GeoScienceWorld
Lithosphere
Volume 2022, Article ID 5249460, 19 pages
https://doi.org/10.2113/2022/5249460

Lithosphere

Research Article

Reservoir Permeability Prediction Based on Analogy and Machine Learning Methods: Field Cases in DLG Block of Jing'an Oilfield, China

Qiao Guo,[1] **Shiqing Cheng**,[1] **Fenghuang Zeng**,[2] **Yang Wang**,[1] **Chuan Lu**,[3] **Chaodong Tan**,[1] **and Guiliang Li**[4]

[1]*China University of Petroleum (Beijing), Beijing 102249, China*
[2]*12th Oil Production Plant of Changqing Oilfield Company, Xi'an 710021, China*
[3]*China National Offshore Oil Corporation Research Institute Co., Ltd., Beijing 100028, China*
[4]*Xi'an Supcon World Technology Development Co., Ltd., Xi'an 710021, China*

Correspondence should be addressed to Shiqing Cheng; chengsq973@163.com

Received 8 June 2022; Revised 26 August 2022; Accepted 6 September 2022; Published 30 September 2022

Academic Editor: Songjian Ao

Reservoir permeability, generally determined by experimental or well testing methods, is an essential parameter in the oil and gas field development. In this paper, we present a novel analogy and machine learning method to predict reservoir permeability. Firstly, the core test and production data of other 24 blocks (analog blocks) are counted according to the DLG block (target block) of Jing'an Oilfield, and the permeability analogy parameters including porosity, shale content, reservoir thickness, oil saturation, liquid production, and production pressure difference are optimized by Pearson and principal component analysis. Then, the fuzzy matter element method is used to calculate the similarity between the target block and analog blocks. According to the similarity calculation results, reservoir permeability of DLG block is predicted by reservoir engineering method (the relationship between core permeability and porosity of QK-D7 in similar blocks) and machine learning method (random forest, gradient boosting decision tree, light gradient boosting machine, and categorical boosting). By comparing the prediction accuracy of the two methods through the evaluation index determination coefficient (R^2) and root mean square error (RMSE), the CatBoost model has higher accuracy in predicting reservoir permeability, with R^2 of 0.951 and RMSE of 0.139. Finally, the CatBoost model is selected to predict reservoir permeability of 121 oil wells in the DLG block. This work uses simple logging and production data to quickly and accurately predict reservoir permeability without coring and testing. At the same time, the prediction results are well applied to the formulation of DLG block development technology strategy, which provides a new idea for the application of machine learning to predict oilfield parameters.

1. Introduction

In the process of exploration and development of oil and gas fields, permeability is the basis for determining production rate, optimizing completion perforation scheme, and selecting the best drainage position [1–5]. The main methods for determining reservoir permeability include indoor core test method, logging or seismic inversion processing method, and well testing analysis method [6–13]. Core permeability is a rock physics concept. Its acquisition is through the collection of core samples from the reservoir, generally using air as infiltration fluid for laboratory measurement of rock samples [14–16]. Logging permeability is a concept of geophysics, which is estimated according to the statistical relationship between logging parameters and reservoir permeability [11, 12, 17–19]. Well testing permeability is a concept of reservoir engineering, which is interpreted from shut-in pressure measurement data [20–22]. The reliability of measurement is not consistent with different ways of obtaining permeability, and core permeability is the main factor. The traditional way to predict permeability is to establish various physical models and calculate permeability

by formula derivation. In terms of physical model establishment, Amaefule et al. proposed the classical FZI (flow zone index) model, which can be used to predict the permeability from core to single well profile and also widely used in permeability prediction of homogeneous reservoirs such as sandstone, carbonate rock, and shale [23, 24]. Based on the FZI model, Rezaee et al. introduced the formation factor (F) and the current zoning index (CZI) to adjust the model [25]. Although traditional physical models are widely used to calculate permeability, they are highly dependent on core test data. Moreover, for the unique seepage characteristics of low permeability reservoirs, the classical and improved Amaefule models still have some limitations [26–28]. In order to solve these problems, later generations proposed more revised models [27, 29, 30]. For example, Nooruddin and Hossain proposed an improved K-C equation, which considered the relationship between cementation index (m), lithology factor (a), and porosity and modified FZI to FZIm [31]. However, the improved method still needs some parameters tested in the experiment, which are difficult to obtain and costly. Therefore, directly finding the correlation between permeability and logging data and production data is the key to economically, rapidly, and accurately predict the permeability of uncored reservoirs in low permeability reservoirs.

In recent years, with the application of big data method in oilfield, this method is more and more applied to permeability prediction [32–39]. Zhou et al. combined principal component analysis (PCA), clustering method, and regression analysis and found that the gas production of Marcellus shale was significantly affected by the number of hydraulic fractures, vertical depth, proppant, and fracturing fluid volume pumped during well stimulation [40]. Male et al. used linear regression model and gradient boosting model to predict the permeability of cemented sandstone in Beihai. The results showed that the gradient boosting model had better prediction effect [41]. Wang et al. used big data technology to predict reservoir parameters and oil well productivity in the western South China Sea oilfield [39]. Khalifah et al. took porosity, pore throat diameter, and formation factors as input structural parameters and used artificial neural network (ANN) and genetic algorithm to predict the permeability of tight carbonate rocks. Compared with the K-C equation, the Berg model and other seven traditional physical models. The results show that the machine learning method has higher accuracy [42]. Based on machine learning algorithm, logging data are also used to predict the permeability of sandstone and carbonate reservoirs [39, 43–48]. For example, Urang et al. used crude oil bulk density and water saturation as input parameters of permeability prediction model, combined with standardized nonlinear regression analysis, and proposed a new model based on artificial neural network [49].

In recent years, many people have used machine learning methods to model and predict permeability as a logging record function [50–55]. Al-Mudhafar used a probabilistic neural network (PNN) to obtain accurate lithofacies classification and then used the generalized boosted regression model (GBM) to predict permeability by establishing a nonlinear relationship between core permeability, well logging data, and lithofacies [52]. Lee et al. proposed a two-step method for logging permeability prediction using nonparametric regression combined with multivariate statistical analysis. Firstly, the logging data is classified into electrofacies types. Secondly, alternating conditional expectation (ACE), generalized additive model (GAM), and neural network (NNET) are used to predict permeability using logging in each electrical phase [56]. Al-Mudhafar et al. used extreme gradient boosting (XGBoost) and adaptive boosting (AdaBoost) configured as classifiers to measure discrete lithofacies distributions based on core data and then combined the predicted lithofacies with recorded logging data and analyzed them through an XGBoost regression model to predict permeability [57]. Al-Anazi et al. established a permeability prediction model using support vector regression (SVR) of well logging in heterogeneous sandstone reservoirs [58]. Perez et al. determined the relative importance of logging in identifying electrical facies, lithofacies, and HFU based on the tree method, explained the lost logging records in the permeability prediction process, and predicted the permeability in the Salt Creek Oilfield Unit (SCFU) in western Texas [59].

Other big data algorithms are also used in permeability prediction, including fuzzy logic model [60], XGBoost [61], SVM [62], and random forests [63]. Rafik et al. applied variable selection to the nonparametric regression ACE to improve permeability predictions [64]. Al-Mudhafar classified lithofacies using multinomial logistic regression (Multinom), logistic boosting regression (LogitBoost), and extreme gradient boosting (XGBoost) and incorporated the discrete phase distribution and logging interpretation derived from LogitBoost into a multivariate permeability model and predict corrected core permeability for logging interpretation of all wells in the reservoir [51, 52]. Zhao et al. trained decision tree regression to predict the permeability of low-permeability sandstones of the Zhuhai Formation in the Wenchang A sag, Pearl River Mouth Basin, using seven machine learning algorithms, including linear regression, backpropagation neural network regression, K-neighbor regression, random forest regression, support vector machine regression, gradient boosting decision tree regression, and extreme gradient boosting (XGBoost), with logging data as input [65]. Zhang et al. proposed a new method of univariate prediction model (UPP) and bivariate prediction model (BPP) and constructed support vector regression (SVR), random forest (RF), and deep residual neural network (ResNet) machine learning models for permeability prediction [17, 19]. Anifowose et al. studied the commonly used and complex ML techniques in estimating the permeability of carbonate reservoirs in the Middle East [62]. This study integrated seismic attributes and cable data to adjust the influence of hyperparameters on technical performance to improve permeability prediction. Previous studies have obtained good permeability prediction results based on various machine learning methods, but the application of predicted permeability in low permeability oilfield development technology adjustment is limited.

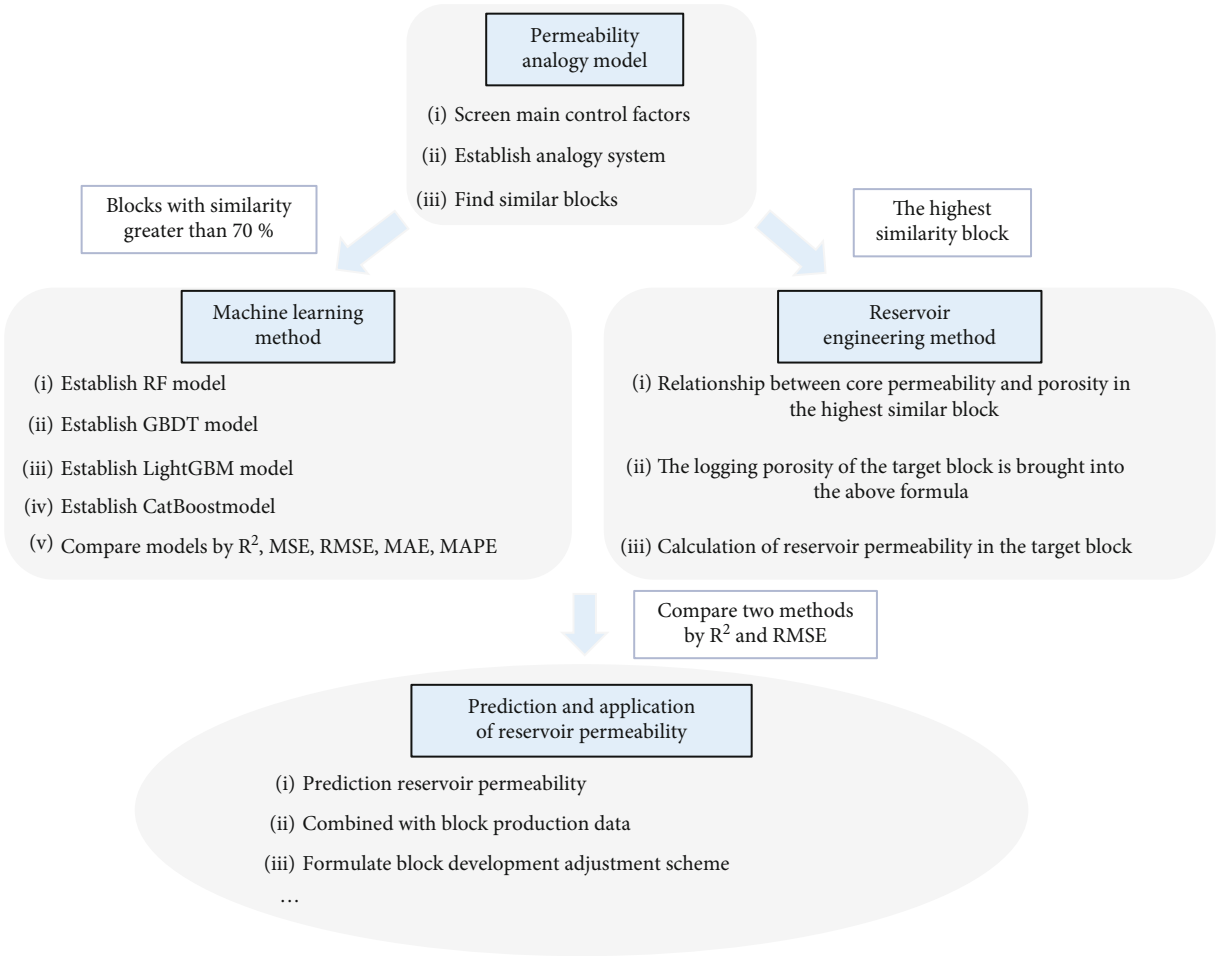

Figure 1: Reservoir permeability prediction and application flowchart.

In this paper, the core test and production data of 24 blocks (analog blocks) in Jing'an Oilfield are counted. Based on big data analogy method, the oil well reservoir permeability of DLG block (target block) in the oilfield is predicted, and the corresponding development technology policy is formulated according to the prediction results. The geological and reservoir characteristics of DLG block are analyzed firstly, and 11 dynamic and static parameters that affect the permeability of the analogy block are counted. The main controlling factors of permeability (porosity, mudstone content, reservoir thickness, oil saturation, liquid production, and production pressure differential) were selected from the above parameters by using Pearson and principal component analysis, and an analogy system of permeability containing six main controlling factors was established. Then, the similarity between the target block and the analogy block is calculated by using the fuzzy matter-element method. According to the similarity calculation results, reservoir engineering method (core permeability and core porosity relationship of the similar block) and machine learning method (RF model, GBDT model, LightGBM model, and CatBoost model) are used to predict the reservoir permeability of DLG block in the target block. Results show that the CatBoost model has higher accuracy in predicting reservoir permeability. Finally, the CatBoost model in machine learning method is selected to predict the reservoir permeability of 121 oil wells in the target block. Using the predicted reservoir permeability, combined with the formation pressure, water content and liquid production contour map, aiming at the problems existing in the current development process, the corresponding development technology strategy is put forward, which provides extremely valuable experience for the efficient development of the block. The methods of predicting reservoir permeability based on machine learning are highly dependent on core experimental parameters and have potential limitations in uncored areas. They fail to make good use of dynamic parameters and predict reservoir permeability without combining oilfield development technology policies. In this paper, the machine learning algorithm is used to make full use of the existing dynamic and static parameter data, find similar blocks by establishing an analogy system, and provide a set of processes for rapid prediction of reservoir permeability without coring and testing. The predicted reservoir permeability is well applied to the formulation of development technology strategy, which provides a new idea for the application of machine learning to predict oilfield parameters. The workflow of permeability prediction and application is shown in Figure 1.

Table 1: Actual values of 24 block analogy parameters.

Block	K	$P1$	$P2$	$P3$	$P4$	$P5$	$P6$	$P7$	$P8$	$P9$	$P10$	$P11$
DLG	0.54	13.50	15.60	21.90	57.90	1.50	5.56	1.96	860	1.20	28.60	2
QK-D7	0.61	13.34	15.19	18.97	57.21	2.58	3.24	2.74	746	1.21	30.24	1
QK-D13	0.42	12.97	14.72	19.01	56.94	3.47	4.26	1.72	698	1.20	32.16	3
QK-D17	0.72	12.56	14.23	17.64	54.23	1.69	4.69	1.69	842	1.19	19.47	2
QK-D14	0.76	14.21	13.96	17.21	53.97	2.42	2.76	2.31	834	1.20	27.84	1
QK-D5	0.43	14.87	12.46	25.42	52.68	3.78	2.49	2.54	867	1.23	34.26	1
QK-D23	0.36	11.94	11.82	26.38	51.02	3.69	3.12	3.21	941	1.18	36.48	3
QK-D4	0.89	11.84	12.21	26.94	63.97	2.97	3.06	3.28	745	1.24	32.16	4
QK-D8	0.84	11.97	18.74	16.38	64.02	2.68	7.28	4.27	747	1.20	39.46	2
QK-D10	0.96	15.06	19.84	27.69	63.86	2.59	1.26	1.46	762	1.23	38.45	3
QK-D6	1.02	18.96	20.16	15.24	48.97	2.47	1.39	1.87	624	1.26	42.16	1
QK-D1	0.73	10.67	21.32	28.14	63.92	1.65	2.91	1.94	659	1.21	24.74	3
QK-D9	0.91	19.63	24.36	28.63	47.26	1.24	3.45	1.25	613	1.18	46.23	3
QK-D3	0.69	9.94	25.47	13.29	46.97	1.03	3.77	3.25	579	1.20	50.12	4
QK-D16	1.14	21.24	9.86	10.29	45.61	1.65	2.15	3.57	781	1.19	62.31	4
QK-D20	0.74	10.06	8.21	30.04	44.28	4.69	1.38	4.16	632	1.20	24.16	2
QK-D11	0.82	10.47	7.46	8.67	43.61	4.79	4.67	2.67	648	1.21	60.48	2
QK-D24	1.96	29.46	9.28	7.69	42.12	5.85	4.21	1.36	653	1.21	55.27	3
QK-D18	2.87	34.62	26.39	31.16	41.68	5.92	6.17	1.17	725	1.19	51.02	1
QK-D2	0.75	9.14	28.47	6.21	39.65	6.87	6.74	1.69	749	1.21	49.61	2
QK-D15	0.85	8.46	30.26	6.03	35.17	8.21	1.32	6.84	865	1.24	61.03	3
QK-D12	3.65	36.74	7.21	34.67	33.12	9.14	5.02	10.21	813	1.23	61.27	3
QK-D22	1.14	20.71	31.65	34.89	34.16	12.36	5.80	9.46	816	1.21	57.68	2
QK-D19	1.56	21.33	35.87	5.42	32.17	14.69	6.02	8.72	742	1.20	50.19	2
QK-D21	3.24	36.74	41.23	4.18	31.06	18.23	5.96	7.69	761	1.17	47.65	3

K refers to permeability; $P1$~$P11$ refer to porosity, shale content, reservoir thickness, oil saturation, liquid production, production pressure difference, oil viscosity, surface crude oil density, oil volume factor, water content, and number of oil layers.

2. Establishment of Permeability Analogy Model

The purpose of block analogy is to find blocks similar to the characteristics of the target block. Due to the same or similar geological characteristics and development methods between the analogy block and the target block, the development experience and exploitation mode accumulated by the analogy block after long-term development adjustment and optimization can indicate the direction for the efficient development of the target block.

How to select analogy parameters accurately has become an important part of block analogy. When judging whether the blocks are similar, it is usually necessary to consider the static characteristic parameters. The "similarity" between the target block and the analog block is determined by parameter comparison, and the most similar block is selected to predict the reservoir permeability of the target block. Due to the limited parameters affecting reservoir permeability, some analog parameters are difficult to collect, and different parameters have different effects on permeability. Therefore, in order to improve the operability and practicability of analogy, it is necessary to select the main influencing factors of permeability from the collected block parameters as the analogy index system.

Table 2: Pearson correlation coefficient and correlation degree classification.

The variation range of correlation coefficient	Correlation intensity
0.8~1.0	Extremely strong correlation
0.6~0.8	Strong correlation
0.4~0.6	Moderately relevant
0.2~0.4	Weak correlation
0~0.2	Very weak or no correlation

In this paper, according to the general situation of Jing'an Oilfield onshore low permeability oilfield, according to the existing research results of geological and engineering factors on permeability, 11 parameters affecting permeability are determined, and the data of 24 blocks affecting permeability in the oilfield are counted. The permeability is mainly affected by the characteristics of static parameters, and the influencing factors include eight static geological parameters and three dynamic production parameters. Static geological

Table 3: Correlation matrix between permeability and influence parameters.

Analysis index	K	$P0$	$P1$	$P2$	$P3$	$P4$	$P5$	$P6$	$P7$	$P8$	$P9$	$P10$
K	1.000	0.834	0.682	0.506	0.479	0.427	0.408	0.356	0.208	0.139	0.067	0.019
$P0$	0.834	1.000	0.124	0.237	0.478	0.146	0.081	0.104	0.216	0.168	0.076	0.073
$P1$	0.682	0.124	1.000	0.013	0.030	0.024	0.017	0.036	0.027	0.018	0.013	0.058
$P2$	0.506	0.237	0.013	1.000	0.124	0.018	0.032	0.102	0.126	0.064	0.082	0.041
$P3$	0.479	0.478	0.030	0.124	1.000	0.176	0.080	0.041	0.052	0.021	0.013	0.036
$P4$	0.427	0.146	0.024	0.018	0.176	1.000	0.361	0.482	0.173	0.079	0.033	0.024
$P5$	0.408	0.081	0.017	0.032	0.080	0.361	1.000	0.017	0.026	0.001	0.032	0.018
$P6$	0.356	0.104	0.036	0.102	0.041	0.482	0.017	1.000	0.018	0.022	0.012	0.001
$P7$	0.208	0.216	0.027	0.126	0.052	0.173	0.026	0.018	1.000	0.014	0.003	0.009
$P8$	0.139	0.168	0.018	0.064	0.021	0.079	0.001	0.022	0.014	1.000	0.001	0.002
$P9$	0.067	0.076	0.013	0.082	0.013	0.033	0.032	0.012	0.003	0.001	1.000	0.004
$P10$	0.019	0.073	0.058	0.041	0.036	0.024	0.018	0.001	0.009	0.002	0.004	1.000

$P0$~$P10$ refer to porosity, shale content, reservoir thickness, oil saturation, liquid production, production pressure difference, oil viscosity, surface crude oil density, oil volume factor, water content, and number of oil layers; K refers to permeability.

parameters including porosity, mudstone content, number of oil layers, oil viscosity, reservoir thickness, oil volume factor, oil saturation, surface crude oil density, and the pore throat ratio data are missing and not included in the analysis. The dynamic production parameters include liquid production, production pressure difference, and water content. The range of each parameter is shown in Table 1.

2.1. Screening of Main Control Factors. In this paper, Pearson correlation coefficient method is used to calculate the correlation coefficient between the influence parameters and permeability, and the parameters with strong correlation with permeability are preliminarily screened [66–68]. This method can measure the linear correlation between two random variables X and Y, and the range of correlation coefficient is $(-1, 1)$. When X is linearly correlated with Y, the correlation coefficient is 1 (positive linear correlation) or -1 (negative linear correlation), the closer the absolute value of the correlation coefficient is to 1, the stronger the correlation between the two, and the specific correlation is shown in Table 2.

The Pearson correlation coefficient method was used to analyze the correlation between 11 dynamic and static influencing factors and permeability in 24 samples. The static influencing factors with strong correlation with the permeability have a great influence on the permeability and are initially selected as the analog index, which is left to the principal component analysis method for verification. The analysis results are shown in Table 3 and Figure 2.

It can be concluded from Table 3 and Figure 2 that among the 11 permeability influencing factors, porosity and shale content are strongly correlated with permeability; the correlation coefficients were 0.834 and 0.682, respectively; reservoir thickness, oil saturation, liquid production, and production pressure difference are moderately correlated with permeability, with correlation coefficient between 0.4 and 0.6; oil viscosity and surface crude oil density are weakly correlated with permeability, and other factors are extremely weakly correlated with productivity.

It can be seen from the correlation matrix diagram (Table 3) that among the various influencing factors in the block, the liquid production is moderately correlated with the oil viscosity, the oil saturation is moderately correlated with the porosity, and the correlation between other influencing factors is weak. This is because the permeability of the block studied in this paper is low, and the data of geological factors and fluid properties have small changes, and the correlation is not strong. At the same time, affected by the quantity and quality of data, the deeper and universal laws between data may not be reflected.

2.2. Establishment of Permeability Analogy System. In this paper, the principal component analysis method is used to establish the permeability analogy system combined with the Pearson correlation analysis results. Principal component analysis is a commonly used method in statistics, which is suitable for dealing with data with high dimensions and strong correlation among variables. For PCA algorithm, if the characteristic parameters with strong input correlation lead to the rank of the input matrix close to 0, the matrix operation results of PCA algorithm will cause a large deviation. Through the Pearson parameter matrix table, it can be concluded that the correlation between the selected permeability factors is not strong, so the 11 features are all used for principal component analysis. The step of this method is to standardize the sample data and then calculate the characteristic root and variance percentage. The calculation results are shown in Table 4, and the main characteristics of accumulating more than 90% variance information are obtained as the analogy index system.

The components with eigenvalues greater than 1 and cumulative contribution rate greater than 90% are selected as principal components. From Table 4, it can be seen that there are 4 components with eigenvalues greater than 1

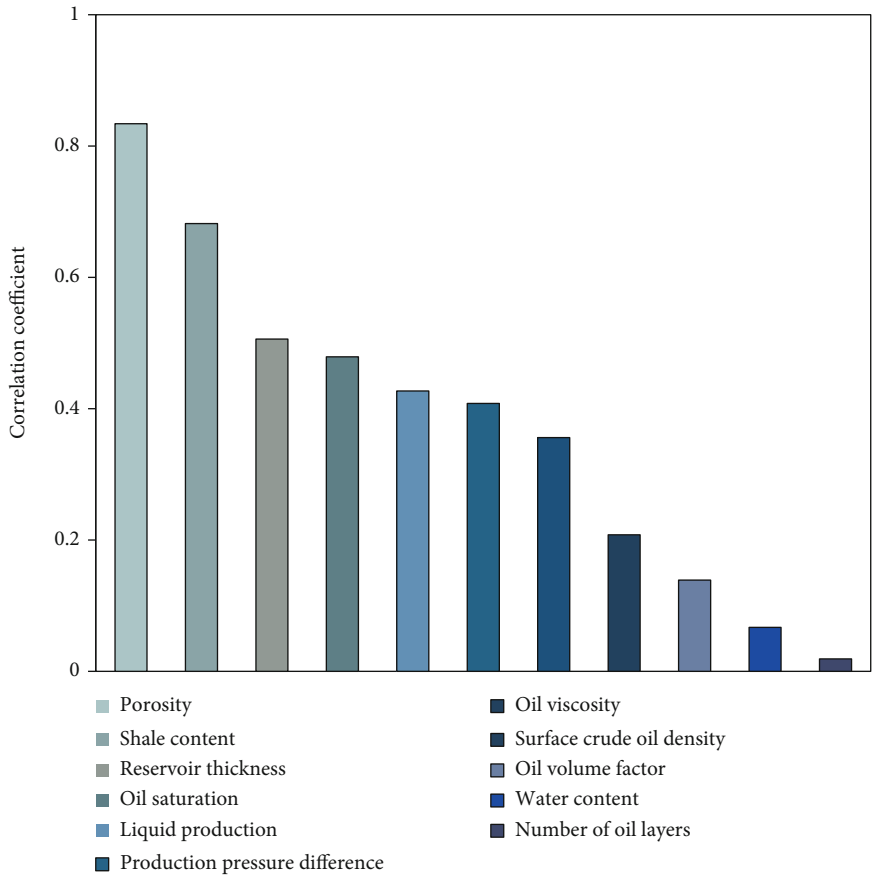

Figure 2: Correlation coefficient between permeability and influence parameters.

Table 4: Total variance interpretation of permeability influence parameters.

Component	Total variance explanation		
	Latent root	Latent root Variance proportion (%)	Accumulation (%)
1	4.306	39.142	39.142%
2	2.858	25.982	65.124
3	1.537	13.975	79.099
4	1.017	9.247	88.346
5	0.813	7.389	95.735
6	0.257	2.339	98.074
7	0.146	1.324	99.398
8	0.054	0.497	99.895
9	0.011	0.099	99.994
10	0.001	0.006	100.000
11	0.000	0.000	100.000

Components 1~11 refer to porosity, shale content, reservoir thickness, oil saturation, liquid production, production pressure difference, oil viscosity, surface crude oil density, oil volume factor, water content, and number of oil layers, respectively.

and 5 components with cumulative contribution rate greater than 90%. From the slope of eigenvalue decline, it can be concluded that the decline slope gradually slows down from the sixth component. In order to extract the information of the original variables as much as possible, combined with Pearson correlation analysis, the first six components with a cumulative contribution rate of 98.047% were finally extracted as the final principal components. The main controlling factors affecting permeability are porosity, shale content, reservoir thickness, liquid production, oil saturation, and production pressure difference.

Table 5: Weight calculation results of permeability principal component parameters.

Name of principal component	Variance explained rate	Cumulative variance interpretation rate	Weight
Porosity	0.391	0.391	0.399
Shale content	0.260	0.651	0.265
Oil saturation	0.14	0.791	0.143
Reservoir thickness	0.092	0.883	0.094
Liquid production	0.074	0.957	0.075
Production pressure difference	0.023	0.981	0.024

Table 6: Target block DLG basic dynamic and static parameter.

Parameter	Value	Parameter	Value
Block area (km^2)	10.5	Geological reserves (10^4t)	780
Average burial depth (m)	1680	Reservoir temperature (K)	325
Rock type	Mainly medium-fine lithic feldspar sandstone	Sedimentary micro	Lower distributary channel, flank of underwater distributary channel, mouth bar, diversion bend
Acid sensitivity	Weak	Speed sensitivity	Weak
Original formation pressure (MPa)	11.5	Saturation pressure (MPa)	7.26
Number of oil layers	2	Reservoir thickness (m)	21.70
Permeability (10^{-3} μm^2)	0.54	Porosity (%)	13.5
Coefficient of variation	1.1	Oil saturation (%)	51.9
Surface crude oil density (kg/m^3)	860	Oil viscosity (mPa·s)	1.96
Oil volume factor	1.2	Shale content	15.6
Water well	41	Well network	Rhombus inverse nine-spot well pattern
Well spacing density ($well/km^2$)	11.52	Oil well	121
Injection-production ratio	3.1	Production pressure difference (MPa)	5.56
Average daily liquid production per well (m^3/d)	1.5	Water content (%)	28.6
Average daily water injection per well (m^3/d)	17	Oil production rate (%)	0.59

Next, the linear combination coefficient of each index in each principal component is calculated, and the weight value of each principal component is finally calculated according to the information such as load coefficient. The results are shown in Table 5.

The establishment of analogy system mainly considers the static characteristic parameters. It can be seen from the weight values of each principal component in Table 5 that the weight values of liquid production and production pressure difference of dynamic characteristic parameters are small, which are lower than 10%. The weight of the four static characteristic parameters is consistent with the actual situation in the field. Therefore, the above six parameters and the calculated weight value are selected to construct the permeability analogy system.

2.3. Finding of Similar Blocks. According to the analog parameter system established by Pearson correlation coefficient method and principal component analysis method, fuzzy matter element method is selected to calculate the similarity between analog blocks and target blocks according to the weight of analog parameters. The target block DLG is an actual block in Jing'an Oilfield. The average burial depth of the reservoir is 1680 m, the average reservoir thickness is 21.7 m, the average porosity is 13.5%, the average permeability is $0.54 \times 10 - 3\,\mu m^2$, the original reservoir formation pressure is 11.5 MPa, and the saturation pressure is 7.26 MPa, which is the same pressure system. The reservoir temperature is 325 K. The basic information of the block is shown in Table 6.

The relevant parameter data of 24 blocks in Jing'an Oilfield are used as analog blocks, and the similarity calculation needs to collect the above analog parameters as the basis. However, some block parameter data are vacant, and the missing data content and quantity of each block are different. Considering the need for null value processing in the similarity calculation process, the fuzzy matter-element method is used to define the fuzzy value of the index

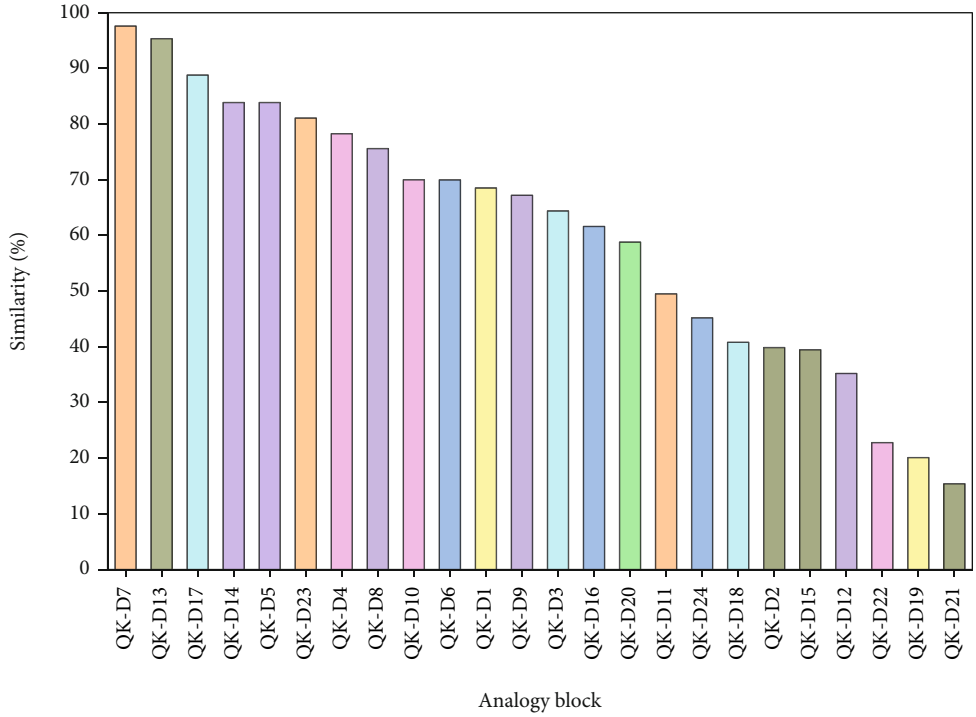

Figure 3: The calculation results of the similarity of the target block DLG search analogy block.

corresponding to the reservoir null value as 0. That is, the parameters corresponding to the null value are not involved in the similarity accumulation.

The similarity between the *n*th block and the target block (DLG block) is expressed by the sum of the product of the fuzzy value of each target block and the corresponding weight of each parameter.

The similarity calculation results show that the QK-D7 target block is the most similar, the similarity is as high as 97.6%, and there are 8 oilfields with similarity greater than 70% (Figure 3). The relationship between core permeability and core porosity in QK-D7 block is selected by reservoir engineering method to predict reservoir permeability in the target block. 421 single well samples from the first 8 blocks with similarity greater than 70% were selected to establish permeability prediction model by big data method.

3. Reservoir Engineering Method

Logging permeability is determined by mathematical statistics and other methods to establish the statistical regression relationship between reservoir permeability and rock porosity, specific surface, bound water saturation, and other parameters. The logging permeability can reflect the average permeability of the formation in a large area near the wellbore, but it cannot well reflect the real average permeability of the formation with strong heterogeneity. The core permeability is measured by coring the target layer and cleaning the lithology and then using air as the medium to measure the absolute permeability of the core. The measurement method can directly measure the absolute permeability of the core at the coring point. It is a more direct measurement method of permeability, which is more direct and accurate than logging permeability, but the cost of this method is high.

Therefore, the specific steps of reservoir permeability prediction by reservoir engineering method in this paper are as follows. The existing relationship between core permeability and porosity in similar blocks is counted, and the logging porosity of the target block is brought into the relationship to correct the existing logging permeability, so as to obtain more accurate reservoir permeability.

The similarity calculation results show that the QK-D7 block is the most similar to the DLG block. The permeability of QK-D7 block is low, with 23 measured core data. The linear relationship between core permeability and porosity is shown in Figure 4.

It can be seen from the diagram that the linear relationship between permeability and porosity in similar blocks is good, R^2 is 0.904. The logging porosity of DLG block is substituted into the relationship between permeability and porosity of QK-D7 core in similar blocks, and the calculated permeability is DLG block reservoir permeability (Table 7).

The core permeability of 11 wells in the DLG block is measured, and the predicted reservoir permeability is compared with the measured core permeability, and permeability is concentrated in $0.5 - 1.0 \times 10^{-3}\ \mu m^2$ (Table 7). From the comparison chart of the two (Figure 5), it can be intuitively seen that the error of the two is small, and the R^2 value is 0.906, which further shows that the method of predicting reservoir permeability using big data analogy combined with reservoir engineering-related formulas is more reasonable.

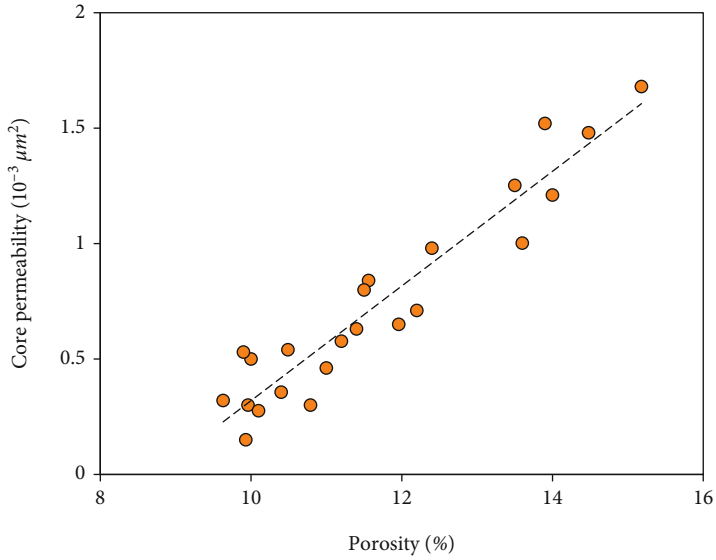

Figure 4: Relationship between core permeability and core porosity of QK-D7 in similar block.

Table 7: Comparison between predicted reservoir and core permeability of DLG block (prediction and correction of DLG logging permeability and core permeability in block QK-D7 by using relationship between core permeability and porosity).

Logging porosity	Predicted reservoir permeability	Core permeability
10.82	0.53	0.58
13.24	1.13	0.92
11.47	0.69	0.82
10.56	0.46	0.69
11.48	0.69	0.60
15.62	1.73	1.59
11.27	0.60	0.73
10.02	0.33	0.47
9.17	0.12	0.15
12.36	0.91	0.72
9.78	0.27	0.25

4. Machine Learning Method

Six parameters of analogy system are selected as input parameters, and permeability is selected as target parameters. Four machine learning algorithms (random forest, gradient boosting decision tree, light gradient boosting machine, and categorical boosting) are used to establish permeability prediction models. Using the similarity calculation results, the data of 421 oil wells in the first 8 blocks with a similarity greater than 70% were selected as the sample data. Due to the small amount of sample data in this paper, the application of random secondary sampling cross-validation method [69–71], using 70% of the sample data as a training set and the remaining 30% as a test set, helps to obtain a reliable and stable model. The random quadratic sampling cross-validation method can be used to evaluate the prediction performance of the model, especially the performance of the trained model on new data. It can reduce the overfitting to a certain extent and can obtain as much effective information as possible from limited data, so as to find suitable model parameters more conveniently. Then, the accuracy of the model was evaluated by the mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and determination coefficient (R^2) of five evaluation indexes, and the prediction model with the best performance was selected. The first four evaluation index value is smaller; the higher the accuracy and the R^2 value closer to 1 indicate that the effect is better; see Table 8.

4.1. Random Forest (RF) Model. The permeability prediction model is established based on random forest algorithm. The parameters of the algorithm are listed in Table 9. 90% of the sample data is used to train the capacity prediction model, and 10% of the sample data is used to test the performance of the model. The RF model prediction results are shown in Figure 6. It can be seen intuitively from the diagram that when the permeability is lower than 2×10^{-3} μm^2, the prediction effect is better.

It can be seen from the prediction results that the training set fits well. When the permeability is greater than 3×10^{-3} μm^2, the prediction result is poor, the test set has a poor fit, and there is an overfitting phenomenon. The performance evaluation results of the model are shown in Table 10. The determination coefficients R^2 of the training set and the test set are 0.946 and 0.884, and the root mean square error is low. The smaller the value of the first four evaluation indexes indicates the higher the accuracy. The relationship between the training set and the sample data is used to train the model. Therefore, the accuracy of the

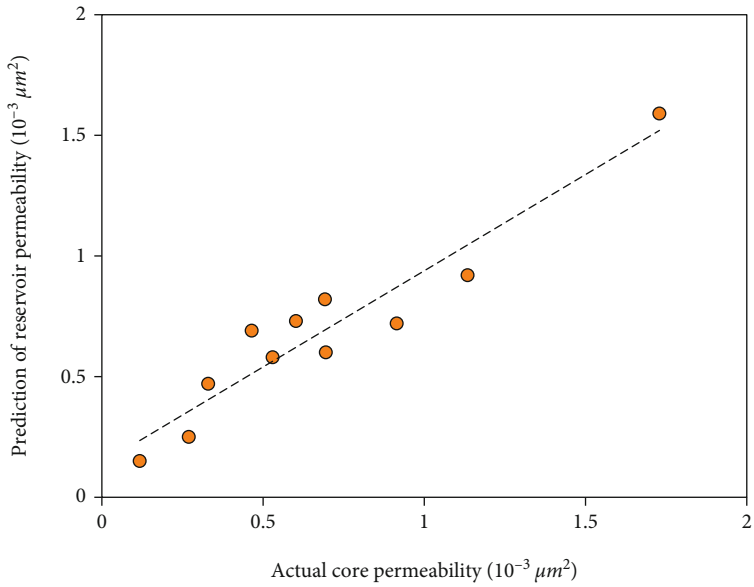

Figure 5: Comparison of predicted reservoir permeability and core permeability in the DLG block.

Table 8: Regression model evaluation index.

Evaluation index	Formula	Evaluation criteria
R^2 score	$R^2 = 1 - \sum_{i=1}^{n_{\text{samples}}} \left(y_{\text{ture}} - y_{\text{pred}}\right)^2 / \sum_{i=1}^{n_{\text{samples}}} \left(y_{\text{ture}} - \overline{y_{\text{ture}}}\right)^2$	The results are between 0 and 1. The larger the value is, the better the effect is.
Mean squared error	$\text{MSE} = 1/n_{\text{samples}} \sum_{i=1}^{n_{\text{samples}}} \left(y_{\text{ture}} - y_{\text{pred}}\right)^2$	The smaller the calculation result, the smaller the error.
Root mean square error	$\text{RMSE} = \sqrt{1/n_{\text{samples}} \sum_{i=1}^{n_{\text{samples}}} \left(y_{\text{ture}} - y_{\text{pred}}\right)^2}$	The result is between 0 and 1. The smaller the value is, the worse the result is.
Mean absolute error	$\text{MAE} = 1/n_{\text{samples}} \sum_{i=1}^{n_{\text{samples}}} \left\lvert y_{\text{ture}} - y_{\text{pred}} \right\rvert$	The larger the calculation results, the greater the error.
Mean absolute percentage error	$\text{MAPE} = 1/n_{\text{samples}} \sum_{i=1}^{n_{\text{samples}}} \left\lvert y_{\text{pred}} - y_{\text{ture}}/y_{\text{ture}} \right\rvert$	The smaller the calculation result, the smaller the error.

Notes: y_{ture} is the actual output value (observed value) for data y, and y_{pred} is the corresponding predicted value.

Table 9: Parameter values of the RF model.

Parameter name	Parameter value
Minimum number of samples for internal node splitting	2
Minimum sample number of leaf nodes	1
Minimum weight of samples in leaf nodes	0
Maximum depth of a tree	10
Maximum number of leaf nodes	50
Threshold of node partition purity	0
Number of decision trees	100

training set in Table 10 is higher than that of the test set, and the variance is small.

4.2. Gradient Boosting Decision Tree (GBDT) Model. The permeability prediction model is established based on GBDT algorithm. The parameters of the algorithm are listed in Table 11. The model is divided into training set and test set, which are the same as the RF model. It can be seen from the prediction results (Figure 7) that compared with the above model, the GBDT model performs better on the training set and test set than the RF model. When the permeability is higher than 2×10^{-3} μm^2, the model prediction effect is better. However, when the permeability is higher than 6×10^{-3} μm^2, the training set effect is poor.

The evaluation results of the performance index of the model are shown in Table 12. The determination coefficients R^2 of the training set and the test set are 0.957 and 0.918. The R^2 value of the training set and the test set is greater than that of the RF model, and the root mean square error is greater than that of the RF model. Compared with the above models, the overall performance of the GBDT model is better than that of the RF model.

4.3. Light Gradient Boosting Machine (LightGBM) Model. The permeability prediction model is established based on LightGBM algorithm. The calculation efficiency of this

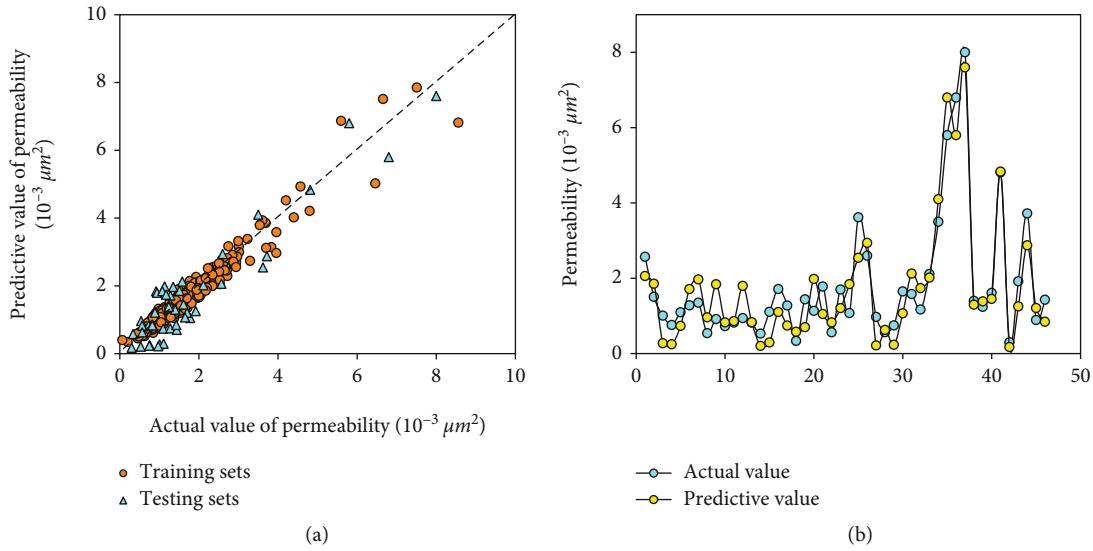

Figure 6: The results of reservoir permeability predicted by the RF model. (a) The comparison of actual permeability and predicted permeability of training set and test set. (b) The prediction results of the test set.

Table 10: Evaluation results of RF model index.

Sample data	MSE	RMSE	MAE	MAPE	R^2
Training sets	0.046	0.214	0.144	10.033	0.946
Testing sets	0.095	0.308	0.224	16.068	0.884

Table 11: Parameter values of the GBDT model.

Parameter name	Parameter value
Learning rate	0.08
Minimum number of samples for internal node splitting	2
Minimum sample number of leaf nodes	1
Minimum weight of samples in leaf nodes	0
Maximum depth of a tree	10
Maximum number of leaf nodes	50
Threshold of node partition purity	0

model is high, and the training parameters are shown in Table 13. The prediction results are shown in Figure 8. It can be seen from the figure that the fitting accuracy of the model training set is very high, and there is an overfitting phenomenon. The prediction error on the test sample is close to that of GBDT model. When searching for the optimal solution, the model is based on the optimal segmentation variable and does not take into account the idea that the optimal solution is the synthesis of all features, resulting in poor consistency of the test set.

The performance evaluation results of the model are shown in Table 14. The determination coefficients R^2 of the training set and the test set are 0.976 and 0.923. The R^2 value of the training set is much larger than that of the other two models, and the root mean square error is close to that of the RF model. Compared with the above two models, the LightGBM model performs best on the training set, and the matching effect on the test set is not good, and overall prediction is better than the other two models.

4.4. Categorical Boosting (CatBoost) Model. The permeability prediction model is established based on CatBoost algorithm. The parameters of the algorithm are listed in Table 15. The difference between the predicted permeability and the actual permeability is small, and the fitting data can be better when the permeability is high, and the prediction results are shown in Figure 9. Compared with LightGBM and GBDT models, the CatBoost model better solves the problems of prediction migration and overfitting. So the CatBoost model has better stability and generalization ability. The training set and test set are both good.

The performance evaluation results of the model are shown in Table 16. The determination coefficients R^2 of the training set and the test set are 0.987 and 0.951, and the root mean square error is small; the model has high accuracy. Based on the above three models, this model has the best overall performance.

5. Comparison of Two Methods

Based on reservoir engineering and machine learning methods that can predict reservoir permeability, the determination coefficient (R^2) and root mean square error

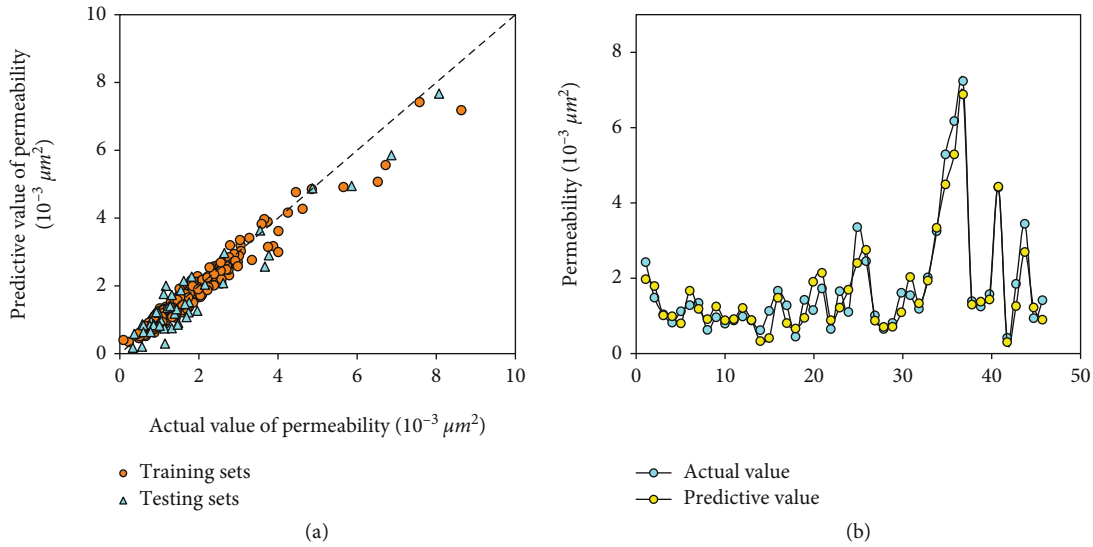

Figure 7: The results of reservoir permeability predicted by the GBDT model. (a) The comparison of actual permeability and predicted permeability of training set and test set. (b) The prediction results of the test set.

Table 12: Evaluation results of GBDT model index.

Sample data	MSE	RMSE	MAE	MAPE	R^2
Training sets	0.133	0.365	0.246	16.366	0.957
Testing sets	0.103	0.321	0.268	20.433	0.918

Table 13: Parameter values of the LightGBM model.

Parameter name	Parameter value
Data segmentation	0.8
Base learner	gbdt
Number of base learners	100
Learning rate	0.1
L1 regular term	0
L2 regular term	1
Sample sampling rate	1
Tree characteristic sampling rate	1
Node splitting threshold	0
Minimum weight of samples in leaf nodes	0
Maximum depth of a tree	10
Minimum sample number of leaf nodes	10

(RMSE) are used to evaluate the accuracy of the two methods (Figure 10).

The closer the R^2 value is to 1, the higher the accuracy of the model is, and the RMSE is the opposite. It can be seen from the figure that the CatBoost model R^2 is the largest, followed by the LightGBM model. The determination coefficients of the reservoir engineering method and the GBDT model are similar, while the RMSE coefficients of the reservoir engineering method and the CatBoost model are similar, far less than those of the other three models. Finally, the CatBoost model with the best generalization and the best performance is selected to predict the permeability of oil well reservoir in the DLG block.

6. Application

According to the comparison results of the above different reservoir permeability prediction methods, the CatBoost model is finally selected to predict the reservoir permeability of 121 wells in the DLG block.

The area of DLG block is 10.5 km^2, and the original geological reserves are 7.8 million tons. The oil-bearing strata are mainly Triassic Yanchang Formation, and the sedimentary microfacies are mainly underwater distributary channel microfacies. The regional structure is a west-dip monoclinic with east high and west low, and the dip angle is less than 1°. According to the sensitivity test, the reservoir has no water sensitivity, no salt sensitivity, weak speed sensitivity, and weak acid sensitivity. The reservoir rock types in this area are mainly medium-fine grained lithic feldspathic sandstone with strong heterogeneity, poor reservoir oil moisture, inactive reservoir edge and bottom water, and small ground saturation pressure difference. The reservoir is mainly controlled by traps formed by the tight layer in the upward direction and the sand body lateral deformation zone. Natural fractures are not developed in general, and the oil well dynamic mainly reflects the artificial fractures. The oil reservoirs in this area are low-permeability lithologic blocks with

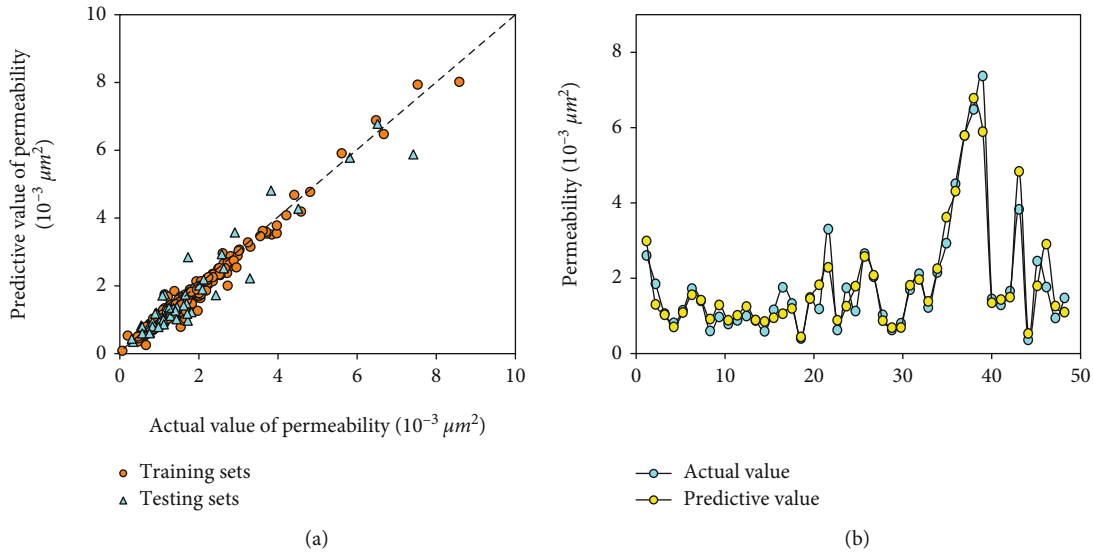

Figure 8: The results of reservoir permeability predicted by the LightGBM model. (a) The comparison of actual permeability and predicted permeability of training set and test set. (b) The prediction results of the test set.

Table 14: Evaluation results of LightGBM model index.

Sample data	MSE	RMSE	MAE	MAPE	R^2
Training sets	0.068	0.158	0.123	7.114	0.976
Testing sets	0.079	0.181	0.219	17.371	0.923

Table 15: Parameter values of the CatBoost model.

Parameter name	Parameter value
Data segmentation	0.9
Iteration times	100
Learning rate	0.1
L2 regular term	1
Maximum depth of a tree	10
Overfitting detection threshold	0
Number of iterations continued after optimization	20

low reservoir abundance, dense lithology, poor reservoir physical properties, and strong intralayer heterogeneity.

At present, there are 162 oil and water wells in this block, including 121 oil wells and 41 wells. At present, it is in the stage of production capacity construction. The rhombic inverted nine-spot well pattern is used, the average daily water injection is 17 m^3/d, and the average single well liquid production is 1.5 m^3/d. The maximum daily liquid production of a single well reaches 5.93 m^3/d, the maximum daily water injection of a single well reaches 21.46 m^3/d, the maximum oil production of a single well reaches 4.27 m^3/d, and the oil recovery rate is 0.34%–0.85%. This paper uses the oil well reservoir permeability predicted in this block to draw the reservoir permeability contour map (Figure 11(a)).

According to the predicted reservoir permeability contour map (Figure 11(a)), it can be concluded that the reservoir permeability value of 51.2% oil wells in the block is higher than 2×10^{-3} μm^2, and the reservoir permeability value of 29.4% oil wells is lower than 1×10^{-3} μm^2. Among them, the reservoir permeability of 133-36 well, lj34-362 well, 134-34 well, lj34-344 well, and l35-34 well is higher, with an average of 3.47×10^{-3} μm^2.

It can be seen from the contour map of formation pressure (Figure 11(b)) that the current formation pressure in this block is about 8 MPa, and the formation pressure in the central region is relatively high.

It can be seen from the water content contour map (Figure 11(c)) that most of the oil wells are in the middle and low water content period. Among 121 oil wells, 60.2% of the wells are with water content less than 20%, 25.3% of the wells are with water content between 20 and 40%, 10.4% of the wells are with water content between 40 and 60%, and 4.1% of the wells are with water content greater than 60%.

Combined with the contour map of fluid production in this block (Figure 11(d)), it can be concluded that the regions with high fluid production are mainly concentrated in the southwest, west, and southeast regions. The regions with high reservoir permeability have low fluid production and low water injection utilization rate.

In general, 75.2% of the oil well liquid production is less than 1.5 t/d, the overall moisture content is low, the oil well liquid supply capacity is insufficient, the permeability of the

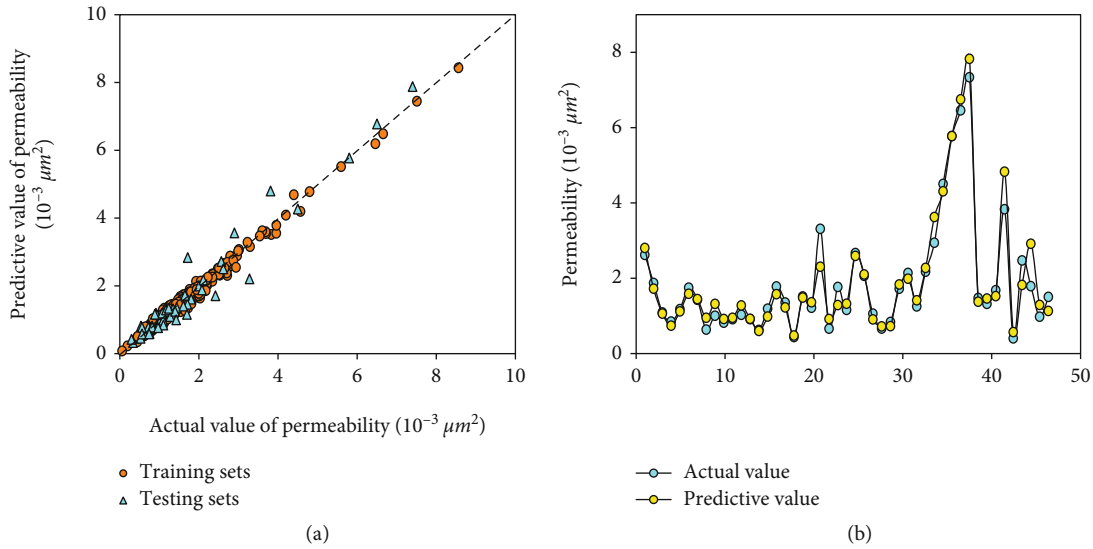

Figure 9: The results of reservoir permeability predicted by the CatBoost model. (a) The comparison of actual permeability and predicted permeability of training set and test set. (b) The prediction results of the test set.

Table 16: Evaluation results of CatBoost model index.

Sample data	MSE	RMSE	MAE	MAPE	R^2
Training sets	0.012	0.108	0.083	6.532	0.987
Testing sets	0.064	0.139	0.207	16.687	0.951

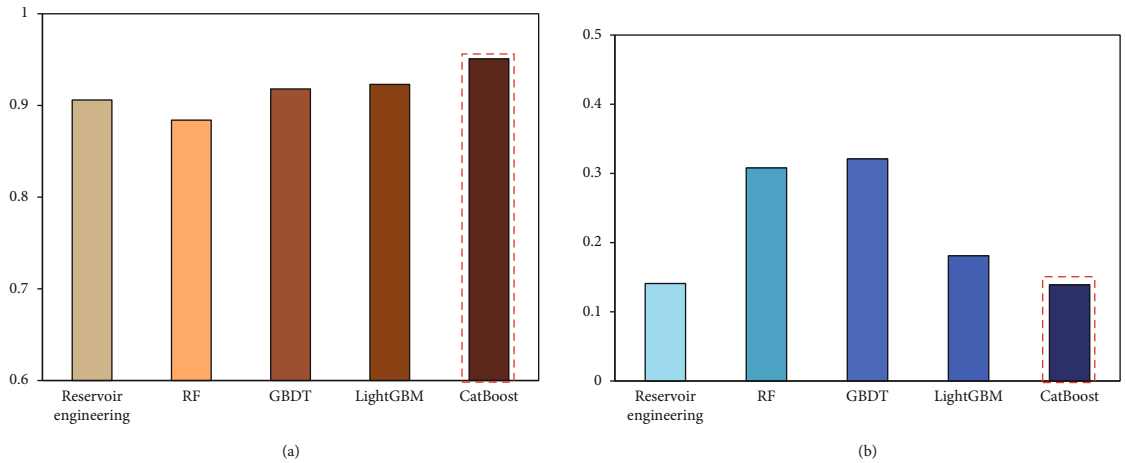

Figure 10: Comparison of index evaluation of different permeability prediction methods. (a) R^2 evaluation results. (b) The result of RMSE evaluation.

central region is lower than the high yield liquid, and the development effect is poor. In view of the above problems, this paper puts forward the corresponding development technology policy based on the current development status of the block:

(1) According to the predicted reservoir permeability contour map and liquid production volume contour map, the study block is divided into the noneffective area in the middle and the effective area in the east and west. For the well pattern infill in the central area, the well spacing between oil wells was shortened from the original 520 m to 173 m. At the same time, the corner wells in the diamond-shaped reverse nine-spot well pattern were transferred to form a row of water injection (Figure 12)

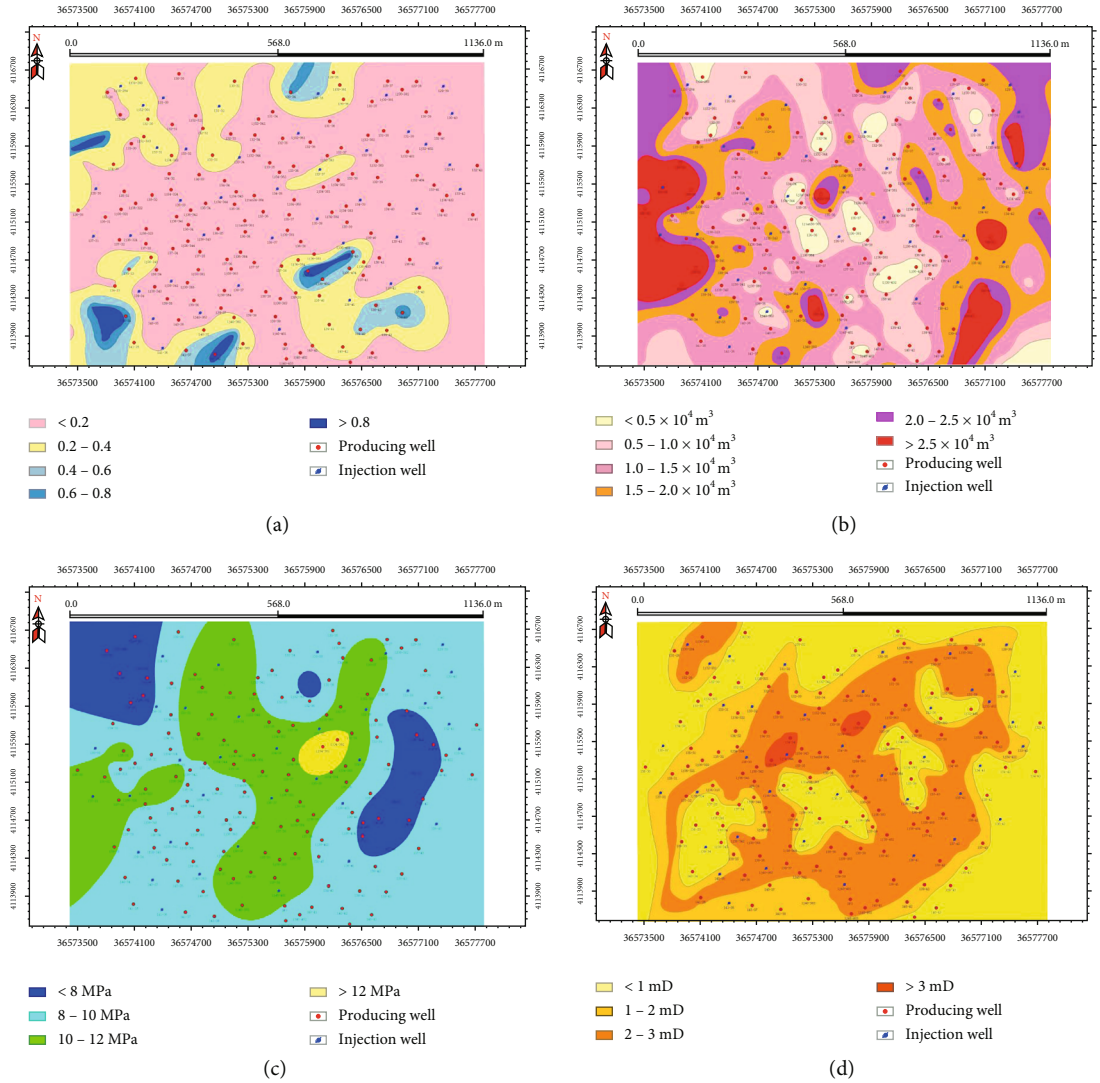

Figure 11: Isoline map. (a–d) The current reservoir permeability, formation pressure, water content, and liquid production isoline map.

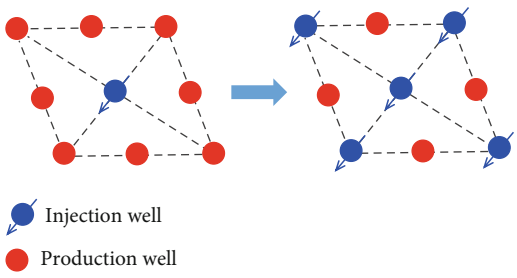

Figure 12: Diamond reverse nine-point well pattern corner well injection adjustment. Blue for the injection wells, and red for the production wells.

The adjustment scheme of well pattern in the eastern and western regions is based on the fixed point encryption, and the well pattern encryption is carried out in the local areas with poor development effect and low formation pressure maintenance level (wells lj36-341, lj40-363, lj38-366, 131-39, lj34-422, and 133-41), so as to improve the formation energy, the control degree of water wells, and the water flooding effect.

(2) For the wells not corresponding to the injection-production horizon in the study area, it is suggested to supplement the injection hole and improve the corresponding relationship of the injection-production horizon. The wells that need to improve the corresponding relationship between injection and production are mainly the following 12 wells. The number of supplementary injection wells and the supplementary injection horizon is shown in Table 17

(3) According to the results of part (1), the current average daily liquid production per well, daily oil production per well, and the average daily water injection per well of injection wells in each region are shown in Table 18

Table 17: Wells and layers needing replenishment.

Well number	Reinjection horizon	Well number	Reinjection horizon
l30-35	Yanchang-6	l36-32	Yanchang-5
l30-39	Yanchang-5/6	l36-41	Yanchang-3/4
l31-30	Yanchang-5/6	l36-42	Yanchang-6
l31-32	Yanchang-4/5/6	l36-45	Yanchang-6
l31-38	Yanchang-5/6	l37-37	Yanchang-4/5
l31-40	Yanchang-5/6	l37-38	Yanchang-5/6

Table 18: Average liquid production, oil production, and water injection per well in target area.

Partition	West effective area	Eastern effective area	Central ineffective area
Average daily liquid production of single well (m^3)	1.85	1.41	1.41
Average daily oil production per well (m^3)	0.94	1.00	0.56
Average daily water injection of single well (m^3)	17.40	16.80	18.30

Table 19: Reasonable water injection and liquid production in target area.

Partition	West effective area	Eastern effective area	Central ineffective area
Reasonable daily liquid production of single well (m^3)	3.01	2.64	2.01
Reasonable single well daily water injection (m^3)	18.60	17.60	16.80

According to the calculation formula of liquid production index in reservoir engineering manual, the reasonable water injection and reasonable liquid production in different regions of the area are shown in Table 19.

According to the calculation results (Tables 18 and 19), the water injection rate of the water injection wells in the target area is evaluated. The water injection rate is smaller in the middle and southwest and larger in the west, southeast, and north. The reasonable water injection rate in each region was determined by the calculation results, in which the reasonable water injection rate in the eastern effective region was 17.4 m^3/d, the reasonable water injection rate in the central effective region was 16.8 m^3/d, and the reasonable water injection rate in the western ineffective region was 18.3 m^3/d. Therefore, the water injection volume in the whole region is adjusted as follows.

In view of the ineffective area in the middle of the target block, the water injection rate of 9 wells (l33-34, l35-36, l36-38, l37-39, l38-40, l35-38, l33-37, l35-37, and lj34-383) was adjusted to 16.5 m^3/d.

In view of the effective area in the western part of the target block, the water injection rates of five water injection wells were adjusted. The water injection rates of well l36-31 were adjusted to 13 m^3/d, those of well l38-33 were adjusted to 12.5 m^3/d, and the water injection rates of well l39-36, well l36-34, and well l40-36 were adjusted to 18 m^3/d.

In the eastern effective area, 6 water injection wells were adjusted to 12.5 m^3/d for wells l30-40 and l34-41 and 18 m^3/d for wells lj32-401, l31-39, l30-39, and lj30-381.

7. Conclusion

In this paper, the original dynamic and static data are fully utilized to establish a permeability analogy system to find similar blocks by using big data analogy method, and the prediction results of reservoir engineering method and machine learning method are compared. Finally, the CatBoost model of machine learning method is selected to predict the permeability of 121 oil wells in the DLG block. The main conclusions are as follows:

(1) In this paper, the logging and production data are counted, and the Pearson and principal component analysis are used to screen the main controlling factors of permeability and establish the analogy system of six parameters including porosity, shale content, reservoir thickness, oil saturation, liquid production, and production pressure difference. Then, the fuzzy matter-element method is used to calculate the similarity between DLG block and the analogy block, and the blocks with similarity greater than 70% are selected as the similar blocks of DLG block

(2) According to the similarity calculation results, reservoir permeability of the DLG block is calculated by reservoir engineering method. In this method, the logging porosity of the DLG block is introduced into the correlation formula between core permeability and core porosity of QK-D7 block with the highest similarity, and the reservoir permeability of

the DLG block is calculated. Compared with 11 measured core permeability in the DLG block, the determination coefficient is 0.906, indicating that the method is accurate

(3) 421 oil wells in 8 blocks with similarity higher than 70% were selected as sample data, and machine learning method is used to predict reservoir permeability in the DLG block. Using randomized quadratic sampling cross-validation method to establish RF, GBDT, LightGBM, and CatBoost models, R^2, MSE, RMSE, MAE, and MAPE are used to evaluate the accuracy of the model. Results show that the CatBoost model has higher accuracy in predicting reservoir permeability, with R^2 of 0.951 and RMSE of 0.139. By comparing the reservoir engineering method, the CatBoost model is finally selected to predict reservoir permeability of 121 wells in the DLG block

(4) Using the predicted reservoir permeability, combined with the current production data such as formation pressure, water content, and liquid production in this block, contour maps are plotted, respectively. In view of the problems existing in the current mining process, the development adjustment measures such as increasing well spacing density, supplementary perforation, and rational proration are proposed, which is beneficial to the efficient development of the block and provides a new idea for the application of machine learning to predict oilfield parameters

Data Availability

The data used to support the findings of this study are included within the article. The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the Science Foundation of China University of Petroleum, Beijing (2462022BJRC004).

References

[1] H. Darcy, *Les Fontaines publiques de lo ville de Dijon*, Dalmont, Paris, 1856.

[2] P. M. Doyen, "Permeability, conductivity, and pore geometry of sandstone," *Journal of Geophysical Research*, vol. 93, no. B7, p. 7729, 1988.

[3] J. H. Baas, E. A. Hailwood, W. D. McCaffrey, M. Kay, and R. Jones, "Directional petrological characterisation of deep-marine sandstones using grain fabric and permeability anisotropy: methodologies, theory, application and suggestions for integration," *Earth-Science Reviews*, vol. 82, no. 1-2, pp. 101–142, 2007.

[4] K. Hoes, D. Dinescu, H. Sol, R. S. Parnas, and S. Lomov, "Study of nesting induced scatter of permeability values in layered reinforcement fabrics," *Composites Part A: Applied Science and Manufacturing.*, vol. 35, no. 12, pp. 1407–1418, 2004.

[5] Y. He, S. Cheng, S. Li et al., "A semianalytical methodology to diagnose the locations of underperforming hydraulic fractures through pressure-transient analysis in tight gas reservoir," *SPE Journal*, vol. 22, no. 3, pp. 924–939, 2017.

[6] C. C. Lim, N. Gowripalan, and V. Sirivivatnanon, "Microcracking and chloride permeability of concrete under uniaxial compression," *Cement and Concrete Composites.*, vol. 22, no. 5, pp. 353–360, 2000.

[7] D. Zhang and K. Li, "Concrete gas permeability from different methods: correlation analysis," *Cement and Concrete Composites.*, vol. 104, article 103379, 2019.

[8] M. Karaaslan, G. K. Wong, and A. Rezaei, "Reduced order model and global sensitivity analysis for return permeability test," *Journal of Petroleum Science and Engineering*, vol. 207, article 109064, 2021.

[9] S. Scholz, J. W. Gillespie, and D. Heider, "Measurement of transverse permeability using gaseous and liquid flow," *Composites Part A: Applied Science and Manufacturing*, vol. 38, no. 9, pp. 2034–2040, 2007.

[10] V. N. Dorovsky, Y. A. Nefedkin, A. I. Fedorov, and M. Y. Podberezhnyy, "A logging method for estimating permeability, velocity of second compressional wave, and electroacoustic constant in electrolyte-saturated porous formations," *Russian Geology and Geophysics*, vol. 51, no. 12, pp. 1285–1294, 2010.

[11] C. Wei, Y. Liu, Y. Deng, S. Cheng, and H. Hassanzadeh, "Analytical well-test model for hydraulicly fractured wells with multiwell interference in double porosity gas reservoirs," *Journal of Natural Gas Science and Engineering*, vol. 103, p. 104624, 2022.

[12] C. Wei, Y. Liu, Y. Deng, S. Cheng, and H. Hassanzadeh, "Temperature transient analysis of naturally fractured geothermal reservoirs," *SPE Journal*, vol. 1-23, pp. 1–23, 2022.

[13] Y. He, J. Qin, S. Cheng, and J. Chen, "Estimation of fracture production and water breakthrough locations of multi- stage fractured horizontal wells combining pressure-transient analysis and electrical resistance tomography," *Journal of Petroleum Science and Engineering*, vol. 194, article 107479, 2020.

[14] S. Siddiqui, S. Talabani, S. T. Saleh, and M. R. Islam, "Foam flow in low-permeability Berea Sandstone cores: a laboratory investigation," *Journal of Petroleum Science and Engineering*, vol. 36, no. 3–4, pp. 133–148, 2002.

[15] Y. Ozaki, "Effect of compaction behaviour of iron powders on permeability of iron cores," *Metal Powder Report*, vol. 57, no. 6, p. 54, 2002.

[16] Y. He, S. Cheng, J. Qin et al., "Interference testing model of multiply fractured horizontal well with multiple injection wells," *Journal of Petroleum Science and Engineering*, vol. 176, pp. 1106–1120, 2019.

[17] G. Zhang, Z. Wang, S. Mohaghegh, C. Lin, Y. Sun, and S. Pei, "Pattern visualization and understanding of machine learning models for permeability prediction in tight sandstone reservoirs," *Journal of Petroleum Science and Engineering*, vol. 200, article 108142, 2021.

[18] T. Zhongyuan, J. Aming, Y. Weilin, L. Yingming, and G. Shuangsheng, "Resistivity correction for drilling fluid

invasion using LWD and wire-line logging data: a case from high-porosity and low-permeability carbonate reservoirs, DLL Oilfield, Oman," *Petroleum Exploration and Development*, vol. 37, no. 4, pp. 430–437, 2010.

[19] Z. Zhang, H. Zhang, J. Li, and Z. Cai, "Permeability and porosity prediction using logging data in a heterogeneous dolomite reservoir: an integrated approach," *Journal of Natural Gas Science and Engineering*, vol. 86, article 103743, 2021.

[20] H. Hamdi, "Well-test response in stochastic permeable media," *Journal of Petroleum Science and Engineering*, vol. 119, pp. 169–184, 2014.

[21] H. N. Dmour, "Practical well test analysis of a hydraulically fractured low permeability gas reservoir: a case history," *Journal of King Saud University-Engineering Sciences.*, vol. 20, no. 1, pp. 47–59, 2008.

[22] W. Bottomley, J. Schouten, E. McDonald, and T. Cooney, "Novel well test design for the evaluation of complete well permeability and productivity for CSG wells in the Surat Basin," *Journal of Natural Gas Science and Engineering*, vol. 33, pp. 1002–1009, 2016.

[23] R. Aguilera, "Incorporating capillary pressure, pore throat aperture radii, height above free-water table, and winland r 35 values on Pickett plots," *AAPG Bulletin*, vol. 86, no. 4, pp. 605–624, 2002.

[24] J. O. Amaefule, M. Altunbay, D. Tiab, D. G. Kersey, and D. K. Keelan, "Enhanced reservoir description: using core and log data to identify hydraulic (flow) units and predict permeability in uncored intervals/wells," in *Proceedings SPE Annual Technical Conference and Exhibition*, Houston, Texas, 1993.

[25] M. R. Rezaee, H. Motiei, and E. Kazemzadeh, "A new method to acquire *m* exponent and tortuosity factor for microscopically heterogeneous carbonates," *Journal of Petroleum Science and Engineering*, vol. 56, no. 4, pp. 241–251, 2007.

[26] F. Song, J.-g. Hou, and S. Ni-na, "Model building for Chang-8 low permeability sandstone reservoir in the Yanchang formation of the Xifeng oil field," *Mining Science and Technology*, vol. 19, no. 2, pp. 245–251, 2009.

[27] F. T. T. Huque, K. Box, J. A. Platts, and J. Comer, "Permeability through DOPC/dodecane membranes: measurement and LFER modelling," *European Journal of Pharmaceutical Sciences*, vol. 23, no. 3, pp. 223–232, 2004.

[28] G. Wang, K. Wang, S. Wang, D. Elsworth, and Y. Jiang, "An improved permeability evolution model and its application in fractured sorbing media," *Journal of Natural Gas Science and Engineering*, vol. 56, pp. 222–232, 2018.

[29] M.-A. Ahmadi, M. R. Ahmadi, S. M. Hosseini, and M. Ebadi, "Connectionist model predicts the porosity and permeability of petroleum reservoirs by means of petro-physical logs: application of artificial intelligence," *Journal of Petroleum Science and Engineering*, vol. 123, pp. 183–200, 2014.

[30] Z. Chao, G. Ma, K. He, and M. Wang, "Investigating low-permeability sandstone based on physical experiments and predictive modeling," *Underground Space*, vol. 6, no. 4, pp. 364–378, 2021.

[31] H. A. Nooruddin and M. E. Hossain, "Modified Kozeny-Carmen correlation for enhanced hydraulic flow unit characterization," *Journal of Petroleum Science and Engineering*, vol. 80, no. 1, pp. 107–115, 2011.

[32] A. M. Handhal, F. R. Ettensohn, A. M. Al-Abadi, and M. J. Ismail, "Spatial assessment of gross vertical reservoir heterogeneity using geostatistics and GIS-based machine-learning classifiers: a case study from the Zubair Formation, Rumaila oil field, southern Iraq," *Journal of Petroleum Science and Engineering*, vol. 208, article 109482, 2022.

[33] C. Bravo, L. Saputelli, F. Rivas et al., "State of the art of artificial intelligence and predictive analytics in the E&P industry: a technology survey," *SPE Journal*, vol. 19, no. 4, pp. 547–563, 2014.

[34] J. Cai, H. Hajibeygi, J. Yao, and S. M. Hassanizadeh, "Advances in porous media science and engineering from InterPore2020 perspective," *Advances in Geo-Energy Research.*, vol. 4, no. 4, pp. 352–355, 2020.

[35] D. A. Wood, "Predicting porosity, permeability and water saturation applying an optimized nearest-neighbour, machine-learning and data-mining network of well-log data," *Journal of Petroleum Science and Engineering*, vol. 184, article 106587, 2020.

[36] E. Bahonar, M. Chahardowli, Y. Ghalenoei, and M. Simjoo, "New correlations to predict oil viscosity using data mining techniques," *Journal of Petroleum Science and Engineering.*, vol. 208, article 109736, 2022.

[37] Z. Ghahramani, "Probabilistic machine learning and artificial intelligence," *Nature*, vol. 521, no. 7553, pp. 452–459, 2015.

[38] J. B. Montgomery and F. M. O'Sullivan, "Spatial variability of tight oil well productivity and the impact of technology," *Applied Energy.*, vol. 195, pp. 344–355, 2017.

[39] Y. Wang, S. Cheng, F. Zhang et al., "Big data technique in the reservoir parameters' prediction and productivity evaluation: a field case in western South China sea," *Gondwana Research*, vol. 96, no. 2021, pp. 22–36, 2021.

[40] Q. Zhou, R. Dilmore, A. Kleit, and J. Y. Wang, "Evaluating gas production performances in marcellus using data mining technologies," *Journal of Natural Gas Science and Engineering*, vol. 20, pp. 109–120, 2014.

[41] F. Male, J. L. Jensen, and L. W. Lake, "Comparison of permeability predictions on cemented sandstones with physics-based and machine learning approaches," *Journal of Natural Gas Science and Engineering*, vol. 77, article 103244, 2020.

[42] H. Al Khalifah, P. W. J. Glover, and P. Lorinczi, "Permeability prediction and diagenesis in tight carbonates using machine learning techniques," *Marine and Petroleum Geology*, vol. 112, article 104096, 2020.

[43] G. Zhang, Z. Wang, H. Li, Y. Sun, Q. Zhang, and W. Chen, "Permeability prediction of isolated channel sands using machine learning," *Journal of Applied Geophysics*, vol. 159, pp. 605–615, 2018.

[44] O. Sudakov, E. Burnaev, and D. Koroteev, "Driving digital rock towards machine learning: predicting permeability with gradient boosting and deep neural networks," *Computers & Geosciences*, vol. 127, pp. 91–98, 2019.

[45] Y. Wang, H. Li, X. Jianchun, S. Liu, and X. Wang, "Machine learning assisted relative permeability upscaling for uncertainty quantification," *Energy*, vol. 245, article 123284, 2022.

[46] Z. Ma, J. Y. Leung, S. Zanon, and P. Dzurman, "Practical implementation of knowledge-based approaches for steam-assisted gravity drainage production analysis," *Expert Systems with Applications*, vol. 42, no. 21, pp. 7326–7343, 2015.

[47] Z. Yuan, W. Qin, and J. Zhao, "Smart manufacturing for the oil refining and petrochemical industry," *Engineering*, vol. 3, no. 2, pp. 179–182, 2017.

[48] W. J. Al-Mudhafar, "Integrating lithofacies and well logging data into smooth generalized additive model for improved permeability estimation: Zubair formation, South Rumaila oil field," *Marine Geophysical Research*, vol. 40, no. 3, pp. 315–332, 2019.

[49] J. G. Urang, E. D. Ebong, A. E. Akpan, and E. I. Akaerue, "A new approach for porosity and permeability prediction from well logs using artificial neural network and curve fitting techniques: a case study of Niger Delta, Nigeria," *Journal of Applied Geophysics*, vol. 183, article 104207, 2020.

[50] A. Abdulraheem, E. Sabakhy, M. Ahmed, A. Vantala, P. D. Raharja, and G. Korvin, "Estimation of permeability from wireline logs in a Middle Eastern carbonate reservoir using fuzzy logic," in *SPE Middle East Oil and Gas Show and Conference*, Bahrain, 2007.

[51] W. J. Al-Mudhafar, "Integrating machine learning and data analytics for geostatistical characterization of clastic reservoirs," *Journal of Petroleum Science and Engineering*, vol. 195, article 107837, 2020.

[52] W. J. Al-Mudhafar, "Integrating well log interpretations for lithofacies classification and permeability modeling through advanced machine learning algorithms," *Journal of Petroleum Exploration and Production Technology*, vol. 7, no. 4, pp. 1023–1033, 2017.

[53] P. E. Lacentre and P. M. Carrica, "A method to estimate permeability on uncored wells based on well logs and core data," in *SPE Latin American and Caribbean Petroleum Engineering Conference*, Trinidad and Tobago, 2003.

[54] D. Wang and R. S. Seright, "Examination of literature on colloidal dispersion gels for oil recovery," *Petroleum Science*, vol. 18, no. 4, pp. 1097–1114, 2021.

[55] Y. Wang and L. F. Ayala, "Explicit determination of reserves for variable-bottomhole-pressure conditions in gas rate-transient analysis," *SPE Journal*, vol. 25, no. 1, pp. 369–390, 2020.

[56] S. H. Lee, A. Kharghoria, and A. Datta-Gupta, "Electrofacies characterization and permeability predictions in complex reservoirs," *SPE Reservoir Evaluation & Engineering*, vol. 5, no. 3, pp. 237–248, 2002.

[57] W. J. Al-Mudhafar and D. A. Wood, "Tree-based ensemble algorithms for lithofacies classification and permeability prediction in heterogeneous carbonate reservoirs," in *Offshore Technology Conference*, Houston, Texas, USA, 2022.

[58] A. F. Al-Anazi and I. D. Gates, "Support-vector regression for permeability prediction in a heterogeneous reservoir: a comparative study," *SPE Reservoir Evaluation & Engineering*, vol. 13, no. 3, pp. 485–495, 2010.

[59] H. H. Perez, A. Datta-Gupta, and S. Mishra, "The role of electrofacies, lithofacies, and hydraulic flow units in permeability prediction from well logs: a comparative analysis using classification trees," *SPE Reservoir Evaluation & Engineering*, vol. 8, no. 2, pp. 143–155, 2005.

[60] I. S. Nashawi and A. Malallah, "Permeability prediction from wireline well logs using fuzzy logic and discriminant analysis," in *Proceedings SPE Asia Pacific Oil and Gas Conference and Exhibition*, Brisbane, Queensland, Australia, 2010.

[61] S. Pan, Z. Zheng, Z. Guo, and H. Luo, "An optimized XGBoost method for predicting reservoir porosity using petrophysical logs," *Journal of Petroleum Science and Engineering*, vol. 208, no. C, article 109520, 2022.

[62] F. Anifowose, A. Abdulraheem, and A. Al-Shuhail, "A parametric study of machine learning techniques in petroleum reservoir permeability prediction by integrating seismic attributes and wireline data," *Journal of Petroleum Science and Engineering*, vol. 176, pp. 762–774, 2019.

[63] S. Sen, M. Abioui, S. S. Ganguli et al., "Petrophysical heterogeneity of the early Cretaceous Alamein dolomite reservoir from North Razzak oil field, Egypt integrating well logs, core measurements, and machine learning approach," *Fuel*, vol. 306, article 121698, 2021.

[64] B. Rafik and B. Kamel, "Prediction of permeability and porosity from well log data using the nonparametric regression with multivariate analysis and neural network, Hassi R'Mel Field, Algeria," *Egyptian Journal of Petroleum*, vol. 26, no. 3, pp. 763–778, 2017.

[65] X. Zhao, X. Chen, Q. Huang, Z. Lan, X. Wang, and G. Yao, "Logging-data-driven permeability prediction in low-permeable sandstones based on machine learning with pattern visualization: a case study in Wenchang A Sag, Pearl River Mouth Basin," *Journal of Petroleum Science and Engineering*, vol. 214, article 110517, 2022.

[66] D. Pavanello, W. Zaaiman, A. Colli, J. Heiser, and S. Smith, "Statistical functions and relevant correlation coefficients of clearness index," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 130-131, pp. 142–150, 2015.

[67] J. L. Rodgers and W. A. Nicewander, "Thirteen ways to look at the correlation coefficient," *The American Statistician*, vol. 42, no. 1, pp. 59–66, 1988.

[68] S. K. Tyagi, "Correlation coefficient of dual hesitant fuzzy sets and its application to multiple attribute decision making," *Applied Mathematical Modelling*, vol. 38, no. 2, pp. 659–666, 2014.

[69] W. J. Al-Mudhafar, "Incorporation of bootstrapping and cross-validation for efficient multivariate facies and petrophysical modeling," in *SPE Low Perm Symposium*, Denver, CO, USA, 2016.

[70] M. Pirrone, A. Battigelli, and L. Ruvo, "Lithofacies classification of thin-layered turbidite reservoirs through the integration of core data and dielectric-dispersion log measurements," *SPE Reservoir Evaluation & Engineering*, vol. 19, no. 2, pp. 226–238, 2016.

[71] G. Wang, Y. Ju, T. R. Carr, and G. Cheng, "Application of artificial intelligence on black shale lithofacies prediction in Marcellus Shale, Appalachian Basin," in *Unconventional Resources Technology Conference*, pp. 1970–1980, Denver, CO, USA, 2014.