the model. If the feature selection is based only on the correlation of the training set, the predictive power of the model on the test wells can be significantly reduced. Therefore, the feature selection method based only on correlation may not be able to effectively solve the cross-well prediction problem. Based on this, this study analyzes the data distribution characteristics between the source domain (training set) and the target domain (test set), and filters out the features with smaller distribution differences to mitigate the negative impact of distribution inconsistencies on model performance.

In this study, a RF classification model is constructed in which the mud log data from the source and target domains are combined and fed into the model, with the goal of enabling the classification model to accurately discriminate whether the data is from the source or target domain (Jaiswal and Samikannu, 2017). The method evaluates the contribution of different features to the classification accuracy by calculating the importance scores of the features, thereby identifying the parameters that have significant differences in distribution between the source and target domains. The workflow of the model is depicted in Fig. 7. Unlike traditional metrics such as Maximum Mean Discrepancy (MMD), Kullback-Leibler (KL) divergence, or Earth Mover's Distance (EMD), which directly quantify the statistical distance between two distributions, this study adopts a classifier-based domain discrepancy evaluation strategy. Specifically, a RF classifier was trained to distinguish between the source and target domain samples based on their mud log features. The importance scores derived from this classifier serve as proxies for measuring how significantly each feature contributes to the domain distinction. This approach allows us not only to identify features that exhibit substantial cross-domain variability but also to evaluate their impact on model performance through ablation experiments. It provides a practical and interpretable alternative to direct statistical metrics, particularly when working with high-dimensional, noisy mud log data.

3.3.2. Feature selection based on feature importance

After data distribution analysis and initial feature screening, this study integrates the XGBoost and SHAP (Shapley Additive exPlanations) approaches for further feature selection. Unlike the conventional importance scoring method, which is based on the weighted average of the decision tree split frequency and split gain, the SHAP method is able to quantitatively evaluate the marginal contribution of each feature to the prediction results,

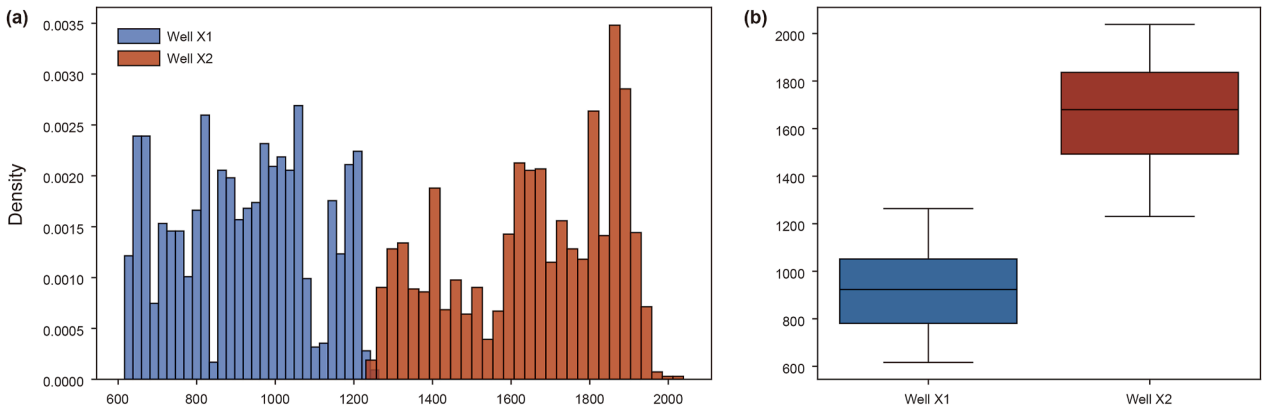


Fig. 6. Data distribution comparison of offset wells.

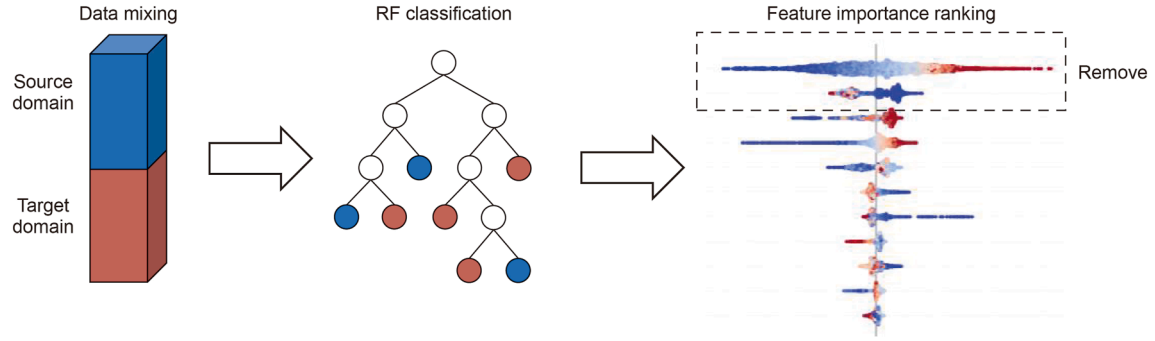


Fig. 7. Feature selection flow based on data distribution.

thus providing a more detailed and intuitive model interpretation (Rzychoń et al., 2022).

The SHAP method derives from the Shapley value in cooperative game theory, and its core concept is to treat the prediction problem as a revenue distribution problem, and to achieve model interpretation by quantifying the marginal contribution of each feature to the prediction outcome. Specifically, for a feature set N and a prediction function f, the SHAP value ϕ_i of each feature $i \in N$ is defined as (Van Zyl et al., 2024; Huang et al., 2024):

$$\phi_i = \sum_{\mathbf{S}\subseteq N\setminus\{i\}} \frac{|\mathbf{S}|!(M-|\mathbf{S}|-1)!}{M!}[f_x(\mathbf{S}\cup\{i\}) - f_x(\mathbf{S})] \tag{7}$$

where $N = \{1, 2, \ldots, M\}$ is the subscripts of the feature variables in the dataset, and M is the total number of feature variables; $\mathbf{S}$ is a subset of the set $\{1, 2, \ldots, M\}$; $|\mathbf{S}|$ is the total number of elements in $\mathbf{S}$; $f_x(\mathbf{S}\cup\{i\})$ is the predicted value of the model when there are only the features in $\mathbf{S}\cup\{i\}$; $f_x(\mathbf{S})$ is the predicted value of the model when there are only features in $\mathbf{S}$. The subtraction of the two is the marginal contribution of the ith feature variable under the subset $\mathbf{S}$.

3.4. Stacking ensemble approach

The core concept of ensemble learning is to integrate multiple individual learners via a particular strategy to achieve superior generalization performance compared to a single learner (Dietterich, 2000). Among ensembles, they can be classified into homogeneous and heterogeneous ensembles based on the types of individual learners and their interrelationships. In contrast to homogeneous ensembles, which mainly depend on the fusion of several identical or similar models through voting, weighting, etc., heterogeneous ensembles exploit the respective advantages of different types of machine learning algorithms to capture the multidimensional characteristics of the data (Zounemat-Kermani et al., 2021). Among them, stacking is a typical heterogeneous ensemble approach. It achieves a more refined delineation of the decision boundary and prediction optimization by constructing a multi-level model structure and passing the prediction results of different base learners as inputs to the top-level meta learner (Chatzimparmpas et al., 2020).

The first aspect concerns the selection and combination of base learners. Ideal base learners should exhibit considerable variability in model structure, algorithmic principles, and fitting capacity, ensuring that each model focuses on data features differently, thereby complementing each other's deficiencies during ensemble and mitigating the overfitting or bias problems that might exist in a single model. The next step is to design the meta-learner. The meta-learner is responsible for re-learning the prediction results produced by the base-learner, and its design requires both a strong fitting capability to capture the non-linear relationships among the base-models and the avoidance of excessive model complexity due to over-reliance. Through the rational design of the meta-learner and its training process, the generalization performance of the overall model can be effectively enhanced (Kardani et al., 2021).

In this study, linear regression models, namely Ridge, Lasso, and ElasticNet, tree-based models, including DT, RF, AdaBoost, XGBoost, LightGBM, and neural network models, such as BPNN, CNN, and LSTM, were initially chosen for the design of base and

meta-learners in the stacking model (Huang et al., 2023; Pei et al., 2024; Dong et al., 2023; Matinkia et al., 2022; Carneiro et al., 2022; Yang and Wen, 2018; Shehadeh et al., 2021).

In addition, given the existence of a learner model for the stacking model and the fact that multiple parameters in each learner exert a significant influence on the predictive performance of the model, this study incorporates a Bayesian optimization algorithm to optimize the hyperparameters of each learner. Bayesian optimization achieves global optimization of a complex objective function by establishing an agent model based on a Gaussian process that effectively explores the hyperparameter space within a limited number of trials. This method not only automates the search for optimal parameter combinations but also demonstrates high efficiency and robustness in scenarios where the parameter space is extensive and evaluation is costly. The pseudocode of stacking's modelling and optimization process is shown in Table 2. Firstly, the hyperparameters of the base learners and meta-learner in the stacking model were optimized using the Bayesian optimization algorithm. Then, the input parameters were fed into each base learner to obtain the prediction results of each base learner, which were combined as the input features of the meta-learner. Finally, these features were input into the meta-learner to obtain the final pore pressure prediction results.

3.5. Evaluation index

In this study, the mean absolute error, root mean square error, absolute percentage error, mean absolute percentage error and R^2 were used to evaluate the merit of the model's prediction performance, and the equations can be expressed as respectively (Liang et al., 2024):

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}|y_i-\widehat{y}_i| \tag{8}$$

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\widehat{y}_i)^2} \tag{9}$$

$$\mathrm{APE}=\frac{|y_i-\widehat{y}_i|}{y_i}\times 100\% \tag{10}$$

$$\mathrm{MAPE}=\frac{1}{n}\sum_{i=1}^{n}\frac{|y_i-\widehat{y}_i|}{y_i}\times 100\% \tag{11}$$

$$R^2=\frac{\sum_{i=1}^{n}(\widehat{y}_i-\overline{y}_i)^2}{\sum_{i=1}^{n}(y_i-\overline{y}_i)^2} \tag{12}$$

where MAE is the mean absolute error, g/cm^3; RMSE is the root mean square error, g/cm^3; APE is the absolute percentage error, %; MAPE is the mean APE, %; R^2 is the R-squared; y_i is the true value of ρ_p, g/cm^3; $\hat{y}_i$ is the predicted value of ρ_p, g/cm^3; $\overline{y}_i$ is the mean of the true value, g/cm^3; n is the number of the data.

Among the above evaluation indicators, the APE is used to evaluate the predictive effect of each data point, while the remaining four indicators are used to evaluate the overall predictive performance of the model.

4. Results

4.1. Pore pressure estimation

The inversion of m was conducted based on the measured values of pore pressure acquired from the WFT in combination with the Bayesian optimization method. Fig. 8 depicts the trend of error variation at different m during 50 iterations, with some points overlapping. It can be observed that as m increases, the

Table 2
Pseudo-code of stacking modeling and optimization process.

Algorithm: Stacking ensemble with Bayesian hyperparameter optimization

Input:
Training set $\boldsymbol{D}$, base learners B_1, B_2, …, B_T, meta learner M, number of folds k, hyperparameter bounds *bounds*;
Output:
Optimized hyperparameters for base and meta learners, predictions of base learners, predictions of meta learner;
Define the Bayesian optimization objective function:
  Function *ObjectiveFunction* (hyperparameters):
    Train/evaluate model with hyperparameters;
    Return performance metric;
Initialize Bayesian optimizer *BayesianOptimizer* with *ObjectiveFunction* and *bounds*;
Split $\boldsymbol{D}$ into k equally sized folds $\boldsymbol{D}_1$, $\boldsymbol{D}_2$, …, $\boldsymbol{D}_k$;
Initialize list $ListPreds = \phi$ to store base learner predictions;
For each base learner $B_t \in \{B_1, B_2, ..., B_T\}$:
  Query *BayesianOptimizer* for suggested hyperparameters θ_{Ri};
  Initialize prediction list $Preds_{Ri} = \phi$;
  For $l = 1$ to k:
    Set test fold $\boldsymbol{D}_{\mathrm{test}} = \boldsymbol{D}_\mathrm{i}$, training folds $\boldsymbol{D}_{\mathrm{train}} = \boldsymbol{D}\backslash\boldsymbol{D}_i$;
    Train R_t on $\boldsymbol{D}_{\mathrm{train}}$ with hyperparameters θ_{Ri};
    Predict on $\boldsymbol{D}_{\mathrm{test}}$ and append results to $Preds_{Ri}$;
  Compute mean cross-validation performance Mr_i across k-folds;
  Update *BayesianOptimizer* with (θ_{Ri}, Mr_i);
  Append $Preds_{Ri}$ to *ListPreds*;
Construct meta-dataset $\boldsymbol{D}_{\mathrm{meta}}$:
  Features: Concatenate predictions from *ListPreds;*
  Target: Original labels of $\boldsymbol{D}$;
Query *BayesianOptimizer* for meta-learner hyperparameters θ_R;
Train meta-learner R on $\boldsymbol{D}_{\mathrm{meta}}$ with θ_R;
Output predictions of meta learner;
Return:
  Optimized hyperparameters {θ_{R1}, …, θ_{RT}, θ_R}, predictions of base learners, predictions of meta learner.

error between the calculated value and the measured value initially decreases and subsequently increases, and the error reaches its minimum at $m = 0.2449$. Hence, the optimal Eaton index for this block can be determined as 0.2449, and this value can be utilized for the subsequent calculation of the pore pressure labelling value to ensure that the calculated pore pressure value approximates the actual pore pressure.

Based on the determined Eaton index, the pore pressures of the six wells were estimated by integrating Eqs. (1) and (2), and the results are presented in Fig. 9, while the percentage errors between the calculated results and the cable test results are tabulated in Table 3. As evident from Fig. 9, the pore pressures of the six wells exhibit a relatively stable fluctuating tendency in the majority of the well sections, and ρ_p is mainly distributed within the range of 0.90–1.10 g/cm^3. However, near 4000 m in X1, X2, X4 and X5, and near 4450 m in X3, ρ_p shows a significant decreasing trend. This phenomenon can be attributed to the presence of geologically discontinuous interfaces in the strata at this depth, such as changes in stratigraphic lithology, variations in stratigraphic density or changes in the depositional environment, resulting in a decrease in pore pressure. In addition, by integrating the measured values of pore pressure obtained by WFT for X1 and X3, it can be seen that there is an excellent agreement between the calculated results and the measured data, and the percentage errors of several points are less than 5%, indicating that the adopted method can more accurately reflect the actual variation of pore pressure of the formation.

4.2. Feature selection

In order to demonstrate the process of feature selection, this study selected the mud log data from wells X2 to X6 and their corresponding pore pressure as the training set, and the data from Well X1 as the test set for model training and testing. To select the optimal combination of input features for pore pressure prediction, we first analyzed the distribution of features in the training and test sets, classified the data into source and target domains using the random forest classification model, and calculated the degree of contribution of each feature to the classification results. By ranking the importance of each feature in descending order and successively removing individual features, this study used the XGBoost model to conduct ablation experiments to evaluate the influence of different features on the predictive performance of the model. The results presented in Fig. 10 visually disclose the trend of the influence of the feature removal process on the model accuracy and the significance level of each feature. In the feature importance bar plot, the degree of influence of PFO, RPM, and HKL is particularly prominent. Meanwhile, the following line graph shows that as these features with large distribution differences are successively removed, the R^2 of the model gradually increases and reaches the optimum after the removal of PFO, RPM, and HKL. From this experimental procedure, it can be observed that when the features exhibit significant outliers or deviations in the data distribution, they tend to introduce significant amounts of noise or bias during the training process, thereby reducing the overall fitness and generalization ability of the model. When these more intrusive features are eliminated, the predictive accuracy of the model shows a significant improvement. It is noteworthy that the predictive performance of the model decreases rather than increases as other features are further removed, suggesting that continued elimination beyond the optimal subset leads to a loss of useful information, resulting in a decrease in model performance.

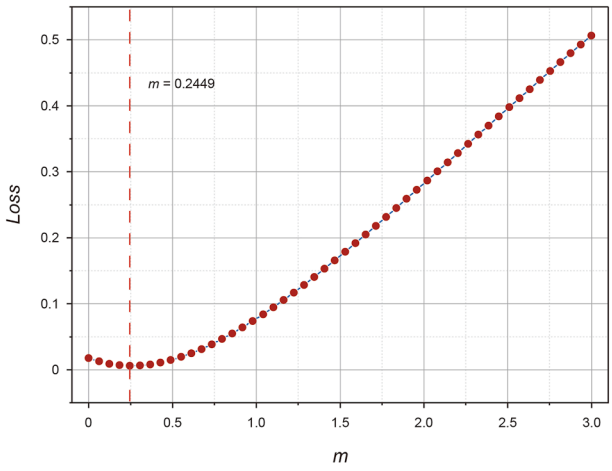


Fig. 8. Effect of the m on the calculation error.

For a further analysis of the features subsequent to the initial screening, in this study the SHAP value of each feature in the pore pressure regression prediction task is calculated using the XGBoost model to quantify the extent of the marginal contribution of diverse features to the model output. As is evident from Fig. 11, R-ROP and WOB possess high contribution values in the regression prediction, suggesting that these two features have remarkable discriminatory capacity in capturing variations in pore pressure. However, the contribution values of HH and SPP are relatively low, indicating that their interpretability and influence are more limited within the framework of the existing model.

To optimize the optimal combination of features, this study subsequently conducted incremental experiments. Specifically, features were sequentially added to the model in accordance with the order of SHAP values from largest to smallest. Each time a new feature was added, the trend of the model's R^2 was recorded, and the results are presented in Fig. 12. The results show that as the number of features increases, the R^2 of the model shows a general tendency to initially increase, then decrease, and finally stabilize. This pattern implies that the addition of new highly contributing features within a certain range can improve the model's ability to learn from the data; however, the introduction of features that are not strongly correlated with the target variables or that are noisy can disrupt the model training process and lead to a deterioration of the prediction performance. It is noteworthy that the inclusion of T causes a significant reduction in the R^2 metric, while the elimination of T leads to a subsequent recovery of R^2, further confirming the strong negative influence of this feature on the predictive performance of the model. Similarly, the addition of HH also decreases the predictive accuracy of the model. Based on the aforementioned experimental results, in this study, T and HH are ultimately excluded from the feature set on the premise of ensuring no loss of critical predictive information, and R-ROP, WOB, TG, PT, MW, PFI, TVD, H, BR, BT, and SPP are considered as the optimal feature combinations for subsequent model construction and validation.

4.3. Model optimization

4.3.1. Base learner optimization

To optimize the base learner of the stacking model, the above 11 models are trained and validated by 5-fold cross-validation on the training set. It is worth noting that in order to avoid the influence of model parameters, the default parameters in the conventional machine learning model are used here, and the neural network model uses 1 hidden layer with the number of neurons being 32. The rest of the hyperparameters remain unchanged.

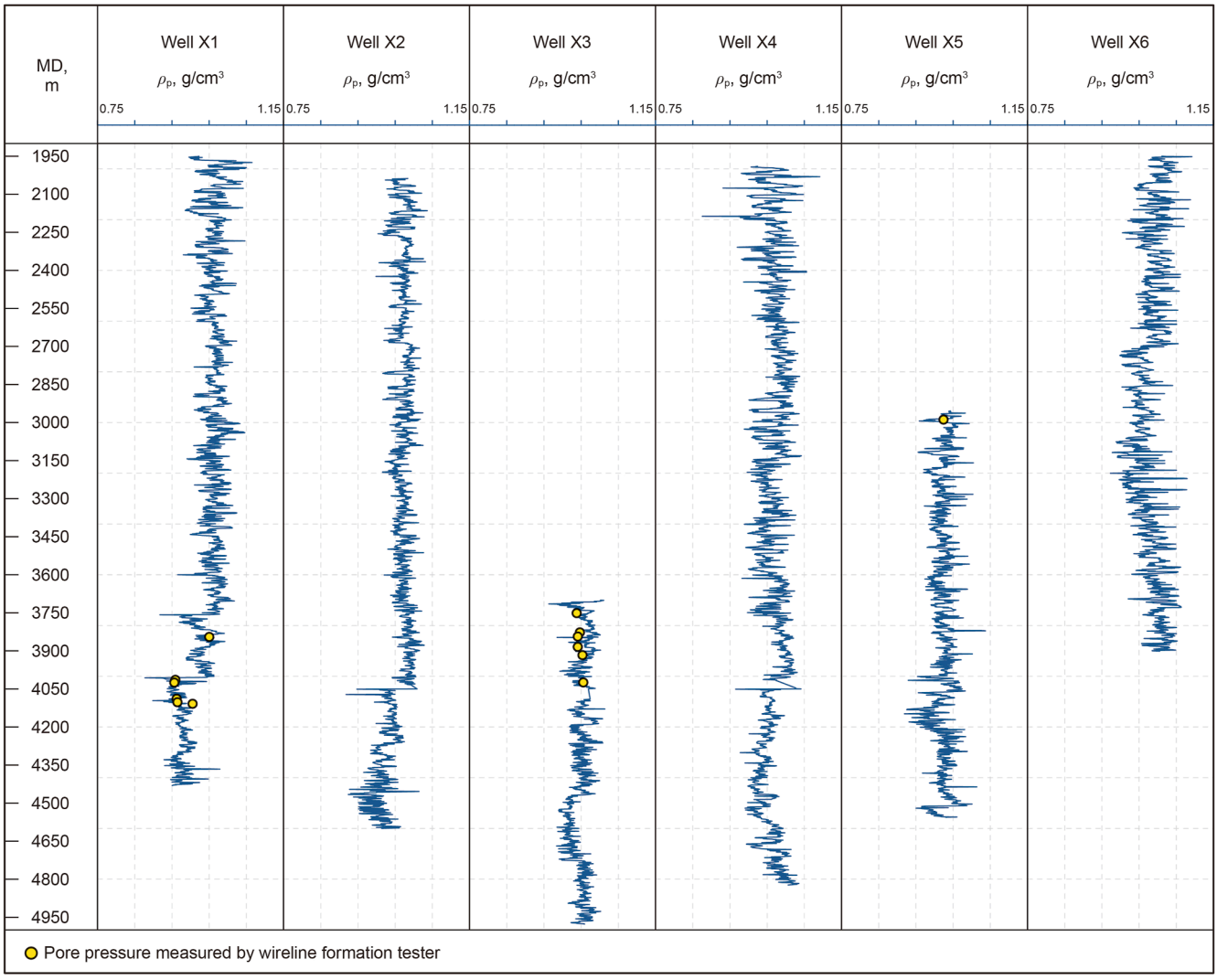


Fig. 9. Results of estimated pore pressure.

Table 3
Verification of estimated pore pressure.

Number	MD, m	WFT, g/cm^3	Estimation, g/cm^3	APE, %
1	2986	0.9793	0.9497	3.02
2	3755	0.9850	0.9936	0.87
3	3825	0.9882	1.0091	2.11
4	3839	0.9865	1.0066	2.04
5	3847	0.9972	0.9918	0.54
6	3885	0.9859	0.9977	1.20
7	3910	0.9984	1.014	1.56
8	4015	0.9209	0.9093	1.26
9	4017	0.9208	0.9198	0.11
10	4021	0.998	0.9819	1.61
11	4088	0.9281	0.9113	1.81
12	4091	0.9282	0.9198	0.90
13	4099	0.9424	0.9079	3.66

During the study, it was found that the prediction results of the ElasticNet and Lasso models deviated significantly from the actual values, resulting in negative R^2 values, indicating that they were unable to effectively capture the strong nonlinear relationship between mud log data and pore pressure gradient, and were therefore excluded from further analysis.

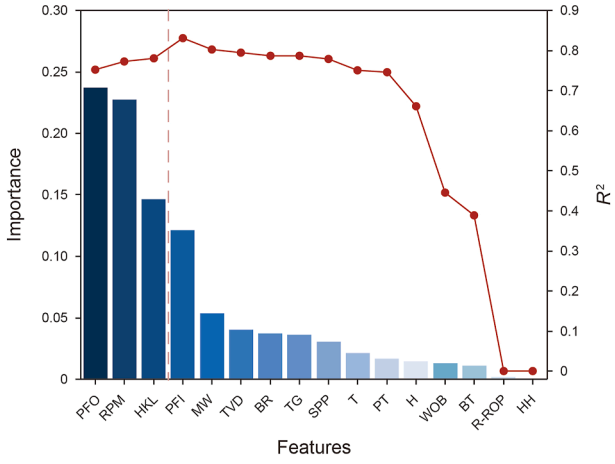


Fig. 10. Feature importance ranking and ablation experiment results.

On this basis, the error scatter density plots of the predicted and actual values of the remaining nine models in the validation set are first presented, as shown in Fig. 13. In addition, the

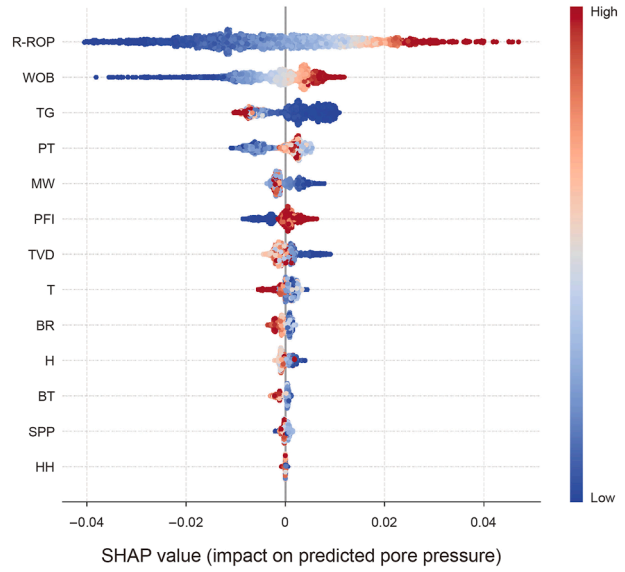


Fig. 11. Feature SHAP value ranking.

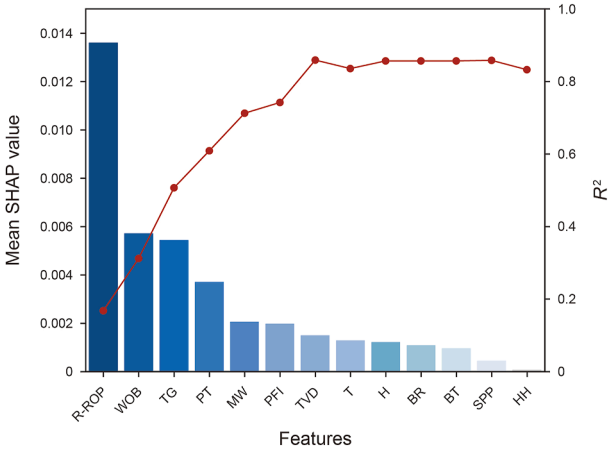


Fig. 12. Feature increment experiment.

evaluation metrics of the prediction results of the validation set were calculated to analyze the prediction performance of the model, and the results are shown in Fig. 14. Combining Figs. 13 and 14, we can see that the fitting results of Ridge and DT are scattered, indicating that the error between the predicted results and the actual values is large, and the fitting accuracy is poor. The prediction effects of RF, AdaBoost, XGBoost and LightGBM models, which optimize the tree model by the ensemble algorithm, are significantly improved, especially the prediction accuracy of XGBoost and LightGBM models is obviously due to other machine learning models. Meanwhile, neural network models such as BPNN, CNN and LSTM generally show high and relatively close prediction performance, indicating that deep learning methods have significant advantages in learning nonlinear relationships between data. Considering the model prediction accuracy, RF, AdaBoost, XGBoost, LightGBM, BPNN, CNN and LSTM with R^2 higher than 0.8 are selected as candidate base learners to further analyze the diversity of model prediction results.

On this basis, to further leverage the diverse and heterogeneous traits of the learners, absolute error correlation analysis is employed for the secondary screening of the aforementioned candidate models, and the correlation results are presented in Fig. 15. In general, the prediction errors of the models show a certain degree of correlation, which is partly due to the non-eliminable noise or bias in the training data, so that different models still show partly similar error patterns in their predictions. In particular, the prediction error correlation of the four tree models is high, reflecting their common decision tree principle and similar ensembles, while the error correlation of the three neural network models is relatively low, indicating that they differ significantly in their network structures and learning mechanisms. BPNN is a typical feed-forward neural network that relies on the back-propagation algorithm to update weights layer by layer and is suitable for handling various types of nonlinear mapping problems; CNN captures local spatial features with a convolutional kernel and excels in multidimensional data analysis; LSTM models the time-dependent relationship in sequence data through a gating mechanism, which is suitable for predicting dynamic data such as time series or deep sequences.

In the stacking model, if the prediction errors of multiple base learners are too similar, it becomes difficult for the meta-learner to extract additional effective features from repetitive information, thereby weakening the performance improvement and increasing the risk of overfitting. Therefore, in this study, a correlation coefficient lower than 0.4 is considered a weak correlation and is used as a criterion for diversifying the base learners. By combining the error correlation with the model's prediction performance on the validation set, LightGBM is ultimately preferred among the four tree models and used as a base learner along with BPNN, CNN, and LSTM for the construction of the subsequent stacking model.

To investigate the impact of the number of base learners on the predictive performance of the stacking model, this study selected Ridge as the meta-learner and conducted a series of ablation experiments on different combinations of base learners. Given that the stacking ensemble strategy requires at least two base learners, the experiments incrementally added additional base learners and assessed the changes in model performance for each combination on the validation set. All base learners were trained using the same dataset, and the stacking model employed a 5-fold cross-validation strategy to generate intermediate features, ensuring no information leakage during the training process.

Table 4 illustrates the primary evaluation metrics for the three groups of base learner combinations on the validation set. The results indicate that as the number of base learners increases, the performance indicators of the stacking model exhibit a consistent optimization trend. When only two base learners are used, the dimensionality of the meta-features is insufficient, which limits the meta-learner's ability to learn an effective combination strategy. This leads to underfitting and even inferior performance compared to single models such as CNN or LSTM. As the number of base learners increases from two to four, the MAE decreases from 0.0060 to 0.0043 g/cm^3, the RMSE decreases from 0.0076 to 0.0057 g/cm^3, the MAPE decreases from 0.6130% to 0.4391%, and R^2 improves to 0.9409. These findings demonstrate that moderately increasing the number of base learners significantly enhances both the prediction accuracy and stability of the stacking model. This improvement can be attributed to the complementary nature of different base models in terms of error distribution, which strengthens the overall model's capacity for information representation. By integrating diverse intermediate predictions, the meta-learner achieves stronger generalization capabilities. Furthermore, the collaboration among multiple models reduces bias while cross-validation-generated meta-features control variance growth, ensuring robustness alongside improved prediction accuracy. In conclusion, the experimental results confirm that combining a diverse and sufficient number of base learners effectively boosts the predictive performance of the stacking

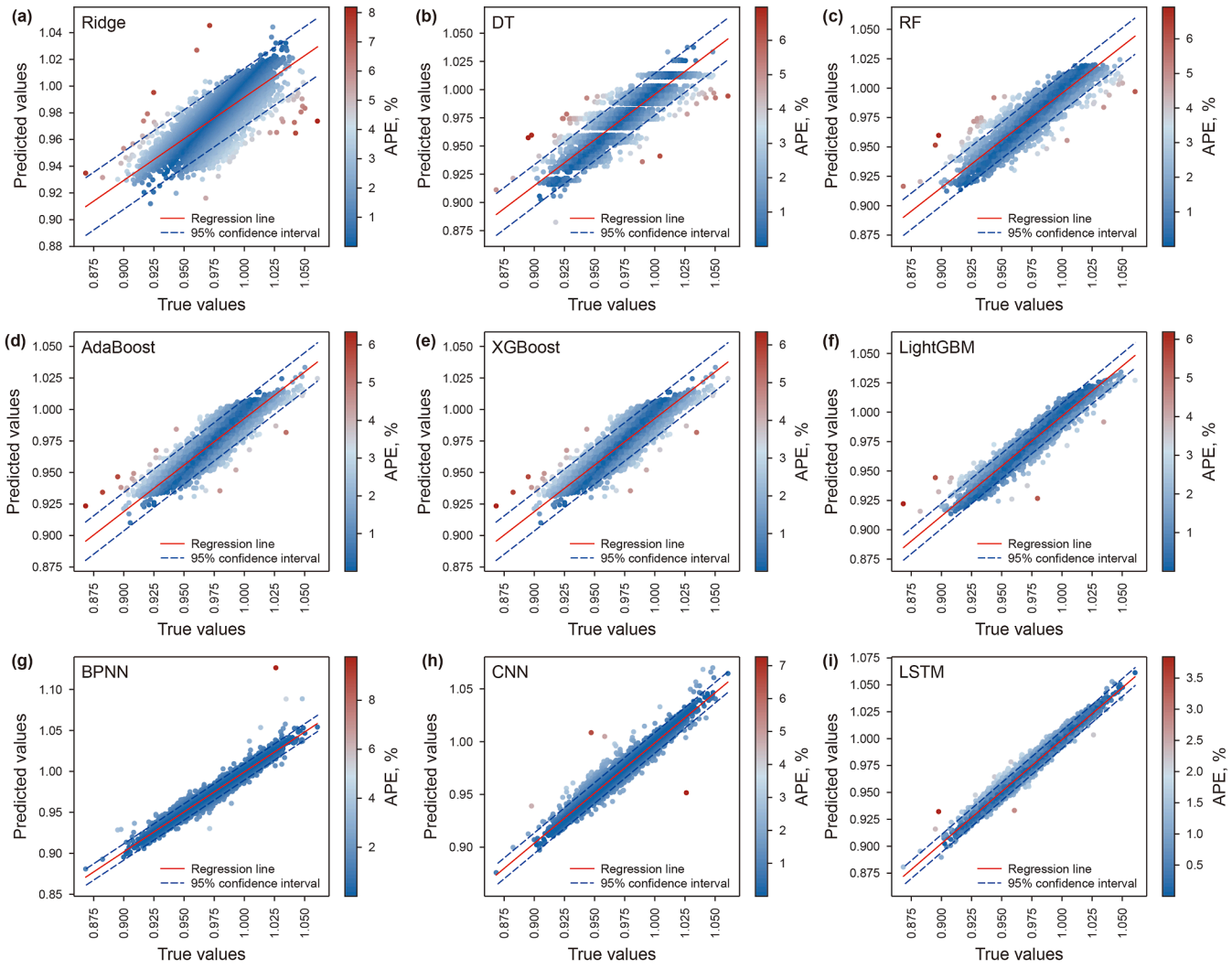


Fig. 13. Scatter density plot of error in prediction results.

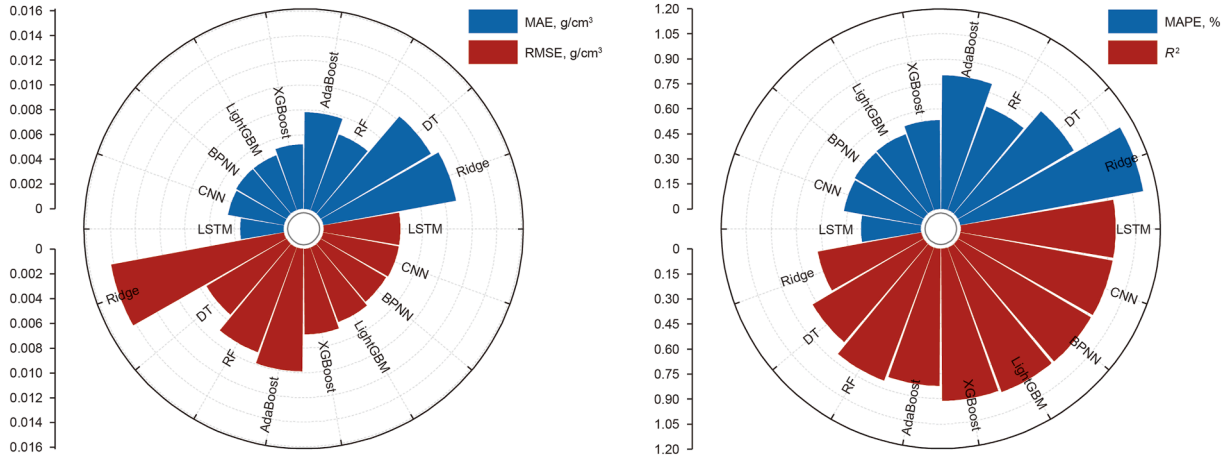


Fig. 14. Base learner evaluation index.

model. In this study, the combination of four highly complementary base learners (LSTM, CNN, BPNN, LightGBM) achieved the best performance across multiple evaluation metrics and is recommended for future research.

4.3.2. *Meta learner optimization*

In the stacking model, the meta learner is responsible for secondary learning or fusion of the prediction results from each base learner, and different meta learners have a significant impact on

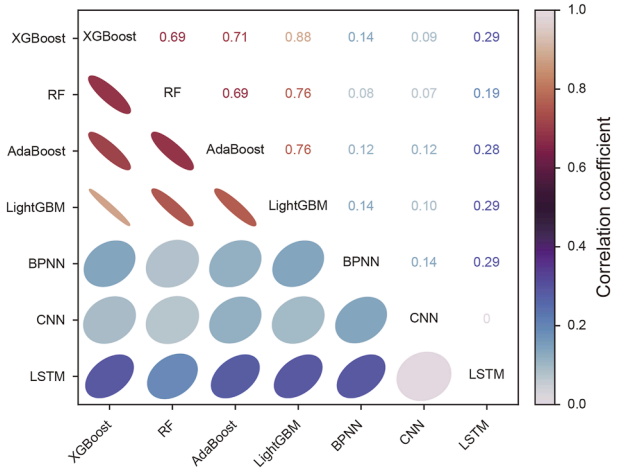


Fig. 15. Base learner error correlation.

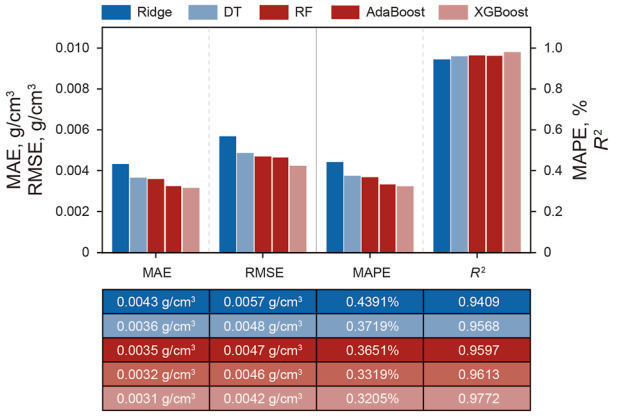


Fig. 16. Evaluation index of different meta learner.

Table 4
Validation performance of stacking models with varying base learner combinations.

Base learners	MAE, g/cm^3	RMSE, g/cm^3	MAPE, %	R^2
LSTM, CNN	0.0060	0.0076	0.6130	0.8942
LSTM, CNN, BPNN	0.0046	0.0059	0.4769	0.9352
LSTM, CNN, BPNN, LightGBM	0.0043	0.0057	0.4391	0.9409

the overall model performance. To select the optimal meta-learners, five previously unused models, namely Ridge, XGBoost, DT, RF, and AdaBoost, were used as meta-learners in this study, and secondary learning was performed on the identical prediction results of the base learners. By combining the prediction results of each base learner in Fig. 14 and the evaluation metrics of the prediction results of the stacking model under different meta-learner in Fig. 16, it can be observed that the prediction performance of the stacking model has been significantly enhanced under the four meta-learners in comparison to each base learner. Among them, there is still a disparity in the prediction accuracy of the Ridge model in contrast to the more complex tree model and ensemble model, which indicates that the linear model excels in capturing the overall trend, but is still marginally inadequate in representing non-linear relationships or complex feature interactions. Compared to the DT with a simple structure, the prediction accuracy of the RF with the bagging strategy, the AdaBoost, and the XGBoost with boosting are all improved. And XGBoost achieves optimal performance in all four metrics due to its collaborative implementation of the gradient boosting tree algorithm with a regularization mechanism.

4.3.3. Hyperparameter optimization

The hyperparameters of the machine learning model play a critical role in the prediction performance of the model. The stacking model includes multiple base learners and a meta learner. Manually adjusting the parameters by hand would consume a considerable amount of time and energy, and may not consistently produce superior results. For this reason, this study uses the Bayesian algorithm to optimize the key parameters of the base and meta learners of the stacking model. The hyperparameter ranges to be optimized for each learner, the final preferred results, and the corresponding RMSE are shown in Table 5.

As shown in Fig. 17(a), before optimization, the predicted values exhibit a relatively scattered distribution around the diagonal reference line, with several points falling outside the 95% confidence interval, indicating deviations between predictions and ground truth. The corresponding residual histogram in Fig. 17(b) reveals a wider distribution with larger deviations from zero, and the KDE curve in Fig. 17(c) demonstrates a visible discrepancy between the distributions of predicted and true values, suggesting systematic bias.

In contrast, Fig. 18 presents the model performance after Bayesian hyperparameter tuning. Fig. 18(a) illustrates that the predicted values are more densely distributed along the diagonal line, with most data points constrained within the 95% confidence band. The residual distribution in Fig. 18(b) is notably narrower and more symmetric, centering tightly around zero, while the KDE curves in Fig. 18(c) show near-complete overlap between predicted and actual values, reflecting the improvement in prediction accuracy and distribution consistency.

These visual results are supported by the quantitative metrics in Table 6. Specifically, after optimization, the MAE and RMSE were significantly reduced from 0.0031 and 0.0042 g/cm^3 to 0.0009 and 0.0016 g/cm^3, respectively. Similarly, the MAPE decreased from 0.3205% to 0.0931%, and the R^2 value increased from 0.9672 to 0.9955. These improvements confirm that Bayesian optimization effectively enhanced the model's fitting ability and generalization performance. In summary, both the graphical and statistical evidence demonstrate that Bayesian hyperparameter optimization substantially improves the predictive performance and reliability of the stacking model.

4.4. Pore pressure prediction results

To further verify the prediction accuracy and generalization ability of the model, this study simulated three scenarios that might occur in on-site drilling operations. Scenario 1 assumes that the well being drilled is located in the center of the block and there are multiple completed wells around it; Scenario 2 assumes that there is only one offset well near the well being drilled, and the other wells are far away; Scenario 3 assumes that the offset wells are all far away from the positive drilling. According to the wellbore trajectories of the six wells shown in Fig. 2, test wells conforming to the above three scenarios are selected. Table 7 shows the training and testing wells used by the model in the three scenarios.

4.4.1. Scenario 1: X1 as the test well

As noted above, Scenario 1 simulates drilling conditions where the well to be drilled lies in the center of the block and numerous

Table 5
Hyperparameter optimization results.

Model	Key parameter	Parameter range	Best iteration	Parameter result	RMSE, g/cm^3
CNN	filters	[16, 64]	38	30	0.0022
	kernel_size	[3,7]		4	
	dense_units	[8, 32]		32	
	learning rate	[0.0001, 0.01]		0.0035	
	batch_size	[8, 64]		16	
	epochs	[50, 300]		298	
LSTM	lstm_units	[32, 128]		85	0.0013
	dense_units	[16, 64]		58	
	learning rate	[0.0001, 0.01]	16	0.003	
	batch_size	[8, 64]		16	
	epochs	[50, 300]		292	
BPNN	dense_units_1	[32, 128]	29	43	0.0023
	dense_units_2	[8, 32]		22	
	learning rate	[0.0001, 0.01]		0.001	
	batch_size	[8, 64]		16	
	epochs	[50, 300]		227	
LightGBM	max_depth	[3, 10]	33	8	0.0023
	learning_rate	[0.01, 0.3]		0.12	
	n_estimators	[50, 200]		189	
XGBoost	max_depth	[1, 10]	47	3	0.0018
	learning_rate	[0.01, 0.3]		0.06	
	n_estimators	[50, 200]		194	

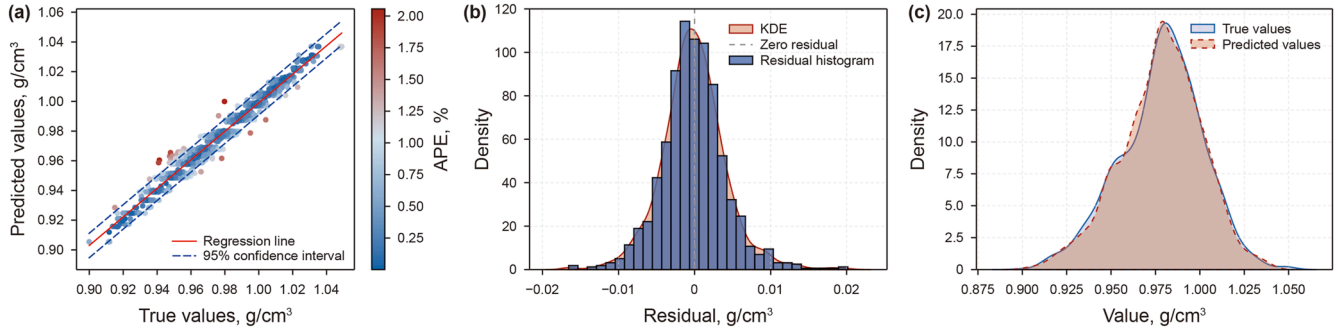


Fig. 17. Visualization of stacking model prediction and error distribution before optimization. (**a**) Scatter density of predicted vs. true values with 95% confidence band. (**b**) Residual histogram with kernel density estimation (KDE). (**c**) KDE comparison between predicted and true values.

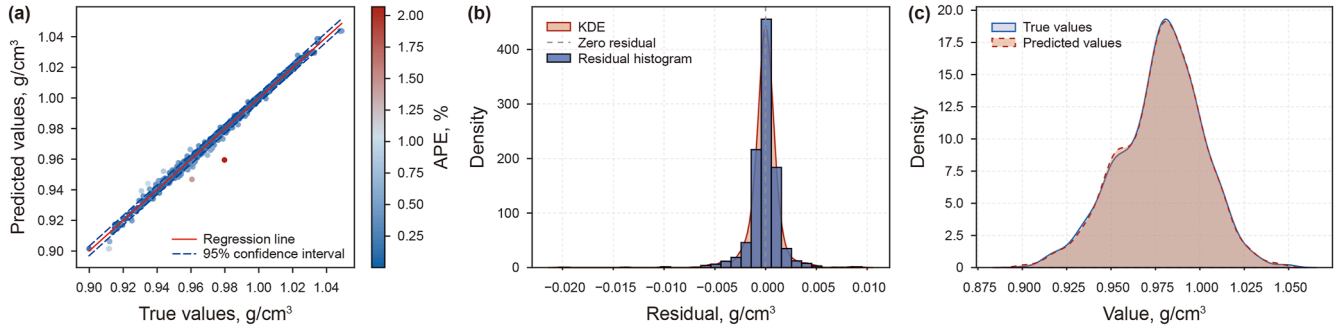


Fig. 18. Visualization of stacking model prediction and error distribution after optimization. (**a**) Scatter density of predicted vs. true values with 95% confidence band. (**b**) Residual histogram with KDE. (**c**) KDE comparison between predicted and true values.

Table 6
Prediction results evaluation index of stacking and Bayesian-stacking model.

Model	MAE, g/cm^3	RMSE, g/cm^3	MAPE, %	R^2
Before optimization	0.0031	0.0042	0.3205	0.9672
After optimization	0.0009	0.0016	0.0931	0.9955

completed offset wells exist around it. Fig. 19 shows the curve of predicted values versus estimated values for Well X1, the corresponding scatter plot, and the prediction result evaluation indicators. As can be seen from Fig. 19(a), it is evident that the method proposed in this study is capable of accurately predicting the pore pressure profile of the well under Scenario 1. The

predicted curve is essentially the same as the actual curve, and the error between the predicted values and the WFT measurement points is relatively small. The APE curve shows that the overall APE between predicted and actual values is less than 1%, with only a few depths showing larger errors, often accompanied by significant variations in pore pressure. From Fig. 19(b), it is easy to see that the fitted scatter points are concentrated near the regression line, and the vast majority lie within the 95% confidence interval, with only a few points deviating, and the maximum APE is only 3.85%. Combined with the evaluation indicators (Fig. 19(c)), the MAPE of the prediction results is only 0.4353%, and the R^2 is as high as 0.9362, indicating that the model has effectively captured the variation law of the pore pressure of Well X1 using mud log data and has achieved accurate prediction of the pore pressure of Well X1. The results indicate that the model can produce more favorable prediction results when there are a significant number of offset wells around the target well. This phenomenon occurs because the geological conditions and drilling environments of the offset wells are relatively analogous to those of the target well, resulting in a slight distributional disparity between the training set data and the test set data. As a result, the model can effectively generalize the pore pressure variation law learned from the mud log data in the training set to the test set. However, when the number of offset wells is limited or the wells drilled are far from the target well, the prediction effect may be insufficient. Therefore, in this study, the predictive performance of the model was further analyzed under scenarios 2 and 3.

4.4.2. Scenario 2: X5 as the test well

Fig. 20 shows the prediction results of the model for Well X5 under Scenario 2. In contrast to Scenario 1, the prediction accuracy of the model has deteriorated. This phenomenon is indicated in the profile plot by a decrease in the degree of fit between the predicted profile and the actual profile, along with an overall increase in the APE. The scatter plot shows that the number of fitted points that deviate from the regression line has increased. Although the maximum APE has decreased from 3.85% in Scenario 1–1.86%, the total number of points with larger errors has increased, as indicated by the increase in the number of red points. In addition, the evaluation metrics more intuitively illustrate the decline in the model's predictive performance. The MAE of the prediction results increased from 0.0042 to 0.0046 g/cm^3, the RMSE increased from 0.0064 to 0.0068 g/cm^3, the MAPE increased from 0.4353% to 0.4646%, and the R^2 decreased from 0.9362 to 0.9078. The above results suggest that despite the consistent number of training wells, the reduction in the number of offset wells closer to the target well has resulted in an increased data distribution disparity between the training and test sets, causing the model to be biased to some extent toward the patterns between mud log data and pore pressure of completed wells farther from the target well in the training set, resulting in a deterioration in the model's prediction performance on the test wells.

4.4.3. Scenario 3: X3 as the test well

In contrast to scenarios 1 and 2, Scenario 3 simulates the drilling process with Well X3 as the test well, where there are no offset wells around the well to be drilled, further increasing the requirements for the generalizability of the model. The prediction results for Well X3 are shown in Fig. 21. It can be seen that although the predicted pore pressure profile has a similar variation tendency to the actual one, there is a significant offset in the predicted values, resulting in an overall increase in the APE. The confidence interval of the fitted scatterplot is significantly wider compared to scenarios 1 and 2, indicating that the disparity and dispersion between the predicted and actual values is greater and the fitting effect is worse. From the evaluation indicators, the MAE of the prediction results has increased to 0.0068 g/cm^3, showing an increase of 61.90% compared to Scenario 1 and 47.83% compared to Scenario 2. The RMSE has increased to 0.0081 g/cm^3, showing an increase of 31.25% compared to Scenario 1 and 19.12% compared to Scenario 2. The MAPE has increased to 0.6856%, showing an increase of 57.50% over Scenario 1 and 47.57% over Scenario 2. The R^2 has decreased to 0.8950, reflecting a decrease of 4.40% over Scenario 1 and 1.41% over Scenario 2. It can be observed that unlike Scenarios 1 and 2, Scenario 3 simulates the actual operating condition where there are no offset wells around the target well. This leads to a lack of reference data from offset wells during the model's prediction process, resulting in significant discrepancies between the data from Well X3 and the training set data. In the absence of sufficient information from offset wells, the model has difficulty effectively learning the applicable patterns for Well X3, resulting in an overall bias and increased error during prediction.

Table 7
Training and test wells used for the 3 scenarios.

Scenario	Train wells	Test well
Scenario 1	X2, X3, X4, X5, X6	X1
Scenario 2	X1, X2, X3, X4, X6	X5
Scenario 3	X1, X2, X4, X5, X6	X3

4.4.4. SHAP analysis

To identify the critical drivers in the pore pressure prediction model, the SHAP method was employed to interpret feature contributions across 3 different scenarios. As illustrated in Fig. 22, despite variations in the training and testing datasets across these scenarios, the features with the highest SHAP values remain remarkably consistent. Specifically, R-ROP, MW, H, and WOB consistently emerge as the most influential features. These 4 mud log parameters are directly associated with pore pressure, and their high SHAP values across scenarios indicate robust and significant marginal effects, confirming their role as universally essential predictors in different operational contexts.

Among them, MW serves as the most direct operational control for managing the mechanical balance between borehole pressure and pore pressure. It is dynamically adjusted in real time to maintain pressure equilibrium, underscoring a strong engineering causality between MW and pore pressure. The model's sensitivity to MW is thus fully aligned with fundamental principles of wellbore pressure control. R-ROP reflects formation drillability and compaction levels, acting as a proxy for lithological hardness and pore closure. Its relationship with pore pressure is non-linear yet stable, allowing the model to detect abrupt lithological transitions or dynamic compaction changes, particularly in transitional geological zones. Depth (H) inherently influences pore pressure through mechanisms such as gravitational compaction, overburden confinement, and fluid migration, which collectively control the evolution of subsurface pressure regimes. WOB, a key mechanical input parameter, modulates the bit-formation interaction. In high pore pressure formations, increased formation resistance necessitates greater WOB to sustain penetration rates, whereas in softer or under-pressured formations, lower WOB suffices. While the top-ranked features remain consistent, notable differences are observed in the relative contributions of lower-ranked features such as BR, SPP, and BT across scenarios. These discrepancies likely arise from variations in the covariance structures of input variables, which may lead the model to exploit different synergistic interactions depending on the specific

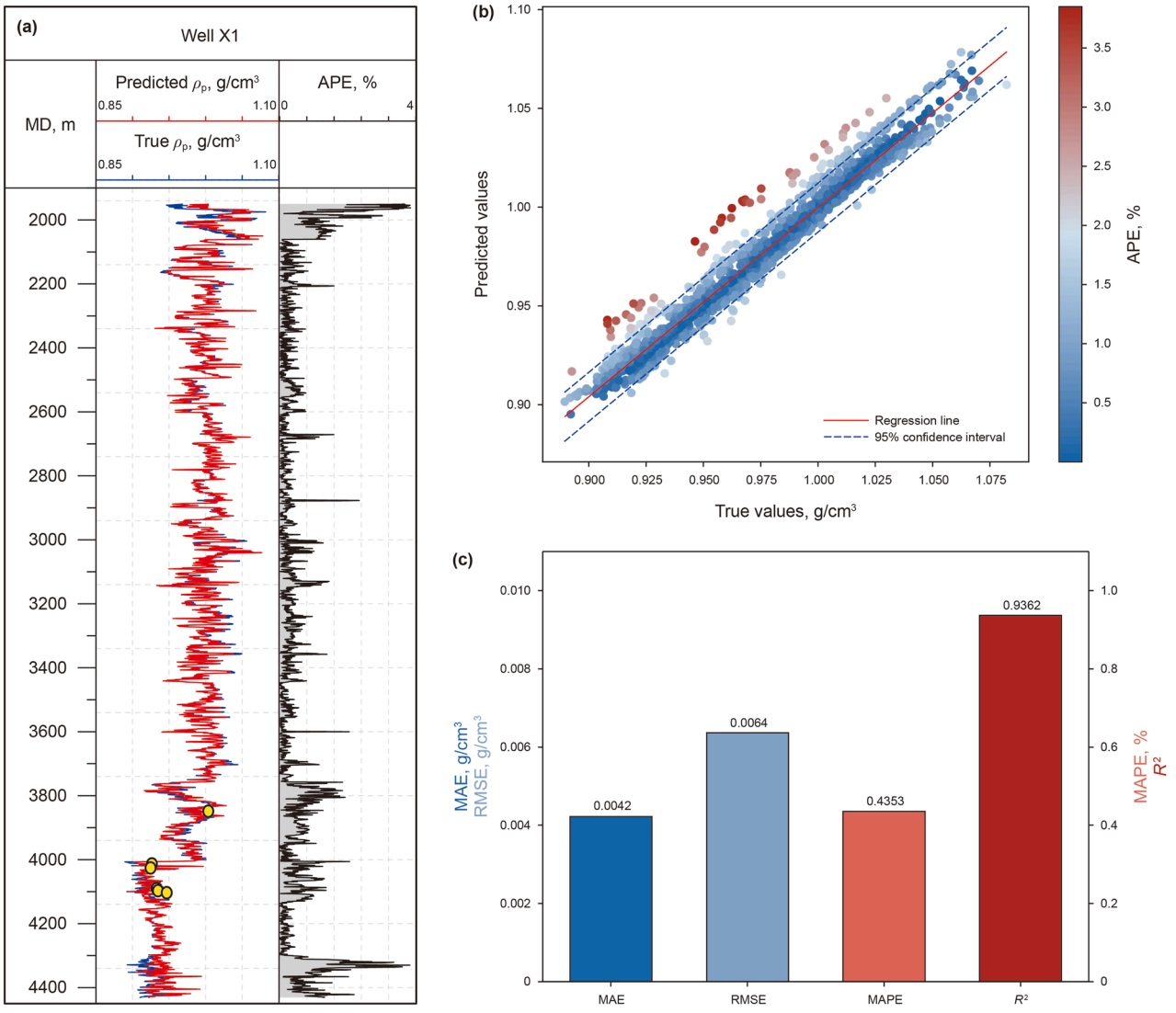


Fig. 19. Prediction results for Scenario 1.

dataset. Furthermore, disparities in the numerical distributions and sampling densities of these features can alter the model's reliance on them during the learning process, resulting in scenario-specific feature importance profiles.

5. Discussions

Real-time pore pressure prediction allows drilling engineers to acquire an early understanding of formation characteristics and drilling conditions at the drill bit. This makes it easier to dynamically adjust drilling parameters to increase drilling efficiency and reduce associated risks. In this study, the isolation forest and wavelet threshold denoising techniques were used to preprocess the original data. To determine the globally optimal Eaton index, the inversion of the Eaton index was performed based on pore pressure data obtained from WFT in combination with the Bayesian optimization algorithm. Subsequently, pore pressure data that were highly consistent with the measured pressures were derived using the d_c exponent method, which provided the data set for machine learning training. In addition, the optimal input parameter combinations were identified by analyzing data distribution characteristics and performing importance analysis. To further improve the model performance, we select BPNN, CNN, LSTM and LightGBM as the base learners, and also use XGBoost as the meta learner to construct the stacking model, and optimize the hyperparameters of the model using Bayesian algorithm. The optimized model was subsequently employed to predict pore pressure in three wells under diverse scenarios. The model exhibited high prediction accuracy and low error, thereby validating the superior prediction performance and robust generalization capacity of the proposed methodology.

In previous studies, most pore pressure predictions based on mud log data involve the subsequent division of the entire data set into a training set and a test set to ensure that the data follow independent and homogeneous distributions and to avoid model prediction bias on the test set. Although this method provides high prediction accuracy, it is unable to capture the trend of pore pressure with depth and does not ensure accurate predictions for new data. Recently, Chen et al. (2022) used a partial least squares regression engine to accurately predict pore pressure continuum profiles by using drilling parameters such as ROP, T, RPM, and pore pressure calculated by the d_c exponent, sigma, and HRSE methods, respectively, as input features. However, this method relies on pore pressures calculated by three empirical methods as intermediate

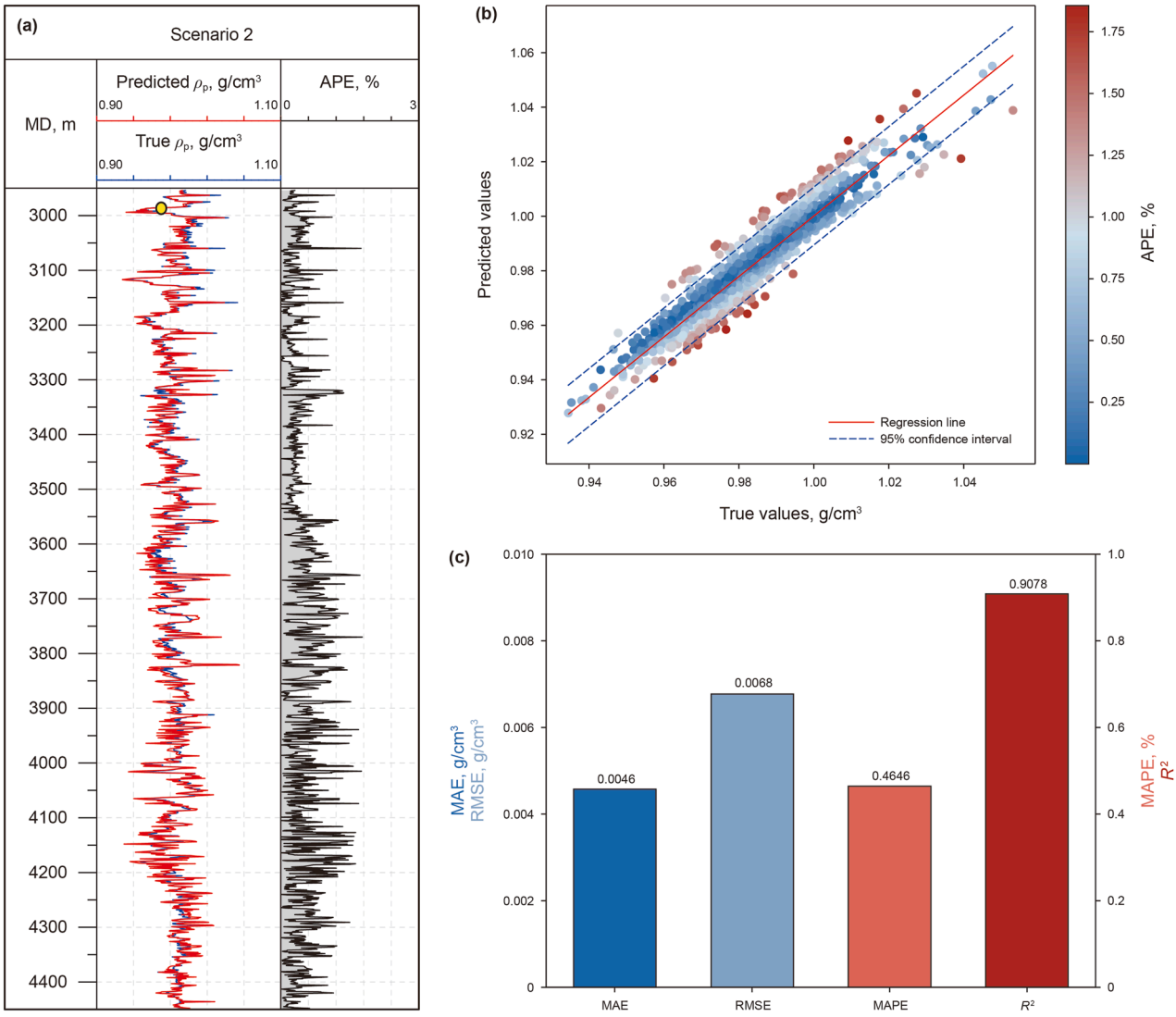


Fig. 20. Prediction results for Scenario 2.

variables, and the calculation error of the empirical formula itself further aggravates the propagation error of the model. In addition, the calculation of intermediate parameters also makes the whole prediction process more complicated. In contrast, this study only uses conventional mud log data without additional calculation of other intermediate parameters, which simplifies the pore pressure prediction process and reduces the propagation error of the model, making the prediction results more reliable. Yang et al. (2025) proposed a sequence-to-point prediction approach based on the drilling parameters and drilling fluid parameters of 14 wells, using three LSTM network structures to accurately predict the pore pressure at the bit position. After considering factors such as bit type and degree of wear, the predicted R^2 could respectively reach 0.89 and 0.75 in two wells. In contrast to the above studies, accurate pore pressure prediction was achieved during the drilling of all three wells, with a predicted R^2 greater than 0.9, using only conventional mud log data without additional calculation of intermediate parameters in this study.

In addition, although the training process of the proposed stacking model—particularly during hyperparameter optimization—is relatively time-consuming, it is conducted offline. In real-world drilling operations, the key concern is the efficiency of model inference. We evaluated the full inference time of the entire stacking ensemble, including all base learners and the meta-learner, under a standard CPU environment (Intel i5, 32 GB RAM). The total prediction time for the full test dataset was approximately 0.64 ms. Given that drilling data is typically collected at intervals of several seconds, this end-to-end latency is sufficiently low to meet the requirements of real-time or near-real-time analysis. Furthermore, the modular design of the model facilitates seamless integration into existing real-time drilling monitoring systems. While the primary goal of this study is to deliver real-time pore pressure estimation as an advisory tool, the demonstrated low-latency inference capability indicates strong potential for integration into more advanced platforms, such as automated drilling decision support systems, with minimal engineering overhead.

Despite the promising results achieved in this study for formation pore pressure prediction based on mud log data, several limitations must be acknowledged. (1) In this study block, the formation is mainly composed of sand and mudstone, and the d_c exponent method and the Eaton method can provide more accurate pore pressure estimates, providing a reliable label for subsequent model training and optimization. However, the applicability

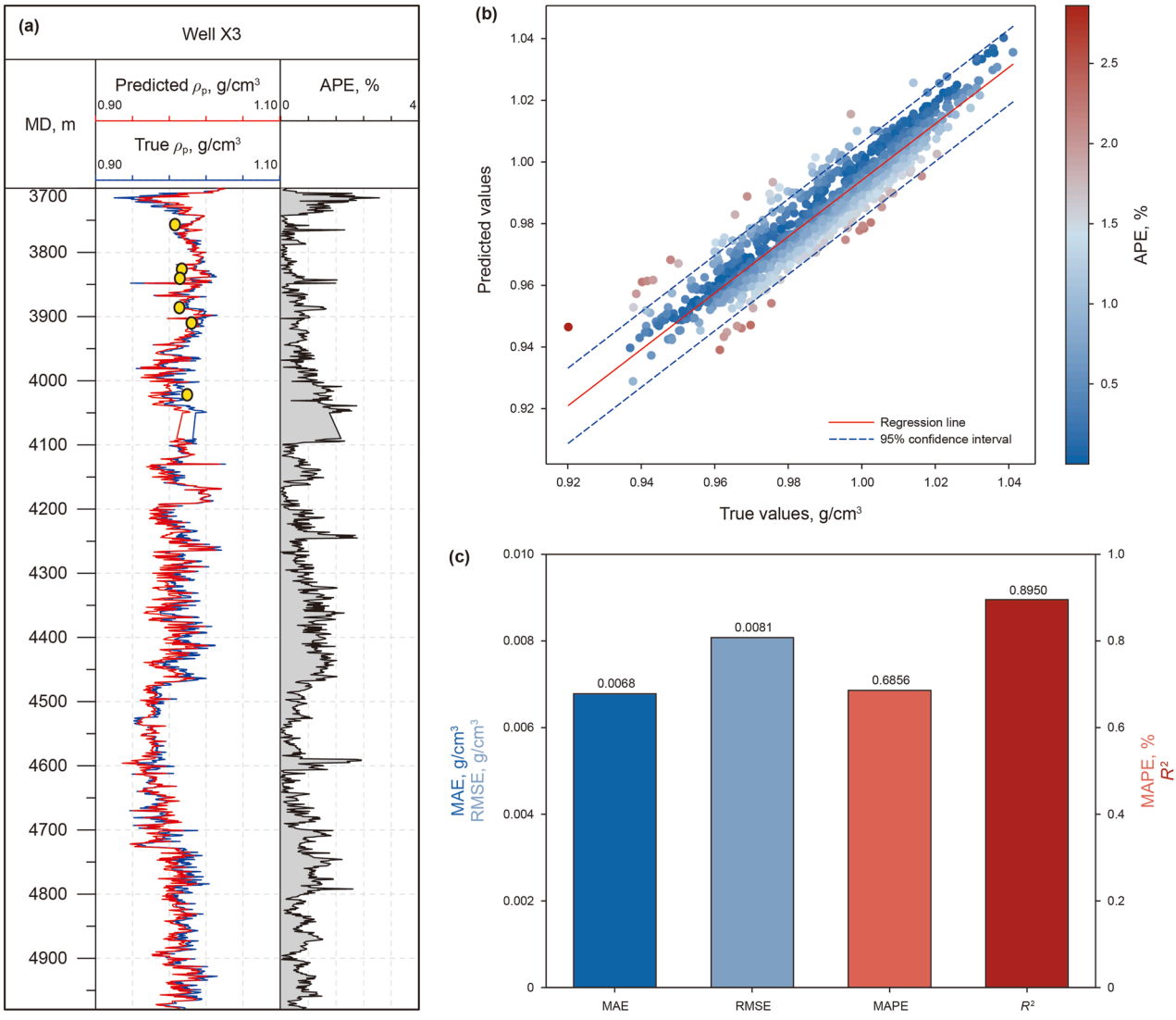
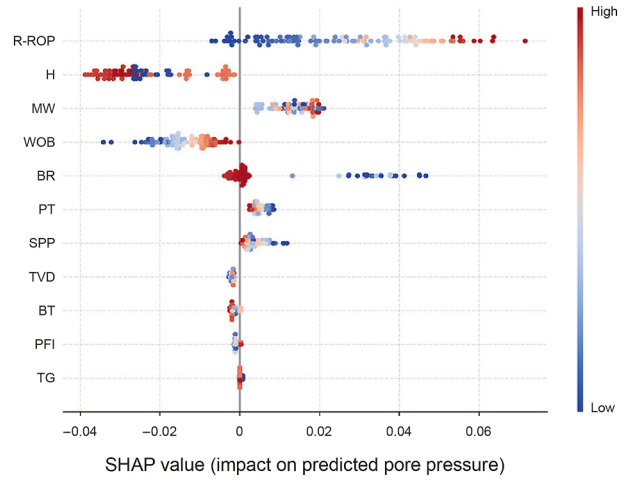


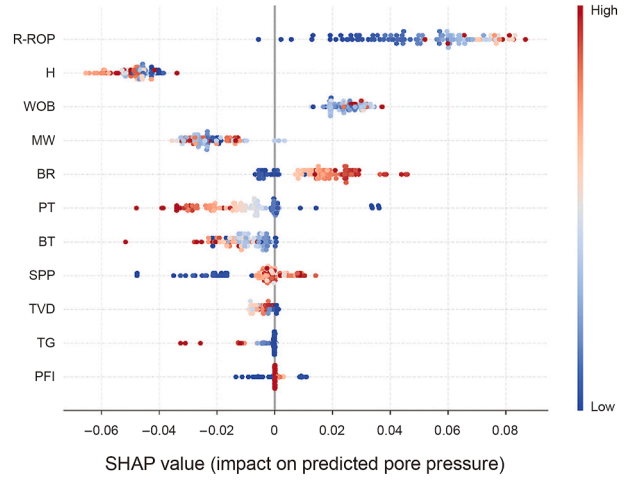
Fig. 21. Prediction results for Scenario 3.

of this method is significantly influenced by the lithology type, and accurate pore pressure labels may not be obtained by using this method in strata with non-classical compaction types, such as carbonate rocks and igneous rocks. (2) The pore pressure labels used in this study were estimation using the d_c exponent method and were not directly measured data, thus having a certain degree of compatibility. Although it is optimized for the inversion of the Eaton index, and is valuable in the context of limited data, it still cannot be fully compatible with all conditions of the block, resulting in limited generalization ability of the model when predicting new wells. (3) In this study, data were collected from a total of 6 wells, and 5 wells were used as the training set each time, which limited the amount of data and thus the accuracy and generalization ability of the model. Although it is shown through validation that the method in this paper can achieve a more accurate pore pressure prediction when using 5 wells as the training set, the prediction may not be satisfactory for new wells with large data deviations. In addition, for other blocks, it is necessary to perform feature screening and model training again using data from that block. (4) Although the stacking model with 4 basic learners achieves a high prediction accuracy, its training process, especially the hyperparameter optimization process, is undoubtedly very time-consuming. (5) This study focuses only on real-time pore pressure prediction during drilling and cannot realize pore pressure prediction before the drill bit. (6) The model's predictive accuracy decreases as the number and spatial proximity of offset wells are reduced. This is particularly evident in Scenario 3, where sparse offset wells result in significant accuracy degradation. This indicates a limitation in the model's cross-well generalization ability.

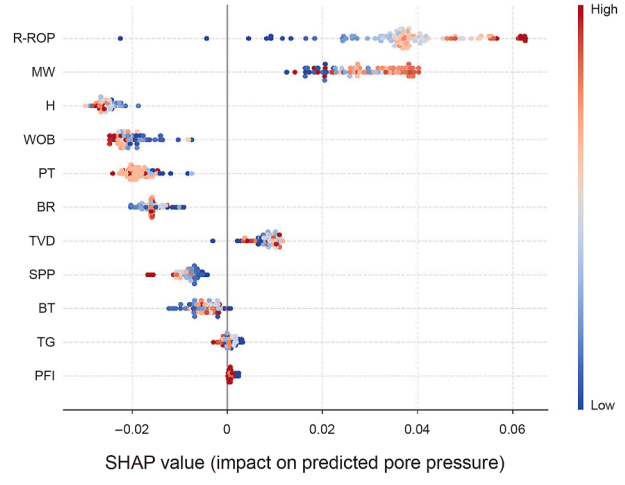
To address these limitations, future work should focus on the following: Firstly, continue to collect data from additional wells within the block to expand the data set in order to increase the learning range of the model and thereby improve the accuracy and generalizability of the model. Secondly, given the limited WFT measurement data, it is impossible to provide more comprehensive calibration support. In future work, uncertainty quantification methods (such as Monte Carlo simulation or Bayesian inference) can be considered to propagate the uncertainties in the input mud log curves and the inversion model, thereby achieving probabilistic distribution prediction of pore pressure and enhancing the credibility and interpretability of the model results. Thirdly, improving robustness under sparse-data conditions by applying domain adaptation and transfer learning techniques. Preliminary



(a) Scenario 1



(b) Scenario 2



(c) Scenario 3

Fig. 22. SHAP values for the 3 different scenarios.

investigations using MMD and adversarial learning showed modest improvement due to current data constraints. Nonetheless, with the expansion of the dataset, these methods offer a promising pathway to enable better cross-well adaptability. Fourthly, attempt to simplify the model and improve the efficiency of model training so that the model can be better applied in the field. Lastly, consider constructing a three-dimensional geological model of the block by integrating seismic attributes to achieve accurate prediction of pre-drilling pore pressure and provide data support for the design of the drilling process.

6. Conclusions

This paper presents a novel machine learning approach for real-time pore pressure prediction based on mud log data, which eliminates the dependence on LWD data and redundant derived parameters. By integrating the predictive capabilities of multiple models through an ensemble learning strategy, the proposed method effectively captures the depth-dependent relationship between mud log data and pore pressure. It enables timely and accurate tracking of pore pressure variations, providing valuable data-driven support to drilling engineers in the field. This, in turn, helps to improve drilling efficiency and reduce operational risks. The key innovations and contributions of this study are summarized as follows:

(1) The raw mud log data were processed using the Isolation Forest algorithm for outlier detection, the wavelet thresholding was employed for the reduction of noise, and the Bayesian inversion approach was then implemented to derive the global optimal Eaton index of the study block. This index was derived based on pore pressure measurements from WFT, and the resultant index value was determined to be 0.2449.

(2) The estimation of pore pressure was accomplished through the integration of the d_c exponent method and the Eaton method. The maximum percentage error in pore pressure estimation was determined to be 3.66%, and the corresponding average error was found to be 1.59% compared to the wireline formation test data.

(3) In the feature selection stage, the optimal combination of input features required for pore pressure prediction was established through the integration of the data distribution characteristics and an analysis of the importance of features in the regression prediction. The results indicated that the optimal combination of input features was determined to be R-ROP, WOB, TG, PT, MW, PFI, TVD, H, BR, BT, and SPP.

(4) A comparative analysis of the prediction performance of nine models on the validation set was conducted, leading to the selection of the BPNN, CNN, LSTM, and LightGBM as base learners. In addition, XGBoost was chosen as a meta learner for the development of the stacking model. The hyperparameters of each learner were optimized using the Bayesian algorithm. The resultant enhancement in the R^2 of the optimized model on the validation set was significant, with an increase from 0.9672 to 0.9955, and a concomitant decrease in the MAPE from 0.3205% to 0.0931%.

(5) The optimized model was utilized to predict the pore pressure of three test wells (X1, X5, X3) under three distinct scenarios. Prediction accuracy is improved when offset wells are densely distributed, spatially balanced, and proximal to the target well; conversely, sparse or distant offset wells result in reduced prediction performance. The MAPE of the prediction results was 0.4353%, 0.4646%, and 0.6856%, respectively, and the R^2 was 0.9362, 0.9078, and 0.8950, respectively.

CRediT authorship contribution statement

Dong-Yang Zhang: Writing – original draft, Visualization, Software, Methodology. **Tian-Shou Ma:** Writing – review &

editing, Supervision, Investigation, Funding acquisition, Conceptualization. **Yang Liu:** Writing – original draft, Validation, Software. **De-Cheng Zhang:** Writing – review & editing, Investigation.

Data availability

Data will be made available on request.

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 work was supported by the National Natural Science Foundation of China (Grant No. 52474010), the Sichuan Science and Technology Program (Grant No. 2025YFHZ0079), and the Program of Introducing Talents of Discipline to Chinese Universities (111 Plan) (Grant No. D18016).

References

Ahmed Abdelaal, A.A., Salaheldin Elkatatny, S.E., Abdulazeez Abdulraheem, A.A., 2021. Pore pressure estimation while drilling using machine learning. In: ARMA/DGS/SEG International Geomechanics Symposium, Virtual.

Ahmed, A., Elkatatny, S., Ali, A., et al., 2019. New model for pore pressure prediction while drilling using artificial neural networks. Arabian J. Sci. Eng. 44 (6), 6079–6088. https://doi.org/10.1007/s13369-018-3574-7.

Alasadi, S.A., Bhaya, W.S., 2017. Review of data preprocessing techniques in data mining. J. Eng. Appl. Sci. 12 (16), 4102–4107. https://doi.org/10.3923/jeasci.2017.4102.4107.

AlBahrani, H., Morita, N., 2022. Risk-controlled wellbore stability criterion based on a machine-learning-assisted finite-element model. SPE Drill. Complet. 37 (1), 38–66. https://doi.org/10.2118/204101-PA.

Alsini, R., Almakrab, A., Ibrahim, A., et al., 2021. Improving the outlier detection method in concrete mix design by combining the isolation forest and local outlier factor. Constr. Build. Mater. 270, 121396. https://doi.org/10.1016/j.conbuildmat.2020.121396.

Azadpour, M., Manaman, N.S., Kadkhodaie-Ilkhchi, A., et al., 2015. Pore pressure prediction and modeling using well-logging data in one of the gas fields in south of Iran. J. Petrol. Sci. Eng. 128, 15–23. https://doi.org/10.1016/j.petrol.2015.02.022.

Bowers, G.L., 1995. Pore pressure estimation from velocity data: accounting for overpressure mechanisms besides undercompaction. SPE Drill. Complet. 10 (2), 89–95. https://doi.org/10.2118/27488-PA.

Carneiro, T.C., Rocha, P.A., Carvalho, P.C., et al., 2022. Ridge regression ensemble of machine learning models applied to solar and wind forecasting in Brazil and Spain. Appl. Energy 314, 118936. https://doi.org/10.1016/j.apenergy.2022.118936.

Chatzimparmpas, A., Martins, R.M., Kucher, K., et al., 2020. StackGenVis: alignment of data, algorithms, and models for stacking ensemble learning using performance metrics. IEEE Trans. Visual. Comput. Graph. 27 (2), 1547–1557. https://doi.org/10.1109/TVCG.2020.3030352.

Chen, X., Cao, W., Gan, C., et al., 2022. A hybrid partial least squares regression-based real time pore pressure estimation method for complex geological drilling process. J. Petrol. Sci. Eng. 210, 109771. https://doi.org/10.1016/j.petrol.2021.109771.

Delavar, M.R., Ramezanzadeh, A., 2023. Pore pressure prediction by empirical and machine learning methods using conventional and drilling logs in carbonate rocks. Rock Mech. Rock Eng. 56 (1), 535–564. https://doi.org/10.1007/s00603-022-03089-y.

Dietterich, T.G., 2000. Ensemble methods in machine learning. In: Multiple Classifier Systems. MCS 2000. Lecture Notes in Computer Science, vol. 1857. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-45014-9_1.

Dong, S.Q., Sun, Y.M., Xu, T., et al., 2023. How to improve machine learning models for lithofacies identification by practical and novel ensemble strategy and principles. Pet. Sci. 20 (2), 733–752. https://doi.org/10.1016/j.petsci.2022.09.006.

Eaton, B.A., 1975. The Equation for Geopressure Prediction from Well Logs. SPE Annual Technical Conference and Exhibition. https://doi.org/10.2118/5544-MS.

Evans, P., 2006. Scaling and assessment of data quality. Biolog. Crystallogr. 62 (1), 72–82. https://doi.org/10.1107/S0907444905036693.

Fillippone, W.R., 1982. Estimation of formation parameters and the prediction of overpressures from seismic data. SEG Technical Program Expanded Abstracts.

Fissha, Y., Ragam, P., Ikeda, H., et al., 2025. Data-driven machine learning approaches for simultaneous prediction of peak particle velocity and frequency induced by rock blasting in mining. Rock Mech. Bull. 4 (1), 100166. https://doi.org/10.1016/j.rockmb.2024.100166.

Gal, M.S., Rubinfeld, D.L., 2019. Data standardization. NYUL Rev. 94, 737. https://doi.org/10.2139/ssrn.3326377.

Huang, F., Xiong, H., Chen, S., et al., 2023. Slope stability prediction based on a long short-term memory neural network: comparisons with convolutional neural networks, support vector machines and random forest models. Int. J. Coal Sci. Technol. 10 (1), 18. https://doi.org/10.1007/s40789-023-00579-4.

Huang, J., Peng, Y., Hu, L., 2024. A multilayer stacking method base on RFE-SHAP feature selection strategy for recognition of driver's mental load and emotional state. Expert Syst. Appl. 238, 121729. https://doi.org/10.1016/j.eswa.2023.121729.

Jaiswal, J.K., Samikannu, R., 2017. Application of random forest algorithm on feature subset selection and classification and regression. In: 2017 World Congress on Computing and Communication Technologies (WCCCT). IEEE, pp. 65–68. https://doi.org/10.1109/WCCCT.2016.25.

Kardani, N., Zhou, A., Nazem, M., et al., 2021. Improved prediction of slope stability using a hybrid stacking ensemble method based on finite element analysis and field data. J. Rock Mech. Geotech. Eng. 13 (1), 188–201. https://doi.org/10.1016/j.jrmge.2020.05.011.

Liang, B., Liu, J., Kang, L.X., et al., 2024. A novel framework for predicting non-stationary production time series of shale gas based on BiLSTM-RF-MPA deep fusion model. Pet. Sci. 21 (5), 3326–3339. https://doi.org/10.1016/j.petsci.2024.05.012.

Liu, C., Yang, Z., Qin, Y., et al., 2024. Excess pore pressure behavior and evolution in deep coalbed methane reservoirs. Int. J. Min. Sci. Technol. 34 (6), 763–781. https://doi.org/10.1016/j.ijmst.2024.06.010.

Liu, F.T., Ting, K.M., Zhou, Z.H., 2008. Isolation forest. In: Proceedings of the 8th IEEE International Conference on Data Mining (ICDM 2008). IEEE Computer Society, pp. 413–422. https://doi.org/10.1109/ICDM.2008.17.

Lu, B., Wang, Z., Zhang, Y., 2022. Progress of theories and methods for prediction and detection of pore pressure in carbonate rock. Acta Pet. Sin. 43 (4), 571–580. https://doi.org/10.7623/syxb202204010 (in Chinese).

Ma, T., Xiang, G., Gui, J., et al., 2024. Physics-constrained distributed neural network model for 3D in-situ stress prediction. Chin. J. Geophys. 67 (8), 3211–3228. https://doi.org/10.6038/cjg2023R0132 (in Chinese).

Ma, T., Xiang, G., Shi, Y., et al., 2022. Horizontal *in situ* stresses prediction using a CNN-BiLSTM-attention hybrid neural network. Geomech. Geophys. Geo-Energy Geo-Resourc. 8 (5), 152. https://doi.org/10.1007/s40948-022-00467-2.

Matinkia, M., Amraeiniya, A., Behboud, M.M., et al., 2022. A novel approach to pore pressure modeling based on conventional well logs using convolutional neural network. J. Petrol. Sci. Eng. 211, 110156. https://doi.org/10.1016/j.petrol.2022.110156.

Mining, W.I.D., 2006. Data mining: concepts and techniques. Morgan Kaufinann 10 (559–569), 4. https://doi.org/10.1109/ICMIRA.2013.45.

Mitchell, R.F., Miska, S.Z., 2011. Fundamentals of Drilling Engineering. SPE Text Book Series, vol. 12. Society of Petroleum Engineers, USA. https://doi.org/10.2118/9781555632076.

Mughees, A., Kamran, M., Mughees, N., et al., 2024. New appliance signatures for nilm based on mono-fractal features and multi-fractal formalism. IEEE Access 12, 108986–109000. https://doi.org/10.1109/ACCESS.2024.3440168.

Otchere, D.A., Ganat, T.O.A., Gholami, R., et al., 2021. Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: comparative analysis of ANN and SVM models. J. Petrol. Sci. Eng. 200, 108182. https://doi.org/10.1016/j.petrol.2020.108182.

Pei, Z.J., Song, X.Z., Wang, H.T., et al., 2024. Interpretation and characterization of rate of penetration intelligent prediction model. Pet. Sci. 21 (1), 582–596. https://doi.org/10.1016/j.petsci.2023.10.011.

Peng, N., Ma, T., Chen, P., et al., 2021. Pore pressure evaluation of formation testing while drilling under supercharged conditions. J. Petrol. Sci. Eng. 203, 108689. https://doi.org/10.1016/j.petrol.2021.108689.

Phan, D.T., Liu, C., AlTammar, M.J., et al., 2022. Application of artificial intelligence to predict time-dependent mud-weight windows in real time. SPE J. 27 (1), 39–59. https://doi.org/10.2118/206748-PA.

Rashidi, M., Asadi, A., 2018. An Artificial Intelligence Approach in Estimation of Formation Pore Pressure by Critical Drilling Data. ARMA US Rock Mechanics/Geomechanics Symposium.

Rehm, B., McClendon, R., 1971. Measurement of Formation Pressure from Drilling Data. SPE Annual Technical Conference and Exhibition. https://doi.org/10.2118/3601-MS.

Rousseeuw, P.J., Croux, C., 1993. Alternatives to the median absolute deviation. J. Am. Stat. Assoc. 88 (424), 1273–1283. https://doi.org/10.1080/01621459.1993.10476408.

Roy, D.K., Ray, G.K., Biswas, A.K., 2010. Overview of overpressure in Bengal basin, India. J. Geol. Soc. India 75 (4), 644–660. https://doi.org/10.1007/s12594-010-0053-5.

Rzychoń, M., Żogała, A., Rog, L., 2022. SHAP-based interpretation of an XGBoost model in the prediction of grindability of coals and their blends. Int. J. Coal Preparat. Utilizat. 42 (11), 3348–3368. https://doi.org/10.1080/19392699.2021.1959324.

Sayers, C.M., Johnson, G.M., Denyer, G., 2002. Predrill pore-pressure prediction using seismic data. Geophysics 67 (4), 1286–1292. https://doi.org/10.1190/1.1500391.

Shajari, M., Najibi, H., 2012. Application of the dc-exponent method for abnormal pressure detection in Ahwaz oil field: A comparative study. Petrol. Sci. Technol. 30 (4), 339–349. https://doi.org/10.1080/10916466.2010.483439.

Sharma, V., 2022. A study on data scaling methods for machine learning. Int. J. Glob. Acad. Scient. Res. 1 (1), 31–42. https://doi.org/10.55938/ijgasr.v1i1.4.

Shehadeh, A., Alshboul, O., Al Mamlook, R.E., et al., 2021. Machine learning models for predicting the residual value of heavy construction equipment: An evaluation of modified decision tree, LightGBM, and XGBoost regression. Autom. ConStruct. 129, 103827. https://doi.org/10.1016/j.autcon.2021.103827.

Teale, R., 1965. The concept of specific energy in rock drilling. Int. J. Rock Mech. Min. Sci. Geomech. Abstracts 2 (1), 57–73. https://doi.org/10.1016/0148-9062(65)90022-7.

Ugwu, G.Z., 2015. An overview of pore pressure prediction using seismically-derived velocities. J. Geol. Min. Res. 7 (4), 31–40. https://doi.org/10.5897/JGMR15.0218.

Van Zyl, C., Ye, X., Naidoo, R., 2024. Harnessing eXplainable artificial intelligence for feature selection in time series energy forecasting: A comparative analysis of Grad-CAM and SHAP. Appl. Energy 353, 122079. https://doi.org/10.1016/j.apenergy.2023.122079.

Yang, H., Feng, Y., Shang, G., et al., 2025. A sequence learning approach for real-time and ahead-of-bit pore pressure prediction utilizing drilling data from the drilled section. SPE J. 30 (2), 524–543. https://doi.org/10.2118/223962-PA.

Yang, X., Wen, W., 2018. Ridge and lasso regression models for cross-version defect prediction. IEEE Trans. Reliab. 67 (3), 885–896. https://doi.org/10.1109/TR.2018.2847353.

Zhang, D., Ma, T., Liu, Y., et al., 2025. Heterogeneous multi-task learning approach for rock strength prediction in real-time during drilling based on mud log data. Geoenergy Sci. Eng. 246, 213608. https://doi.org/10.1016/j.geoen.2024.213608.

Zhang, G., Davoodi, S., Band, S.S., et al., 2022. A robust approach to pore pressure prediction applying petrophysical log data aided by machine learning techniques. Energy Rep. 8, 2233–2247. https://doi.org/10.1016/j.egyr.2022.01.012.

Zhang, J., 2011. Pore pressure prediction from well logs: Methods, modifications, and new approaches. Earth Sci. Rev. 108 (1–2), 50–63. https://doi.org/10.1016/j.earscirev.2011.06.001.

Zhang, X., Li, H., Tan, X., et al., 2023. Coupling effects of temperature, confining pressure, and pore pressure on permeability and average pore size of Longmaxi shale. Deep Undergr. Sci. Eng. 2 (4), 359–370. https://doi.org/10.1002/dug2.12047.

Zhang, X., Lu, Y.H., Jin, Y., et al., 2024. An adaptive physics-informed deep learning method for pore pressure prediction using seismic data. Pet. Sci. 21 (2), 885–902. https://doi.org/10.1016/j.petsci.2023.11.006.

Zhao, J., Wu, T., Pu, W., et al., 2024. Application status and research progress of CO2 fracturing fluid in petroleum engineering: A brief review. Petroleum 10, 1–10. https://doi.org/10.1016/j.petlm.2023.12.001.

Zounemat-Kermani, M., Batelaan, O., Fadaee, M., et al., 2021. Ensemble machine learning paradigms in hydrology: A review. J. Hydrol. 598, 126266. https://doi.org/10.1016/j.jhydrol.2021.126266.