Quantitative Method of Classification and Discrimination of a Porous Carbonate Reservoir Integrating K-means Clustering and Bayesian Theory

FANG Xinxin[1, 2, *], ZHU Guotao[3], YANG Yiming[3], LI Fengling[3] and FENG Hong[2]

[1] *China Coal Research Institute, Beijing 100013, China*
[2] *Xi'an Research Institute of China Coal Research Institute, Xi'an 710077, China*
[3] *Oilfield Technology Service Company of Xinjiang Oilfield Company, PetroChina, Karamay, Xinjiang 834000, China*

Abstract: Reservoir classification is a key link in reservoir evaluation. However, traditional manual means are inefficient, subjective, and classification standards are not uniform. Therefore, taking the Mishrif Formation of the Western Iraq as an example, a new reservoir classification and discrimination method is established by using the K-means clustering method and the Bayesian discrimination method. These methods are applied to non-cored wells to calculate the discrimination accuracy of the reservoir type, and thus the main reasons for low accuracy of reservoir discrimination are clarified. The results show that the discrimination accuracy of reservoir type based on K-means clustering and Bayesian stepwise discrimination is strongly related to the accuracy of the core data. The discrimination accuracy rate of Type I, Type II, and Type V reservoirs is found to be significantly higher than that of Type III and Type IV reservoirs using the method of combining K-means clustering and Bayesian theory based on logging data. Although the recognition accuracy of the new methodology for the Type IV reservoir is low, with average accuracy the new method has reached more than 82% in the entire study area, which lays a good foundation for rapid and accurate discrimination of reservoir types and the fine evaluation of a reservoir.

Key words: upstream, resource exploration, reservoir classification, carbonate, K-means clustering, Bayesian discrimination, Cenomanian–Turonian, Iraq

Citation: Fang et al., 2023. Quantitative Method of Classification and Discrimination of a Porous Carbonate Reservoir Integrating K-means Clustering and Bayesian Theory. Acta Geologica Sinica (English Edition), 97(1): 176–189. DOI: 10.1111/1755-6724.14941

1 Introduction

The Cretaceous Mishrif Formation in the W Oilfield in the Mesopotamian sub-basin of Iraq (Fig. 1a–c) is a typical porous carbonate reservoir, which is clearly controlled by sedimentary facies belts (Gao et al., 2013). This reservoir has been seriously reformed after diagenesis, and its heterogeneity and anisotropy are extremely great. This complexity of its own structure and composition increases the difficulty of reservoir classification and recognition (Wang et al., 2019).

There have been many previous studies on the classification of carbonate reservoirs, which can essentially be attributed to two categories (Chen et al., 2018): the first is based on the type and quality of reservoir space, i.e., good, medium or poor. For example, Jodry (1972) classified and evaluated carbonate reservoirs based on the relationship between pore structure and rock types as non-reservoir, poor reservoir, medium reservoir and good ones (Jodry, 1972; Zhao and Liu, 2018); the other is based on genesis, i.e., the type of lithofacies assemblage (Szabó and Nehéz, 2019). For example, according to the evolutionary history of the carbonate rock and its main geological factors, Feng et al. (1995) divided it into five types: granular shoal, dolomitization-reef, cavernous dolomite, paleo weathering dissolution, and fracture (Feng et al., 1995; Yi and Chong, 2018). Smith et al. (2003) and Awoleke and Lane (2011) divided carbonate reservoirs into limestone and dolomite types according to lithology. Wang and Zhang (2019) divided the world's carbonate reservoirs into six types: below unconformity, dolomite, oolitic and aggregate shoal, reef, micro-porous, and micro-fracture. These classification methods are mainly based on geological descriptors, which are little combined with statistical algorithms, and are suitable for reservoir classification for a single well. However, for many wells, these classification methods do not describe the reservoir types clearly. Therefore, petrophysical and petrographical studies, along with reservoir quality index (RQI) and reservoir flow indicator (FZI), which are characterized by the effective pore radius (r35), were employed to solve these problems (El Sharawy et al., 2016, 2019; Nabawy et al., 2018; Abuamarah and Nabawy, 2021). However, these methods are more applicable to sandstone than carbonate, and their precision also depends on Mercury Injection Capillary Pressure (MICP) experimental data.

In recent years, with the development of artificial

* Corresponding author. E-mail: fx15827573109@163.com

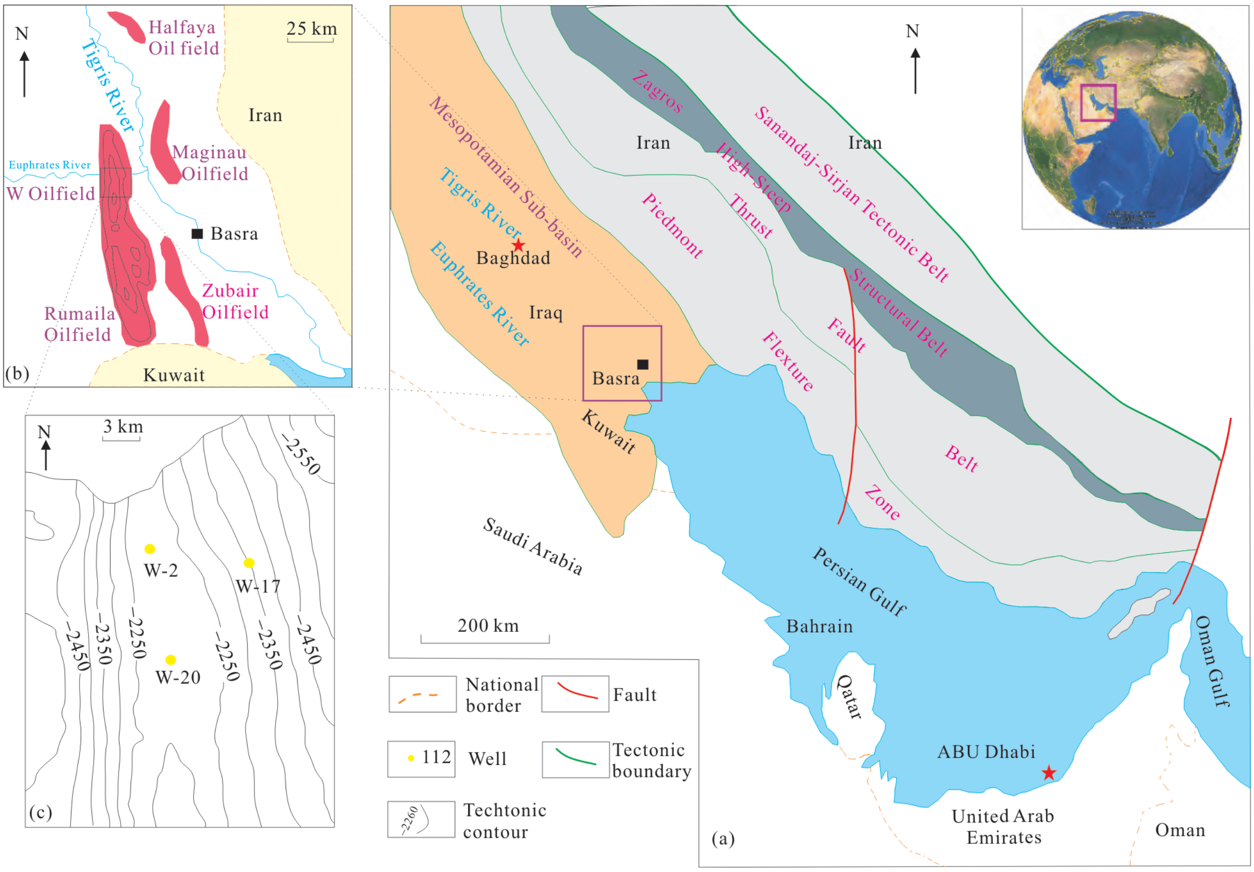

Fig. 1. Structural location map of the Western oilfield in Iraq.
(a) Location map and tectonic unites of the Arabian Plate, the West Qurna oilfield belongs to the Middle Arabian Basin; (b) geographic location of the W oilfield; (c) the top structural map of the Mishrif Formation, the W oilfield.

intelligence technology, data mining algorithms such as cluster analysis have been adopted for reservoir classification (Wang and Yang, 2018). Cluster analysis, also known as group analysis, based on similarity, is a statistical analysis method to study (sample or index) classification (Xun and Yu, 2008). Luo and Ren (1999) studied the recognition of fracture-vuggy carbonate reservoirs by using cluster analysis and the standard back propagation (BP) neural network method respectively. Sima et al. (2012) used standard BP neural network method to achieve the recognition and prediction of clastic reservoir flow units in the first member of the Funing Formation of Well Fang 4 in the Huangyu Oilfield, Jiangsu Province. The statistical methods commonly used at present include the fuzzy clustering, BP neural network, Bayes stepwise discriminant and its derivative methods, such as: Rogistiv discriminant, multivariate statistical, multiple population stepwise discriminant analysis, self-organizing neural grid, etc. (Abdulaziz et al., 2019; Khan and Rehman, 2021). Compared with the fuzzy clustering and BP neural network methods, the Bayes stepwise discriminant method is a statistical analysis method that integrates effective parameter selection and quantitative recognition functions (Liao and Zhang, 2004; Sames and Saussus, 2010; Rimstad and Avseth, 2012; Zhang and Du, 2021). By comparing the posterior probabilities of different types of reservoir samples, the attribution of the samples can be judged with high recognition accuracy and good stability.

At present, there are many difficulties in the classification and recognition of carbonate reservoirs in the Mishrif Formation in W Oilfield. First of all, it is difficult to establish a classification criterion that can be applied to non-cored wells by analysis of a core well and popularized. Secondly, the workload of reservoir classification is large and the efficiency of manual classification is low. Thirdly, the effect of classification is lack of intuitive verification methods. Therefore, in this study, based on the core, logging, and logging data of 10 cored wells, and the physical properties of the reservoir including sedimentation and diagenesis, we investigate and propose a method for automatic classification and verification of reservoir types based on the theory of K-means clustering and Bayesian discrimination with the aim of improving the efficiency of reservoir classification and the accuracy of recognition.

2 Geological Setting

The Persian Gulf Basin is developed on the Arabian plate. It has been filled with sediments since the late Precambrian, and has experienced three stages of evolution: stable craton, passive continental margin, and foreland compression. From east to west, the basin can be divided into the Sanandaj–Sirjan structural belt, the Zagros high-steep structural belt, the thrust fault belt, the piedmont flexure belt, and the Mesopotamian sub-basin (Al-Sakini, 2010; Kakeman et al., 2021). The W Oilfield is located about 50 km northwest of Basra, Iraq. The regional tectonic location belongs to the Mesopotamian sub-basin in the northern Persian Gulf Basin, and the northeastern part is adjacent to the piedmont flexure structural belt (Fig. 1a, b). The Mesopotamian sub-basin is relatively weak in tectonic activity, and the W Oilfield has a simple structure (Thehni, 1998). It belongs to a long-axis anticline with a nearly north–south spread and gentle faults (Aqrawi et al., 1998) (Fig. 1a, c). The oil field is adjacent to the Rumaila oil field to the south and is bounded by the Euphrates River to the north (Fig. 1b).

Source rock: Southeastern Iraq was tectonically stable during the deposition of the Mishrif Formation, so the structure of the deposits was gentle and asymmetric, the western flank being steeper than eastern flank, with no fault (Aqrawi et al., 1998; Sun et al., 2013; Mahdi and Aqrawi, 2014). The Mishrif Formation is controlled by the sedimentary cycles, recording a long term, second-order shallowing-upward cycle with a regional unconformity sedimentary interface at the top contact with the Khasib Formation, and the bottom has an integrated conformable contact with the underlying Rumaila Formation (Fig. 2b). The Mishrif Formation comprises two third-order sequences with potential ties to two flooding surfaces (Sharland et al., 2001), with carbonate rocks developed in a Middle Cretaceous passive continental margin depositional environment. During deposition, the climate in the area was warm and humid, with many crustaceans developed. Consequently, the carbonate rocks of the Mishrif Formation generally contained bio-gravel or clastics. The formation is currently buried at a depth of about 2400 m, and the formation thickness is about 360 m (Fig. 2a, b). According to the combination of characteristics of lithology and lithofacies, from bottom to top, it can be divided into six layers: CRI, CRII, mB1, upper mB2 and lower mB2 (Fig. 2b).

3 Methodology and Available Data

3.1 Collected data

A cumulative 230 core samples have been collected to prepare thin-sections. Pore throat structure analysis was based on high-pressure mercury injection capillary pressure (HPMI). Helium porosity and air permeability were measured on the 230 core samples at lab. Logging data came from cored wells W-171, W-118. Non cored wells W-271, W-19, W-17, W-218, W-239, W-131, W-58, and W-190 were chosen as test wells. All studied thin sections were stained with a mixture of alizarin red S and potassium ferricyanide to differentiate calcite and dolomite. In addition, these thin sections were impregnated with blue-dyed resin to characterize pore type and rock structure. Classification method by Dunham (1962) was selected to distinguish carbonate rock type, which is based on mud– or grain-supported types. This classification stresses the grain-to-matrix relationship, based on relative content of composition, highly generalized depositional characteristics and rock fabrics, in addition to reflecting the hydrodynamic conditions and genesis.

3.2 Methodology

The recognition of reservoir types includes reservoir classification and verification of classification results. Based on core, thin section and physical property parameters of cored wells, reservoir classification based on the geology was performed. Then the qualitative and quantitative relationship between the well logging curve and each type of reservoir was studied, and then the reservoir logging classification standard was established by the K-means clustering method. The discriminant formula of each reservoir type was established using the Bayesian discriminant rule. The logging classification criteria and verification formula of each reservoir type were applied to non-cored wells, and then the reservoir type recognition of non-cored wells was achieved. The method flow consists of four key steps (Fig. 3).

Abbreviations used in text and figures: BP-back propagation; DT-acoustic time difference; FZI-reservoir flow indicator; GR-natural gamma ray; ILD-deep lateral resistivity; ILM-medium lateral resistivity; NPHI-neutron porosity; OHMM-unit of resistivity; RHOB-density; RQI-reservoir quality index; Type I, II, III, etc.-reservoir type.

3.2.1 Reservoir classification of cored wells

Core, thin section and physical property data can reflect reservoir lithology, porosity and permeability characteristics and pore space characteristics, which are important reference data for reservoir quality evaluation. Based on the analysis of cored well data, a reservoir classification of cored well was carried out, which then laid a foundation for reservoir classification and recognition of non-cored wells.

While classifying cored wells with core, thin section, and physical property data, it is also necessary to combine the logging characteristics corresponding to various types of reservoirs. This is done with the aid of qualitative or quantitative relationships between reservoirs and logging curves, to help reservoir classification, forming a relatively accurate reservoir classification result for cored wells that can be discriminated based on logging data.

3.2.2 Establishment of reservoir classification criteria based on logging parameters

Logging parameters are an important bridge for comparison and analysis between cored well and non-cored well. According to the classification results of cored wells with the geological data, the quantitative characterization criteria of a single well is established by using the logging data, and then applied to the non-cored wells with the same type of logging data for reservoir recognition. The specific steps are as follows:

(1) in order to eliminate the systematic error between logging data of different time and different instruments, it is necessary to standardize the logging parameters (Umer et al., 2019);

(2) selection of sensitive logging curve types. There are two main bases for optimizing the types of logging curves.

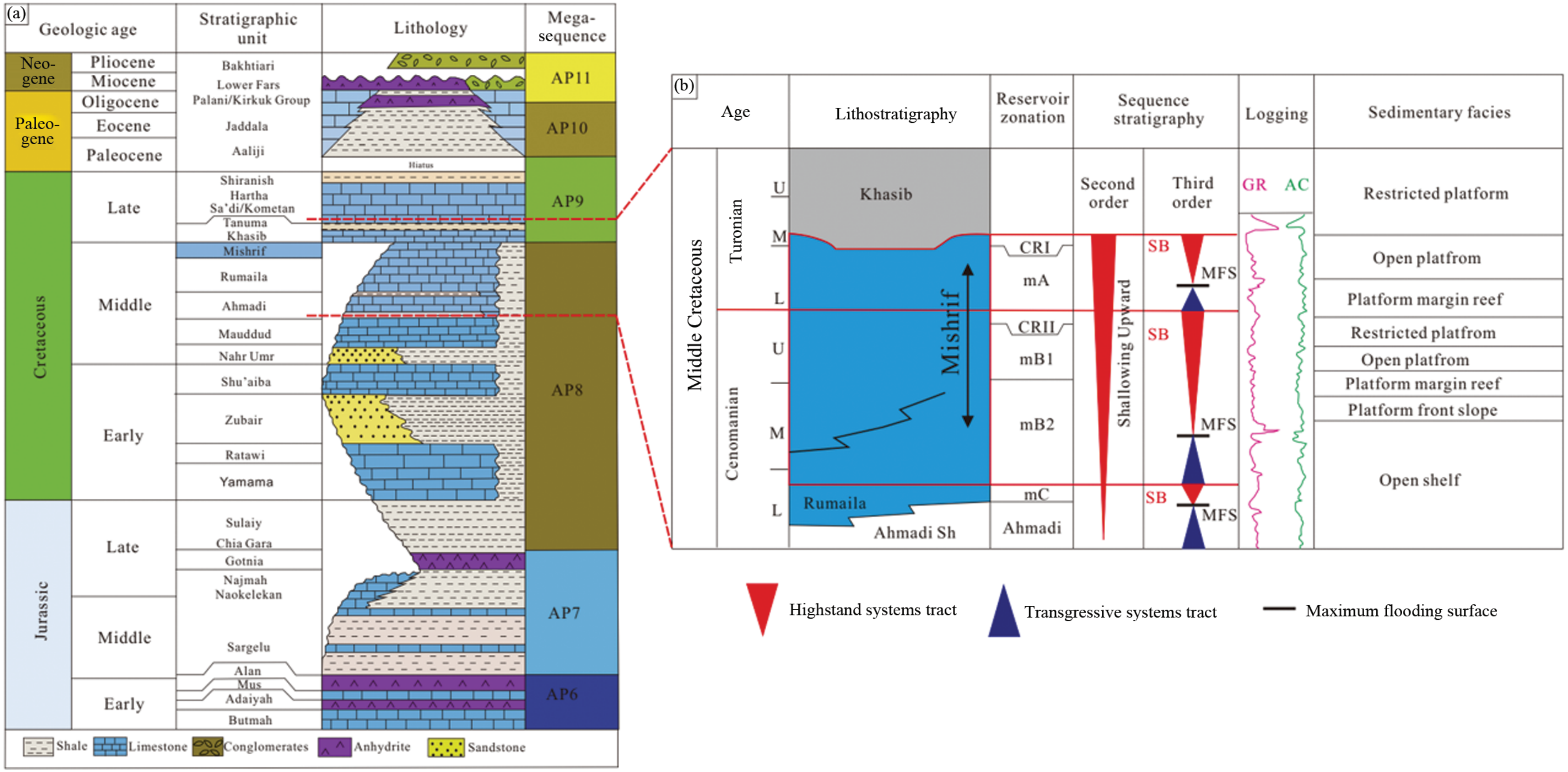

Fig. 2. Stratigraphic column and sequence stratigraphic map for the Mesopotamian basin.
(a) Stratigraphic column; (b) sequence stratigraphic and reservoir zonation of the Mishrif Formation in W oilfield. Notes: CRI, CRII, mA, mB1, mC–unit of formation; AC–acoustic time difference; GR–natural gamma ray; MFS–maximum flooding surface; SB–sequence boundary.

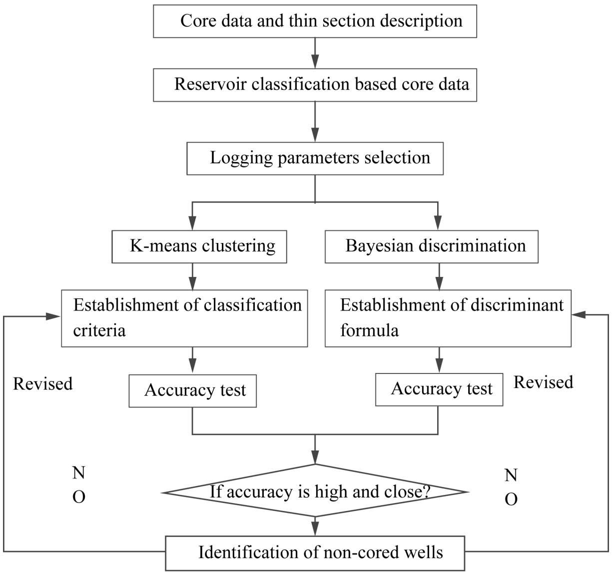

Fig. 3. Reservoir type recognition process for a single well.

One is the geological meaning expressed by a single logging curve. The second is the effect of several combinations of logging curves to distinguish different types of reservoirs. The optimized logging curves can reflect the sedimentary characteristics of the reservoir, being able to effectively distinguish the reservoir types by reasonable combination.

(3) establishment of reservoir logging classification standard. For the study of reservoir classification, the classification scheme suitable for the study area is a unified research platform, and each reservoir type is uniquely determined. Therefore, this paper adopts the classical clustering algorithm that can give the number of clustering, i.e., K-means clustering, which is a clustering analysis algorithm for iterative solution. The specific steps are as follows:

• determine the number of clustering, i.e., K;

• select a vector from the data set as the initial cluster center, i.e., B_1, B_2, B_3...B_K. The vector value of the cluster center can be set randomly, and its value will impact the clustering result;

• assign the samples that need to be classified X_i (i = 1,2,3…,n) one by one to a certain clustering center B_j (0 < j ≤ k),

$$\| X_i - B_j \| = \min_{1<s<k} \| X_i - B_s \| ;$$

• calculate the new vector value of each clustering center

$$B_j = \frac{1}{N_j} \sum_{X \in S_j} X ,$$

where N_j is the number of samples contained in the j-th clustering domain S_j; and

• if the clustering center no longer changes, terminate the process, otherwise return to step (3).

(4) reservoir logging classification criteria are established. According to the logging classification criteria initially established by the K-means clustering method, the reservoir of the cored well is divided. Next, calculate the accuracy rate, compared with the classification results based on core data. Then continue to modify the classification criteria until the accuracy rate meets the requirements, and establish the final reservoir logging classification criteria.

3.2.3 Establishment of discriminant formula of reservoir types

After establishing the logging classification criteria of a reservoir, it is necessary to establish a reasonable verification method to check the accuracy of the classification results. This paper adopts Bayesian discrimination method.

The Bayesian discrimination method constructs a discriminant function with the determined variable data, so that the function has some optimal properties to obtain the posterior probability of the unknown variable, so as to distinguish as much as possible the sample points belonging to different categories (Larsen and Ulvmoen, 2006; Wang et al., 2014; Liu and Chen, 2016).

The principle to establish the discriminant formula of reservoir types is as follows: if there are M parent classes, then L samples shall be taken, and each sample shall belong to one of the M parent groups. If each sample has p observation indices (x_1, x_2, x_3...x_p), then each sample can be regarded as a point in the p-dimensional space {Q}, and L samples constitute a p-dimensional space {T}. Simultaneously, each sample is regarded as an independent normal random vector, and then the m-th class x_m {m = 1,2...M} is a multivariate normal distribution. If there is a class from a certain new sample x= (x_1, x_2, x_3..., x_p), then the posterior probability of sample x belonging to the m-th class can be calculated according to Bayesian.

$$p(\mathrm{m}) = \frac{p(\mathrm{x|m})\, p(\mathrm{m})}{\sum_{\mathrm{j=1}}^{\mathrm{M}} p(\mathrm{x|M_j})\, p(\mathrm{M_j})} \tag{1}$$

where p(m) is the prior probability of the m-th class. p(k|m) is the probability density function when x belonging to the m-th class. $p(M_i)$ is the prior probability of the M_j-th data point. $p(x|M_j)$ is the probability density function when x belonging to the M_j-th data point.

Among the calculated M posterior probabilities, if the posterior probability p(k|x) when x belongs to the k-th class is the largest, then the sample x is classified into the k-th class.

The discriminant formula of the probability density function of the m-th class is as follows:

$$p(\mathrm{x|m}) = \beta_{\mathrm{m}} + \sum_{\mathrm{i=1}}^{\mathrm{p}} \beta_{\mathrm{m}} x_{\mathrm{mi}} \quad \mathrm{i} = 1,2,3...p \tag{2}$$

where β_m is the discriminating coefficient and x_{mi} is the i-th observation index of the m-th class.

According to the reservoir classification results obtained by the classification standard, the logging data are respectively substituted into the discriminant formulas of various types of reservoirs to obtain their posterior probabilities, which can further judge the types of reservoirs and calculate the accuracy. Subsequently, the discriminant formula is adjusted until the accuracy meets the requirements. Finally, the final discriminant formula of reservoir type is determined.

3.2.4 Recognition of reservoir types in non-cored wells

Discrimination criteria of each reservoir is applied to non-cored wells to achieve recognition results of reservoir types from non-cored wells. Next, the Bayesian discriminant formula is applied to verify various types of reservoirs, obtaining the type attribution of all logging data points and the accuracy rate. Reservoir units with low accuracy can be modified manually based on expert experience until the accuracy meets the requirements, and, finally, the non-cored well reservoir type recognition is completed.

4 Applied Cases

Taking the Mishrif Formation carbonate reservoir in W Oilfield of Iraq as an example, the method of recognition of reservoir types based on K-means clustering and Bayes discrimination is further expounded below.

4.1 Reservoir classification of cored well

4.1.1 Reservoir classification based on core data

According to classification method in Dunham (1962) based on different types of grain and mud supported during deposition of rocks, the rocks in the Iraqi study area are divided into grainstone, packstone, wackestone, and mudstone. Rudstone is also commonly developed in the W oil field, which has greater porosity and permeability than grainstone (Dunham, 1962). Based on the actual drilling core description, only grainstone, packstone, and mudstone provide reservoirs in the W Oilfield, and mudstone is a non-reservoir.

The pore development and physical property distribution of the Mishrif Formation in the study area are obviously controlled by the distribution of sedimentary facies and late-stage diagenetic transformation. There are six types of sedimentary sub facies in Mishrif Formation reservoir: limited platform, open platform, platform margin reef, platform front slope, and open shelf. High-quality reservoirs are developed in the reef sub facies at the margin of the platform. Rudstone and grainstone are grain-supported lacking lime mud, and mainly contain rudists (e.g., Theni, 1998), bivalves and echinoids. Packstone are grain-supported with the lime and mud filled into the space between grains, such as benthic forams, gastropods, and echinoids. Wackestone is mud-supported with more than 10% bioclastic grains, and contains bivalves, benthic and planktonic forams. The rock types of such a reservoir are rudstone and grainstone. Medium-quality reservoirs are developed in open platform sub facies and the rock type is packstone. Low quality reservoirs are developed in the limited platform and platform front slope, with wackestone and mudstone rock types. The distribution of porosity and permeability of a reservoir is controlled by the rock types (Fig. 4). The porosity and permeability of a rudstone reservoir are the highest, with the porosity greater than 22.3%, and the permeability greater than 99.8 mD. The second is a grainstone reservoir, with porosity ranging from 18.7% to 23.1% and permeability ranging from 21.2–428 mD. The least effective is the wackestone reservoir with porosity less than 8.9% and permeability less than 0.78 mD. However, the distribution range of the packstone reservoir overlaps with that of the grainstone reservoir and the wackestone reservoir, indicating that the environment in which the packstone developed was similar to the two former rock types.

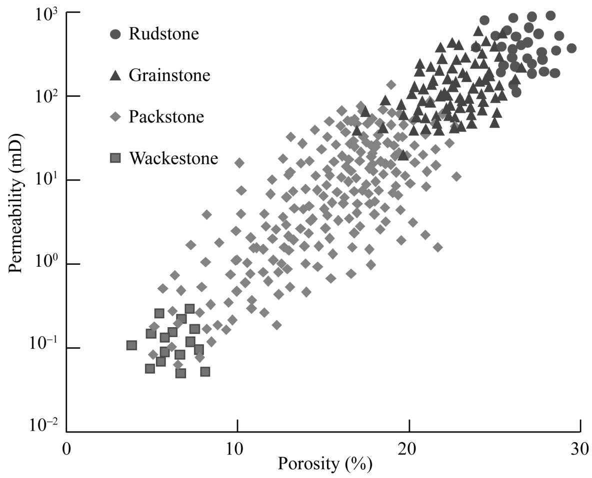

Fig. 4. Porosity and permeability characteristics of the Mishrif Formation reservoir (mD-Unit of permeability, millidarcy).

On the basis of the sedimentary environment, diagenesis had a strong impact on the late transformation of carbonate reservoirs in the Mishrif Formation in the study area, which is mainly reflected in two aspects: differential cementation and differential dissolution. Diagenesis is an important factor controlling the development of the main pore space (intergranular dissolved and moldic pores) of the formation.

Considering the impacting factor of diagenesis, according to the degree of transformation of carbonate reservoirs by strong dissolution, weak dissolution and strong cementation, rudstone reservoirs with intergranular dissolved and moldic pores under strong dissolution can be classified as high-quality reservoirs, whereas grainstone reservoirs can be classified as medium- to high-quality reservoirs. Packstone reservoir with intergranular and moldic pores under predominant medium dissolution and medium cementation is regarded as medium-quality reservoir. The packstone and wackestone reservoirs with moldic and intragranular pores developed under weak dissolution and strong cementation are classified as low-quality reservoirs. A wackestone reservoirs dominated by extra-strong cementation and with micro-pores developed is classified as a poor reservoir.

Based on the rock types under the control of sedimentation and the reservoir physical properties under the control of diagenesis, the reservoirs can be divided into five types by comprehensively considering the transformation of the reservoirs under the control of sedimentation and diagenesis (Table 1): Type I is a high-quality reservoir; Type II is a medium- to high-quality reservoir; Type III is a medium-quality reservoir; Type IV

Table 1 Reservoir classification and characteristics of the Mishrif Formation based on core data

Reservoir types	Porosity (%)	Permeability (mD)	Facies	Dissolution degree	Cementation degree	Rock types	Pore types
Type I	> 22.3	> 99.8	Platform margin reef	High	Low	Rudstone	Intergranular, moldic, dissolution
Type II	18.7–23.1	43.2–428	Platform margin reef	High	Low	Grainstone	Intergranular, moldic, dissolution
Type III	16.3–22.3	10.2–44.6	Restricted platform Open platform	Moderate	Moderate	Packstone	Intergranular, moldic, dissolution, micropore
Type IV	12.6–17.8	3.2–11.3	Platform front slope	Low	Moderate	Packstone	Intergranular, moldic, intragranular, micropore
Type V	7.6–13.9	0.08–4.1	Open shelf	Low	High	Packstone Wackstone	Moldic, micropore, intragranular

is a medium-porosity and low-permeability reservoir; and Type V reservoir has low porosity and ultra-low permeability, and such a reservoir is poor. Their characteristics are discussed below.

4.1.2 Relationship between reservoir types and logging facies

Compared with clastic reservoirs, carbonate reservoirs are characterized by multiple types of pores, great variation of logging response characteristics, strong reservoir heterogeneity (especially fracture and cavern reservoirs), poor correspondence between logging response and reservoir physical properties, and similar logging response characteristics of true and false reservoirs. These factors restrict the division of carbonate reservoirs based on logging data, which can easily lead to wrong or missing reservoirs. The classification of carbonate reservoirs based on logging data needs to make full use of all kinds of logging data to analyze the logging response characteristics of different reservoir types, so as to lay a foundation for subsequent reservoir recognition. Based on the determined reservoir classification scheme and the accurate recognition of reservoir types in 12 core Wells of Mishrif Formation in the study area, the logging response characteristics of each reservoir type in Mishrif Formation in the study area are summarized, and the conversion model between reservoir types and logging facies is established (Fig. 5).

The petrophysical characteristics of carbonate rocks are the physical basis of reservoir logging response evaluation. Different types of reservoirs have different logging response characteristics due to their different rock and pore types. The rock types of Type I and Type II reservoirs are mainly rudstone and grainstone, which are developed in high-energy deposition environment, such as plat form margin reef. Its main bioclastic composition is rudist, with long axis greater than 2mm and micritization locally. Rudstone contains rudist with its body size greater than 4 mm (Fig. 5). The composition of limestone is high, showing low natural gamma ray and high resistivity. On the other hand, Type I and Type II reservoirs suffer from strong dissolution, and develop a large number of intergranular, moldic and dissolution pores, with good porosity and high neutron value, high acoustic time difference value and low-density value. The rock type of Type III reservoir is packstone, which is developed in restricted platform and open platform, affected by the weakening of the sedimentary hydrodynamic force, the mud content increases with the characteristics of medium natural gamma and medium resistivity. In addition, affected by medium cementation, the primary intergranular pores of Type III are damaged to some extent, with medium and high neutron values, medium and low sonic time difference values, and medium density values. The rock types of Type IV and Type V reservoirs are packstone, wackestone, which is developed in platform slope and open shelf, with high mud content and the characteristics of high natural gamma and low resistivity. The types of bioclasts in the packstone are diverse, mainly including benthic foraminifers, bivalves, echinoderms, and sponge spicules, gastropods, pelletoid (Fig. 5). These two types of reservoirs are strongly affected cementation, primary pores are seriously destructed, leading to the characteristics of low neutron value, low acoustic time difference, and high-density value (Fig.5).

4.1.3 Selection of sensitive logging curves

For all kinds of logging data, the histogram method is firstly used to standardize them. Secondly, a set of well logging curves are selected to make cross plot of each reservoir, and the results of distinguishing reservoir types with different sets of well logging data are compared and analyzed. It can be seen from the cross plot that resistivity (ILD), acoustic time difference (DT), natural gamma ray (GR), neutron porosity (NPHI) and density (DEN) can distinguish the five types of reservoirs well (Fig. 6). Therefore, five logging parameters, such as resistivity, density, acoustic time difference, neutron porosity and natural gamma ray, are selected as sensitive logging parameters to study the classification criteria of reservoir types.

4.1.4 Establishment of classification criteria

In order to make the clustering results more accurate, the average value of sensitive logging data corresponding to each reservoir is calculated as the given initial clustering center, and the final clustering results of different types of reservoirs are obtained through the K-means clustering method (Table 2).

4.1.5 Establishment of discriminant formula

The Bayesian discriminating method is applied to establish the discriminant models for different types of reservoirs according to the selected sensitive logging parameters:

$$P_1 = 3788.967NPHI + 3002.375RHOB - 5.834GR + 2.023ILD + 0.683DT - 3101.436 \quad (3)$$

$$P_2 = 3653.523NPHI + 2176.436RHOB - 3.498GR + 1.889ILD + 0.765DT - 3034.587 \quad (4)$$

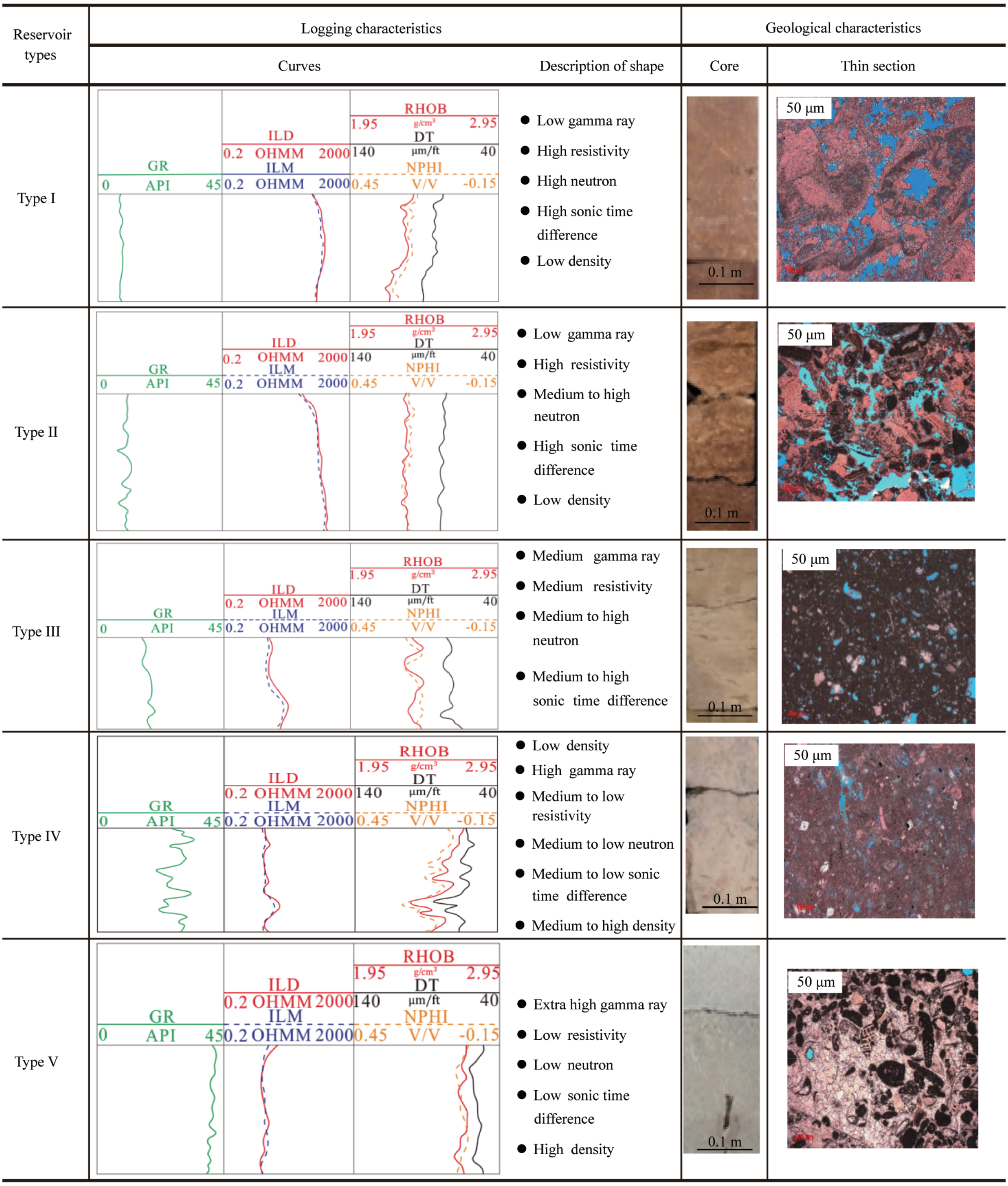

Fig. 5. Relationship between reservoir types and logging characteristics of the Mishrif Formation.

$$P_3 = 3590.478\text{NPHI} + 2172.423\text{RHOB} - 3.325\text{GR} + 1.234\text{ILD} + 0.698\text{DT} - 2989.689 \quad (5)$$

$$P_4 = 3497.346\text{NPHI} + 2019.532\text{RHOB} - 3.921\text{GR} + 1.653\text{ILD} + 0.712\text{DT} - 2901.145 \quad (6)$$

$$P_5 = 3392.793\text{NPHI} + 2102.876\text{RHOB} - 4.128\text{GR} + 1.286\text{ILD} + 0.467\text{DT} - 2890.563 \quad (7)$$

4.2 Test of discriminating model

The established reservoir type discriminating model was used to discriminate the 205 samples participating in the training. Due to the limited number of Type V cores, 25 samples were selected, and 45 samples were selected for the remaining four types of reservoirs. The accuracy of discrimination is as follows (Table 3). Judging from the overall discriminating results, the Bayesian discriminating

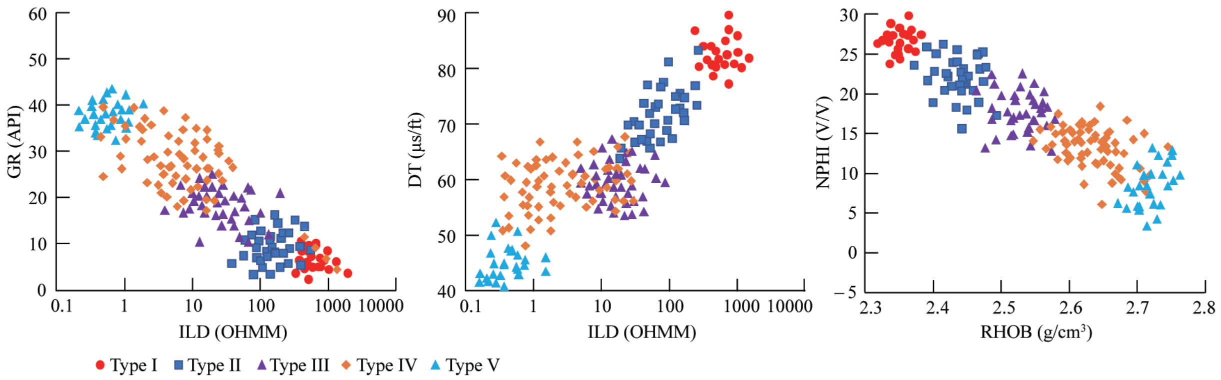

Fig. 6. Cross plot of logging characteristics of the different reservoir types in the Mishrif Formation.

Table 2 Initial clustering center and final clustering results of K-means clustering

Reservoir type	Clustering section	GR (API)	ILD (OHMM)	DT (μs/ft)	RHOB (g/cm³)	NPHI (V/V)
Type I	Initial clustering center	7.8	992.3	81.4	2.3	25.3
	Final clustering results	< 9.2	> 954.8	> 78.6	< 2.38	> 24.7
Type II	Initial clustering center	9.6	636.9	76.2	2.42	24.1
	Final clustering results	3.3–13.2	103.2–1003.5	67.8–80.2	2.37–2.54	16.2–25.4
Type III	Initial clustering center	14.4	68.9	63.4	2.51	18.8
	Final clustering results	10.4–25.4	10.8–104.8	57.8–74.3	2.44–2.58	12.4–23.8
Type IV	Initial clustering center	25.8	8.4	59.3	2.62	10.4
	Final clustering results	23.8–36.2	1.8–11.8	53.8–64.7	2.57–2.73	5.6–15.4
Type V	Initial clustering center	38.6	0.94	47.1	2.74	8.9
	Final clustering results	> 35.8	0.12–1.8	< 53.2	2.69–2.85	4.2–13.4

Table 3 Bayesian discriminating results

		Predicted	Core					
			Type I	Type II	Type III	Type IV	Type V	Total
Discriminant results of original samples	Number	Type I	38	7	0	0	0	45
		Type II	3	39	3	0	0	45
		Type III	0	5	34	6	0	45
		Type IV	0	0	5	32	8	45
		Type V	0	0	0	3	22	25
	Percentage (%)	Type I	84.4	15.6	0.0	0.0	0.0	100.0
		Type II	6.7	86.7	6.6	0.0	0.0	100.0
		Type III	0.0	11.1	75.6	13.3	0.0	100.0
		Type IV	0.0	0.0	11.1	71.1	17.8	100.0
		Type V	0.0	0.0	0.0	12.0	88.0	100.0
Cross confirmation of discrimination results	Number	Type I	39	6	0	0	0	45
		Type II	2	39	4	0	0	45
		Type III	0	3	35	7	0	45
		Type IV	0	0	6	34	5	45
		Type V	0	0	0	4	21	25
	Percentage (%)	Type I	86.7	13.3	0.0	0.0	0.0	100.0
		Type II	4.4	86.7	8.9	0.0	0.0	100.0
		Type III	0.0	6.7	77.8	15.5	0.0	100.0
		Type IV	0.0	0.0	13.3	75.6	11.1	100.0
		Type V	0.0	0.0	0.0	16.0	84.0	100.0

accuracy rate and the cross-confirmation accuracy rate are both over 80%, and the two discriminating accuracy rates are very close, indicating that the established reservoir type discriminating model is very stable and meets the requirements of oilfield production and development. It can be seen from Table 3 that the discriminating accuracy and cross-confirmation accuracy rate of Type I, Type II, and Type V reservoirs all exceed 80%, while these two-accuracy rate of Type III reservoirs are 75.6% and 77.8%, respectively. The discriminating accuracy rate of Type IV reservoirs is 71.1%, and the accuracy rate of cross confirmation is 75.6%.

The reservoir type discriminating method proposed in this paper was applied to reservoir prediction of W-271 well, another cored well in the study area, which did not participate in training (Fig. 7). Comparing the prediction results with the core analysis, it can be seen that the classification of the Type IV reservoir at the depth of 2256 m is Type V, and the classification of Type I reservoir at the depth of 2333 m is Type II. Although some depth points are misjudged, the overall effect is good.

4.3 Recognition of reservoir types in non-cored wells

After standardized processing of sensitive logging data selected by cross-plot method, the logging classification standard obtained by K-means clustering method was used

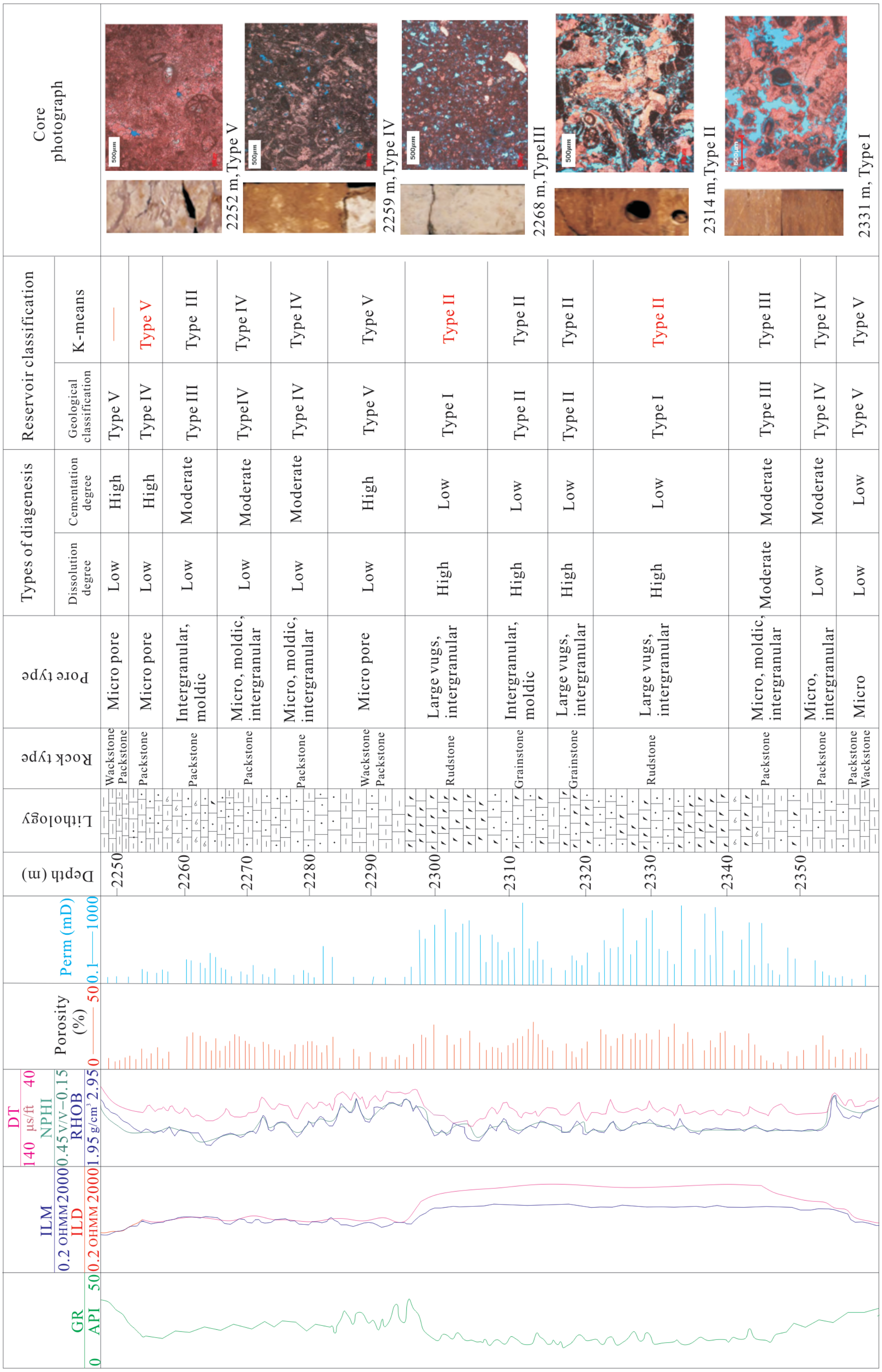

Fig. 7. Discrimination results of cored well W-271.

as constraint conditions to classify non-cored wells. The discriminant formula obtained by Bayesian discriminating method was used to verify the discriminating accuracy of various types of reservoirs in non-cored wells.

Taking wells W-131, W-58, and W-190, W-181 as examples, the accuracy of the recognition of various types of reservoirs in the Mishrif Formation of the three wells was calculated respectively. Among them, the results of recognition from well W-131 is shown in Fig. 8. The lateral distribution of the reservoir type is obtained by combining the prediction results of non-cored single well reservoir type (Fig. 9).

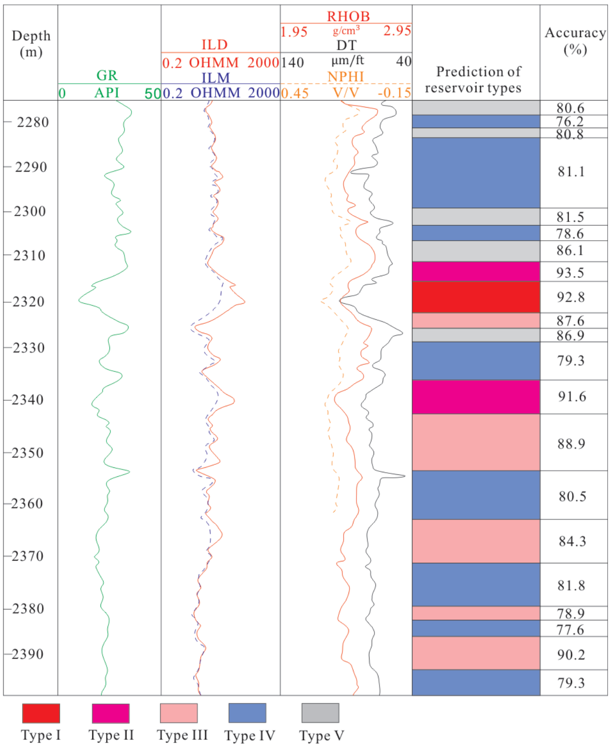

Fig. 8. Reservoir types discrimination results of non-cored well W-131.

It can be seen from the table that the average accuracy of recognition of five types of reservoirs is 76.5%–93.3%. The average accuracy of recognition of each single type of reservoir is 70% to 94%. Compared with the other four types of reservoirs, the recognition accuracy rate of Type IV reservoirs is the lowest, mainly because Type IV reservoirs are formed in the cross-transition zone from the margin of the high-energy sedimentary facies belt to the low-energy sedimentary facies belt, which leads to more complex rock fabric. They have similar parts to the rock fabric formed at the margin of the high-energy facies belt and the low-energy facies belt, so their logging response characteristics also show similar ones, which decreases the recognition rate of Type IV reservoirs and increases the difficulty of recognition. Although the recognition accuracy of Type IV reservoirs is generally low, it can be seen from the Fig. 8 that as the thickness increases, the accuracy is increasing. The greater the thickness of the reservoir, the impact caused by the depositional environment in the transition zone will be eliminated, thereby further improving the accuracy of reservoir recognition.

5 Discussion

The established discriminant model of reservoir types has a very high accuracy rate and overall consistency in the W oil field. However, the two main factors, sedimentary environment and diagenesis, may affect the discrimination effect. When affected by these factors, he characteristics of logging curve of different reservoirs could generate some similarity and not discriminant the reservoir types, lowering the accuracy of the model. The discrimination accuracy rate of the model would be further improved by increasing the number of core samples in the

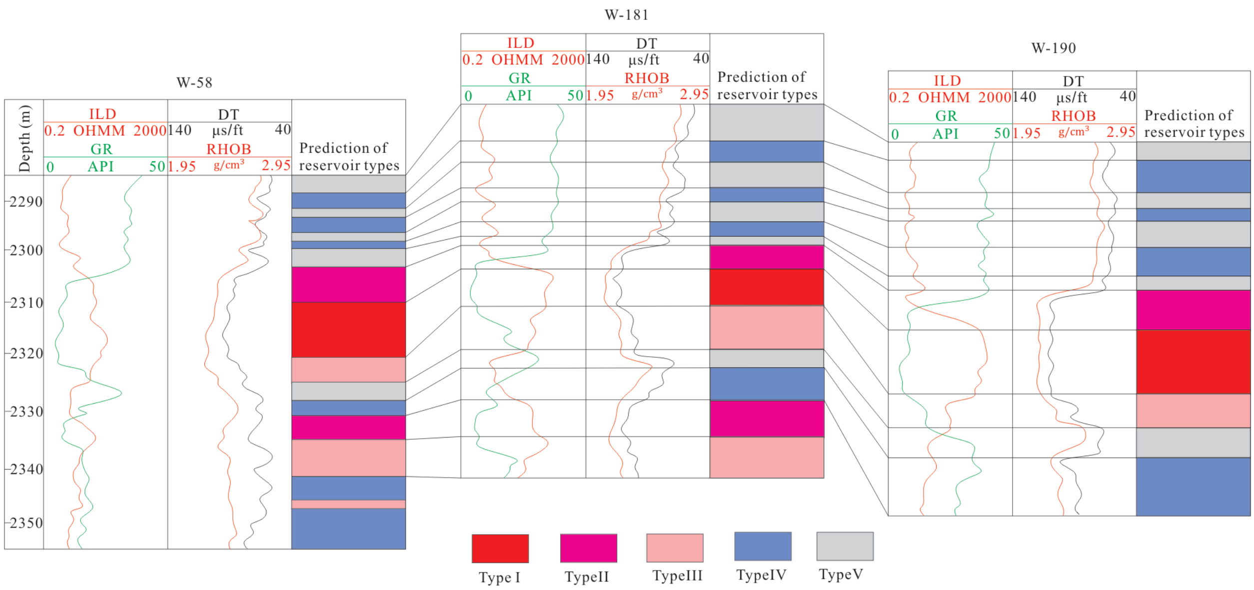

Fig. 9. Prediction results of non-cored wells with proposed methods.

future. In terms of the calculation method, the covariance inverse matrix data A_{ij}^{-1} of the core samples may affect the independence of the discriminant model in the calculating process, and the smaller value of the normalized logging data will result in the large value of the covariance inverse matrix data A_{ij}^{-1}, which will lead to some systematic errors.

According to Fig. 7, it can be seen that the conclusions of external verification are consistent with those of a self-check. The discrimination accuracy of Types I, II, and V reservoirs is the highest, which is above 80%. The discrimination accuracy of Types III and IV reservoirs is lower than that of Type V. The main reason for this error is mainly related to the sedimentary environment and diagenesis. The rock type of Type IV reservoir is packstone, which is supported by particles, and affected by the weaker sedimentary hydrodynamic force, with mud content higher than that of grainstone, mainly found in open platform and platform front margin, where the hydrodynamic force is gradually weakening. In the transition region from platform margin reef to limited platform gradually, the packstone formed is located at the edge of a high-energy environment, and the pore type and structure are very similar to grainstone, so that the reservoir in this area has been misjudged as a Type III reservoir. With hydrodynamic waning, the mud content gradually increased, in addition it was affected by the compaction, cementation, which seriously damaged the primary pore. So, logging response characteristics and pore structure of Type IV and Type V reservoirs have similar parts to some extent, which increases the possibility that a Type IV reservoir can be recognized as Type V reservoir, ultimately affecting the discriminant accuracy. The pore structure of Types I, II and V reservoirs is dissimilar to other two types of reservoirs, and the logging response characteristics are unique and specific. Therefore, the discrimination accuracy of these three types of reservoirs is relatively high.

On the basis of inheriting the sedimentary environment, the diagenesis impacts strongly in the later reformation of the carbonate reservoirs of the Mishrif Formation in the studied area, which is mainly reflected in two aspects: differential cementation and differential dissolution. The better the reservoir quality, the stronger the dissolution and the weaker the cementation. Dissolution is excellently conducive to the transformation of pore structure. In terms of diagenesis, according to the degree of transformation of a carbonate reservoir by strong dissolution, weak dissolution, weak cementation and strong cementation, the rudstone and grainstone with intergranular dissolution pores and mold pores developed under strong dissolution can be considered as a high-quality reservoir, i.e., Type I. The grainstone with intergranular pore and mold pore development dominated by weak dissolution is considered as a good reservoir, i.e., Type II. The packstone with partial primary pores cemented under weak cementation is considered as medium reservoir, i.e., Type III. Under strong cementation, all the primary intergranular pores are cemented and sealed marl, which is considered as weak reservoir, i.e., Types IV and V.

It can be seen from the prediction results that although the Bayesian discrimination method still has certain shortcomings, compared with the results of core analysis, the overall discrimination results show that this method is still well adapted to the quantitative recognition of reservoir types in the study area.

As the pore structure of grainstone and rudstone is much better than in wackestone and mudstone, with no infilled lime mud, lower shale content, and greater porosity, Type I and Type II are characterized by very high deep resistivity. While the pore structure mud-supported of Types IV and V leads to higher mud content and fewer effective porosity, these two reservoirs have very low resistivity (Table 2). Correspondingly, the natural gamma ray (GR) of Type I and Type II is lower than in Type IV and Type V, in addition acoustic time difference (DT) of Type I and Type II is higher than Type IV and Type V (Figs. 7, 8). The higher porosity and permeability, the better the oil quality of the reservoir. Therefore, in the studied area, based on typical logging curves, such as resistivity, gamma ray, and acoustic time difference, we could distinguish the type of reservoir and rock. The Type I reservoir is most well developed in the upper mB2, distributed in a wide range of sheets; in addition, the thickness of the upper part of the structure is better than that of the lower part, which conforms to the sedimentary background of the platform marginal reef in this section (Fig. 9). Reservoir thickness and sedimentary environment are greatly related to the accuracy of reservoir prediction. The mB1 and mB2 Upper is a high-energy facies zone with less mud content and large reservoir thickness. The pore structure formed by mB1 and mB2 has less crossover with other types of reservoirs, which improves the recognition accuracy. Therefore, these characteristics are conducive to recognizing the Type I and Type II reservoirs.

However, because of the impact of actual geological characteristics, this method needs more core data, which will impact the accuracy of prediction. As it has some similarity in sedimentary characteristics in Middle East oilfield, the new method could be referred to other oilfields in the Middle East. The structure of the new method could also be applied to sandstone reservoirs.

6 Conclusions

This study is based on the core, logging, and logging data of 10 cored wells in the Mishirif Formation, W oilfield, Iraq, and the physical properties of the reservoir under the comprehensive influence of sedimentation and diagenesis. We put forward a method for automatic classification and verification of reservoir types to improve the efficiency of reservoir classification and accuracy of recognition.

(1) The reservoir type recognition method based on K-means clustering and Bayesian discrimination mainly includes reservoir classification of cored well based on core data, establishment of logging classification criteria for the reservoir, establishment of discriminating formulas of reservoir types, and recognition of reservoir types in non-cored wells. This method can be applied to the establishment of reservoir classification criteria of a single

well and the establishment of discriminating formulas in other study areas.

(2) Three problems need to be paid attention to when using the method proposed to recognize the reservoir types in non-cored wells. Firstly, when using core data to classify the reservoir, the depositional characteristics of the reservoir must be fully understood. Next, in order to obtain the clustering result with high accuracy, it is necessary to adjust the logging curve combination and set the appropriate initial clustering center when establishing the logging classification criteria. The accuracy of the discriminating formula established by the Bayesian discriminant method depends largely on whether the parameters selected for modeling are appropriate or not. The higher the matching degree between the selected parameters and the deposition and diagenesis of the reservoir, the higher the accuracy of the discrimination will be.

(3) According to the analysis of test results, the various geological factors, especially sedimentation and diagenesis, are important reasons for the decrease in recognition accuracy. In Iraq, the transition zone of the sedimentary facies belt with the development of IV reservoirs leads to the existence of intersecting areas of lithology, in addition of cementation, making its petrophysical characteristics very similar to those of Type V reservoirs, resulting in a decrease in the difference of logging curves, and finally the recognition results of Type IV reservoirs are poor.

Acknowledgments

The paper is funded by the National Key Research and Development Program (Grant No. 2018YFC0807804-2). Dr. Susan Turner (Brisbane) is thanked for her language assistance and research/detective work.

Manuscript received Jul. 13, 2021
accepted Mar. 1, 2022
associate EIC: XIAO Yitian
edited by FEI Hongcai

References

Abdulaziz, A.M., Mahdi, H.A., and Sayyouh, M.H., 2019. Prediction of reservoir quality using well logs and seismic attributes analysis with an artificial neural network: A case study from Farrud Reservoir, Al-Ghani Field, Libya. Journal of Applied Geophysics, 161: 239–254.

Abuamarah, B.A., and Nabawy, B.S., 2021. A proposed classification for the reservoir quality assessment of hydrocarbon-bearing sandstone and carbonate reservoirs: A correlative study based on different assessment petrophysical procedures. Journal of Natural Gas Science and Engineering, 88: 1–16.

Al-Sakini, 2010. Summary of petroleum geology of Iraq and the Middle East. Kirkuk, Iraq: Report of Northern Petroleum Company, 1–22.

Aqrawi, A.A.M., Thehni, G.A., and Sherwani G.H., 1998. Mid-Cretaceous rudist-bearing carbonates of the Mishrif Formation: An important reservoir sequence in the Mesopotamian basin, Iraq. Journal of Petroleum Geology, 21: 57–82.

Awoleke, O.O., and Lane, R.H., 2011. Lane analysis of data from the Barnett Shale using conventional statistical and virtual intelligence techniques. SPE, 14: 544–556.

Chen, Y.L., Zeng Y., and Duan, Y.M., 2018. Pore structure characteristics and reservoir classification of dolomite reservoirs in fourth member of Leikoupo Formation, Longmen Mountain front, western Sichuan Basin. Petroleum Geology & Experiment, 40: 621–631.

Dunham, K.C., 1962. Classification of carbonate rocks according to depositional textures. AAPG, 1: 108–121.

El Sharawy, M.S., and Nabawy, B.S., 2016. Geological and petrophysical characterization of the Lower Senonian Matulla Formation in southern and central Gulf of Suez, Egypt. Arabian Journal of Geoscience, 41: 281–300.

El Sharawy, M.S., and Nabawy, B.S., 2019. Integration of electrofacies and hydraulic flow units to delineate reservoir quality in uncored reservoirs: A case study, Nubia Sandstone Reservoir, Gulf of Suez, Egypt. Natural Resources Research, 28: 1578–1608.

Feng, F.K., Wang, T.B., and Zhang S.Y., 1995. Natural Gas Geology of China. Beijing: Geology Press, 1–218.

Gao, J.X., Tian, C.B., and Zhang, W.M., 2013. Middle Cretaceous of Persian Gulf basin characteristics and development model of carbonate reservoir in Mishrif Formation. Geological Science, 48: 304–316.

Jodry, R.L., 1972. Pore Geometry of Carbonate Rocks (Basic Geologic Concepts). Oil and Gas Production from Carbonate Rocks, New York, Elsevier, 1–408.

Kakeman, U., Jafarian, A., Husinec, A., Hossein Adabi, M., and Mahmoudi, A., 2021. Facies, sequence framework, and reservoir quality along a Triassic carbonate ramp: Kangan Formation, South Pars Field, Persian Gulf Superbasin. Journal of Petroleum Science and Engineering, 30(198): 1–15.

Khan, N., and Rehman, K., 2021. Application of fuzzy logic and neural networks for porosity analysis using well log data: An example from the Chanda Oil Field, Northwest Pakistan. Earth Science Informatics, 14: 2183–2199.

Larsen, A.L., and Ulvmoen, M., 2006. Bayesian lithology/fluid prediction and simulation on the basis of a Markov-chain prior model. Geophysics, 71: 69–78.

Liao, T.P., and Zhang, F.R., 2004. Complex lithologic identification based on fuzzy clustering algorithm. Journal of Daqing Petroleum Institute, 28: 58–60.

Liu, X.Y., and Chen, X.H., 2016. Reservoir physical property prediction based on kernel-Bayes discriminant method. Acta Petrolei Sinica, 37: 878–886.

Luo, L., and Ren, X.G., 1999. Logging identification of carbonate reservoir types. Well Logging Technology, 23: 355–360.

Mahdi, T.A., and Aqrawi, A.A.M., 2014. Sequence stratigraphic analysis of the mid‐cretaceous Mishrif Formation, southern Mesopotamian basin, Iraq. Journal of Petroleum Geology, 37: 287–312.

Nabawy, B.S., Basal, A.M.K., Sarhan, M.A., and Safa, M.G., 2018. Reservoir zonation, rock typing and compartmentalization of the Tortonian Serravallian sequence, Temsah Gas Field, offshore Nile Delta, Egypt. Marine and Petroleum Geology, 92: 609–631.

Rimstad, K., and Avseth, P., 2012. Hierarchical Bayesian lithology/fluid prediction: A North Sea case study. Geophysics, 77: 70–85.

Sames, M., and Saussus, D., 2010. Uncertainties in the quantitative interpretation of lithology probability volumes. The Leading Edge, 29: 576–583.

Sharland, P.R., Archer, R., and Case, D.M., 2001. Arabian Plate Sequence Stratigraphy. Geo Arabia Special Publication, Bahrain, 1–387.

Sima, L.Q., Xiao, H., and Yuan, L., 2012. Study on reservoir flow unit based on BP neural network technology. Well Logging Technology, 36: 421–425.

Smith, L.B., Eberli, G.P., and Masaferro, J.L., 2003. Discrimination of effective from ineffective porosity in heterogeneous Cretaceous carbonates, Al Ghubar Field, Oman. AAPG Bulletin, 87: 1509–1529.

Sun, Z.J., Fu, M.Y., and Leng, Q.C., 2013. Influence of early diagenesis on reservoir properties of Rumaila Formation in an oilfield. Complex Hydrocarbon Reservoirs, 6: 24–28.

Szabó, N.P., and Nehéz, K., 2019. Cluster analysis of core measurements using heterogeneous data sources: An

application to complex Miocene reservoirs. Journal of Petroleum Science and Engineering, 178: 575–585.
Thehni, G.A., 1998. Mid-Cretaceous rudist-bearing carbonates of the Mishrif Formation: An important reservoir sequence in the Mesopotamian basin Iraq. Journal of Petroleum Geology, 21: 57–68.
Umer, A., Zhu, P.M., and Qamar, Y., 2019. Classification of reservoir facies using well log and 3D seismic attributes for prospect evaluation and field development: A case study of Sawan gas field, Pakistan. Journal of Petroleum Science and Engineering, 175: 338–351.
Wang, F.F., Li, J.Y., and Chen, X.J., 2014. Identification of Bayesian lithofacies based on Markov chain priori model. Petroleum Geophysical Exploration, 49: 183–194.
Wang, J., and Zhang, C.M., 2019. Pore structures and classification of carbonate reservoirs in the Middle East. Well Logging Technology, 43: 631–637.
Wang, X.D., and Yang, S.C., 2018. Lithology identification using an optimized KNN clustering method based on entropy-weighed cosine distance in Mesozoic strata of Gaoqing field, Jiyang depression. Journal of Petroleum Science and Engineering, 166: 157–174.
Xun, Z.F., and Yu, J.F., 2008. The application of cluster and discriminant analyses in logging lithology recognition. Journal of Shandong University of Science and Technology, 27: 10–13.
Yi, H., and Chong, Z., 2018. Pore structure classification and logging evaluation method for carbonate reservoirs: A case study from an oilfield in the Middle East. Energy Sources, 41: 1701–1715.
Zhang, H.Y., and Du, W., 2021. Split K-means clustering algorithm based on density weighting. Computer Simulation, 38: 253–264.
Zhao, J., and Liu, Y.Y., 2018. Identification and classification prediction of fractured-vuggy reservoir type in carbonate reservoirs. Carsologica Sinica, 37: 584–593.

About the first and corresponding author

FANG Xinxin, male; born in 1989 in Hubei Province; Ph.D. in geology; graduated from Changjiang University; junior engineer; His research interest is geological evaluation, carbonate reservoir evaluation, hydraulic fracturing. E-mail: fx15827573109@163.com.