Study on the Interpretation Method of Low Resistivity Contrast Oil Reservoir Based on Support Vector Machine—Taking the Chang 8 Tight Sandstone Reservoir of Yanchang Formation in Huanxian Area, Ordos Basin as an Example

Ze Bai (✉ baize@aust.edu.cn)

Anhui University of Science and Technology

Maojin Tan

China University of Geosciences

Yujiang Shi

China National Petroleum Corporation (China), Changqing oilfield

Xingning Guan

Anhui University of Science and Technology

Research Article

Keywords: Low resistivity contrast oil layer, Support vector machine, Log interpretation, Tight sandstone reservoir

Posted Date: October 21st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-985788/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Study on the interpretation method of low resistivity contrast oil reservoir based on support vector machine—Taking the Chang 8 tight sandstone reservoir of Yanchang Formation in Huanxian area, Ordos Basin as an example

Ze Bai^{1*}; Maojin Tan²; Yujiang Shi³; Xingning Guan¹

1. *School of Earth and Environment, Anhui University of Science & Technology, Huainan 232001, China*
2. *School of Geophysics and Information Technology of China University of Geosciences, Beijing 100083, China.*
3. *A Research Institute of Exploration and Development, PetroChina Changqing Oilfield Company, Xi'an 710018, China.*

Correspondence and requests for materials should be addressed to Ze Bai (email: baize@aust.edu.cn)

Abstract: Low resistivity contrast oil reservoirs are subtle reservoirs that have no obvious difference in physical and electrical properties from water layers. It is difficult to identify based on the characteristics of the geophysical well logging response. Especially in tight sandstone reservoirs with low porosity and low permeability, the log interpretation effect of low resistivity contrast oil layers is worse. In recent years, data mining technology has been increasingly applied in oil exploration and development, especially for some complex reservoirs with unclear logging response characteristics, and how to use data mining technology to effectively solve some complex problems is of great significance in oilfields. Therefore, support vector machine (SVM) technology was applied to interpret the low resistivity contrast oil layer in this paper. First, the input data sequences of logging curves were selected by analyzing the relationship between reservoir fluid types and logging data. Then, the SVM classification model for fluid identification and SVR regression model for reservoir parameter prediction were constructed. Finally, the two models were applied to the logging interpretation of the Chang 8 tight sandstone reservoir of the Yanchang Formation in the Huanxian area, Ordos Basin. The application results show that the fluid recognition accuracy of the

SVM classification model is higher than that of the logging cross plot method and BP neural network method. The calculation accuracy of permeability and water saturation predicted by the SVR regression model is higher than that based on the experimental fitting model, which indicates that it is feasible to carry out logging interpretation and evaluation of the low resistivity contrast oil layer by the SVM method. The research results not only provide an important reference and basis for the review of old wells but also provide technical support for the exploration and development of new strata.

Key words: Low resistivity contrast oil layer; Support vector machine; Log interpretation; Tight sandstone reservoir.

1、 Introduction

With the increasing volatility of international oil prices and the continuous reduction of oil reservoir scale, it is particularly important and urgent to carry out logging interpretation and evaluation research for low resistivity contrast oil reservoirs (pratama et al., 2017; Bai et al., 2021). The low resistivity contrast oil reservoir has the characteristics of little difference in porosity and resistivity logging response from the water layer, and the oil saturation of low resistivity contrast oil layer is relatively low, so it is difficult to identify and evaluate such oil layers by using logging response (mode et al., 2015; hakimov et al., 2019). At present, the low porosity and low permeability reservoir represented by tight sandstone has become the main battlefield to ensure the supply of oil and gas resources. However, the complex pore structure and strong heterogeneity reduce the sensitivity of the logging response to pore fluid, resulting in more low resistivity contrast oil layers developed in tight sandstone reservoirs, and it is more difficult to interpret and identify this kind of reservoir by using conventional logging interpretation methods (Bai et al., 2019; Shi et al., 2021).

In recent years, data mining technology has been increasingly applied in oil exploration and development, especially for unconventional reservoirs with unclear logging response characteristics, and how to use data mining technology to effectively solve some complex problems existing in the actual production of oil fields is of great significance (Kadhim et al., 2017; shahriari et al., 2020). Some classical optimization

algorithms, such as the neural network method, support vector machine and fuzzy clustering method, provide a new technology for the identification of low resistivity contrast oil reservoirs (Wang et al., 2015; Liu et al., 2017). Guo et al. (2015) predicted the water saturation at the lower limit of three water models by using the generalized neural network (GRNN) and particle swarm optimization support vector machine (PSO-SVM), which is in good agreement with the core analysis results in the Sulige tight sandstone reservoir. Chen and Peng (2020) used a BP neural network to train and learn the mathematical characteristics of logging curves of low resistivity oil reservoirs, which improved the accuracy of fluid identification and reservoir parameter prediction. Singh et al. (2016) used the stepwise linear regression, multilayer feed forward neural (MLFN) network method to predict the 2D distribution of P-wave velocity, resistivity, porosity, and gas hydrate saturation. Miah et al. (2020) used the multilayer perception artificial neural network (MLP-ANN) and kernel function-based least-squares support vector machine (LS-SVM) techniques to develop predictive models for water saturation, and the prediction performance was better than that of other models. At present, there are many machine learning algorithms based on theoretical mathematics, and each has its own advantages. However, the key to applying this kind of method to log interpretation of actual formation is to select appropriate training data as input (baneshi et al., 2013; Roslin and esterle, 2016).

The Ordos Basin is the second largest sedimentary basin in China, and now more than half of China's energy transfer comes from the Ordos Basin (Yang et al., 2013; Fu et al., 2021). The Huanxian area is located in the southwestern Ordos Basin, and the regional geological structure crosses the Tianhuan Depression and Yishan Slope from west to the east (Figure 1a). The tight sandstone reservoir of the Chang 8 member of the Yanchang Formation developed in the Huanxian area has a large sedimentary thickness, and the oil source mainly comes from the overlying Chang 7 high-quality source rock, which has great exploration and development potential (Zhao et al., 2013; Chai et al., 2020). With the deepening of oil and gas exploration and development in this area, the problem of logging interpretation and evaluation of low resistivity contrast reservoirs has become

increasingly prominent. Figure 1 (b) shows the relationship between the resistivity and density logging response of different pore fluids established according to the oil test data in the study area. The reservoir density and resistivity of the low resistivity contrast oil layer and water layers are low, and it is difficult to identify and evaluate low resistivity contrast oil layers by using conventional logging data, which seriously restricts the exploration progress and development of oil resources in this area. It is important to develop more effective methods to provide new logging technical support for the exploration and development of low resistivity contrast oil layers in unconventional reservoirs.

(a)

Figure 1. (a) Geographical location of the research area and (b) cross plot of reservoir resistivity and density from different fluid types

In this study, the support vector machine (SVM) learning method based on VC dimension theory in statistical learning and the structural risk minimization principle (SRM) were used to establish the interpretation model. By analyzing the relationship between logging response and pore fluid, the training data were optimized, and the SVM classification model for fluid identification and support vector machine regression (SVR) model for reservoir parameter prediction were established. The application results show that the model established by the SVM learning method is more effective than the conventional reservoir interpretation method in fluid identification and parameter evaluation, which proves that the model is effective and feasible in the identification and parameter prediction of low resistivity contrast oil reservoirs.

2、Method and theory

Different from the neural network method to solve the number of hidden nodes of neurons, the basic idea of a support vector machine for reservoir parameter prediction is to map the input space to a high-dimensional space by introducing a kernel function and then solve a linearly separable hyperplane or function in this high-dimensional space, which can separate all data types in the original space. The greater the separation

distance is, the better the classification effect. Finally, the nonlinear discrimination ability of the original spatial data is realized (Vapink, 1998).

Taking $T = \{(\mathbf{x}_i, y_i) | i = 1, 2, \dots, n\}$ and $\mathbf{x}_i \in R^P$ as the input data, where $\mathbf{x}_i$ is the logging data related to the predicted parameters, and y_i is the core analysis data, that is, the target value.

Suppose that in high-dimensional space, the hyperplane or line function that can separate the two types of samples satisfies:

$$g(\mathbf{x}_i) = \langle \mathbf{w}_{ij} \mathbf{g} \mathbf{x}_i \rangle + b_{ij} \quad (1)$$

where $\mathbf{w}_{ij}$ is the weight vector representing high-dimensional unknown coefficients and b_{ij} is a constant term. To use function (1) to distinguish all input data samples without error, function $y_k(\langle \mathbf{w} \mathbf{g} \mathbf{x} \rangle + b) - 1 \geq 0$ should be satisfied. When the classification interval is maximum, function $\phi(\mathbf{w}) = \frac{1}{2} \mathbf{w}^T \mathbf{w}$ should be minimum. In this way, the problem of solving the optimal hyperplane in high-dimensional space is transformed into the minimum value problem of the following convex programming function:

$$\phi(\mathbf{w}, \xi) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{k=1}^n \xi_k \quad (2)$$

Which satisfies the following constraint condition:

$$y_k(\langle \mathbf{w} \mathbf{g} \mathbf{x} \rangle + b) \geq 1 - \xi_k, \quad k = 1, \dots, n \quad (3)$$

where ξ_k is a nonnegative relaxation variable introduced when the sample data are linearly inseparable; C is a penalty parameter, and the greater its value is, the heavier the penalty for misclassification. The first term in the objective function (2) is to increase the classification interval, which effectively controls the generalization ability of the model. The second term is the training error to reduce the experience risk.

To map the training data set to the high-dimensional space, a kernel function needs to be introduced; that is, the convex programming problem of equation (2) is transformed into a quadratic programming problem. The expected weight vector can be

written as $\mathbf{w} = \sum_{i=1}^n (\alpha_i^* - \alpha_i) \mathbf{x}_i$, and finally, the analytical expression of the support

vector machine regression function is as follows:

$$f(\mathbf{x}) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(\mathbf{x}, \mathbf{x}_i) + b \quad (4)$$

where α_i and α_i^* are the nonnegative Lagrange multipliers and $K(\mathbf{x}, \mathbf{x}_i)$ is a kernel function satisfying the Mercer condition. The commonly used kernel functions mainly include the polynomial kernel function, Gaussian kernel function, radial basis function kernel function and sigmoid kernel function. Fig. 2 shows the flowchart of constructing the classification model and regression model by using the SVM method.

Figure 2. Flowchart of constructing the classification model and regression model by using the SVM method.

2.1 The construction of SVM classification model

Fluid identification using SVM is a multiclassification problem, but the SVM method initially solves the two classification problems. Therefore, it is necessary to extend the SVM and construct a reasonable multiclassification coding scheme. At present, there are four main methods to construct SVM multiclassifiers: "one against one", "one against rest", "SVM decision tree" and "one-time solution method". When solving practical multiclassification problems, the "one-to-one" method has a better effect than other methods (Hsu et al., 2002; Peng and Zhang, 2007). Therefore, this method is selected to construct an SVM multiclassifier in this paper, and the basic idea is that if there are class k data, class I data and class J data are selected to construct a

classifier, where $I < J$, so $k(k-1)/2$ classifiers need to be trained. For class I and class J data, a two classification problem needs to be solved, and the voting method is used to solve this problem; that is, if the function judges that it belongs to class I, the number of votes of class I is increased by 1. Otherwise, the number of votes of class J is increased by 1, and the final output result is the class with the largest number of votes.

To build the SVM classification model for fluid identification, we must first determine the input logging data or parameters sensitive to the pore fluid. Considering that the study area is mainly conventional logging curves, nuclear magnetic resonance logs and array acoustic logs are not widely used in the whole area. Therefore, according to the characteristics of logging curve, the fluid identification factors sensitive to fluid type are selected as the input data, including $(K/\phi)^{1/2}$, D_R , QT , R_t , ΔSP , R_{wa} and R_{wa_SP} , where $(K/\phi)^{1/2}$ is the comprehensive physical property index, which K represents the permeability, and ϕ is the porosity of reservoir. QT is the total hydrocarbon logging value, the greater the value, the greater the probability of possible oil and gas. R_t is the resistivity logging value. The specific calculation methods of other parameters are as follows:

$$\Delta SP = \frac{SP_{Shale} - SP}{SP_{shale} - SP_{sand}} \quad (5)$$

where ΔSP is the relative amplitude of the spontaneous potential. When the salinity difference of formation water is small, the higher the oil saturation of the reservoir is, the smaller the ΔSP value; SP is the spontaneous potential logging value; and SP_{Shale} and SP_{sand} are the spontaneous potential values of pure mudstone and pure sandstone, respectively.

$$D_R = \frac{AT90}{AT10} \times \frac{AT90}{AT20} \times \frac{AT90}{AT30} \times \frac{AT90}{AT60} \quad (6)$$

where D_R is the resistivity difference parameter, and $AT10$, $AT20$, $AT30$, $AT60$ and $AT90$ are the resistivity logs at depths of 10 in, 20 in, 30 in, 60 in, and 90 in the wellbore, respectively. The value of D_R is large for the oil layer, while the value for

the water layer is small.

$$R_{wa} = \frac{R_t * \phi^m}{ab} \quad (7)$$

where R_{wa} is the apparent formation water resistivity calculated by the Archie formula when the reservoir water saturation is assumed to be 100%, m is the cementation index, and a and b are the cementation indices.

$$R_{wa_SP} = \frac{R_{mf}}{10^{U_{SSP}/K}} \quad (8)$$

where R_{wa_SP} is the resistivity of the pure water layer calculated by spontaneous potential logging data. R_{mf} is the resistivity of the mud filtrate, and U_{SSP} is the static spontaneous potential value. K is the diffusion adsorption electromotive force coefficient. In water-saturated layers, R_{wa} is equal to or less than R_{wa_SP} , and with the increase in reservoir oil saturation, R_{wa} is higher than R_{wa_SP} .

The output characteristics are represented by digital labels representing different fluid types, in which the number 2 represents the oil layer, the number 1 represents the oil-water layer, the number -2 represents the water layer, and the number -1 represents the dry layer. According to the oil test conclusion of the target interval in the study area, the input logging parameters are matched and combined with the numbers representing different pore fluid types to form the input training set of the model. To ensure the effectiveness and representativeness of the input training set, 204 training samples are selected in the study area, of which 185 are training sample sets and 19 are test sample sets. Table 1 shows the logging parameters and oil test results of these 19 test sample sets.

Table 1 The logging parameters and oil test results of these 19 test sample sets

NO.	Testing interval (m)	$(K / \phi)^{1/2}$	Dr	QT	Rt	ΔSP	R_{wa}	R_{wa_SP}	Oil testing results
1	2502.7-2503.5	0.231	0.621	21.161	146.661	0.675	2.364	0.258	Oil layer
2	2516.4-2520.7	0.170	2.344	1.863	119.823	0.702	1.348	0.327	Oil layer

NO.	Testing interval (m)	$(K/\phi)^{1/2}$	Dr	QT	Rt	ΔSP	R_{wa}	R_{wa_SP}	Oil testing results
3	2565-2571.3	0.274	1.648	9.201	61.624	0.697	1.146	0.274	Oil layer
4	2356.1-2360.5	0.139	1.014	0.53	16.735	0.657	0.294	0.248	Oil layer
5	2531.6-2533	0.178	1.755	1.131	61.118	0.770	0.627	0.184	Oil layer
6	2590-2595.8	0.272	0.814	0.483	45.347	0.770	0.666	0.246	Oil layer
7	2397.6-2401.8	0.166	1.087	2.409	32.619	0.760	0.369	0.229	Oil layer
8	2469.4-2472.9	0.170	0.947	1.281	9.353	0.852	0.215	0.272	oil-water layer
9	2813.5-2816.2	0.216	1.492	0.945	13.018	0.713	0.241	0.236	oil-water layer
10	2652.8-2656.8	0.359	0.878	2.545	13.015	0.858	0.253	0.155	oil-water layer
11	2607.4-2609.8	0.216	1.492	0.945	13.018	0.713	0.236	0.241	oil-water layer
12	2614.2-2618	0.496	0.687	2.713	7.585	0.776	0.248	0.156	oil-water layer
13	2544.3-2548.7	0.297	0.992	1.179	208.161	0.797	2.750	0.292	oil-water layer
14	2602-2605.3	0.151	1.055	1.06	43.868	0.812	0.463	0.244	oil-water layer
15	2696.4-2698.8	0.756	0.707	0.375	4.898	0.76	0.174	0.294	Water layer
16	2595-2600.5	0.144	0.670	0.313	8.736	0.833	0.149	0.142	Water layer
17	2665-2667.2	0.583	0.390	3.643	6.832	0.793	0.166	0.197	Water layer
18	2819.1-2822	0.208	0.838	2.254	7.603	0.864	0.183	0.098	Water layer
19	2527-2529.2	0.080	0.303	2.552	104.618	0.421	0.692	0.290	Dry layer

The input sample set data have different physical meanings and different dimensions and orders of magnitude, and it is necessary to normalize the original data before learning and training. The normalization method selected in this paper is the mapminmax function, and its normalization formula is:

$$\$ = 2 * (x - x_{\min}) / (x_{\max} - x_{\min}) - 1 \quad (9)$$

where $\$$ is the normalized data, x is the input data, $x_{\max}$ and $x_{\min}$ are the

maximum and minimum values of the input data, and the range of normalized data is between -1 and 1.

The libsvm toolbox in MATLAB software is used for SVM classification model learning and training, and the radial basis function is selected as the kernel function, that is, $K(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2})$. The combination of grid search and k-fold cross validation is used to determine the best penalty factor (C) and kernel function parameters ($\sqrt{2}\sigma$), that is, the different combinations of penalty factor and kernel function parameters are selected to calculate the mean square errors obtained through training, and one group with the smallest mean square error is obtained as the optimal parameters. where the values of C are $(2^{-5}, 2^{-4}, \dots, 2^{12})$ and the values of $\sqrt{2}\sigma$ are $(2^{-5}, 2^{-4}, \dots, 2^5)$. Fig. 3 shows the plan maps of the mean square error and correlation coefficient trained by the 4-fold cross validation method under different C and $\sqrt{2}\sigma$ parameter combinations. By looking for the penalty factor and kernel function parameters with the smallest mean square error and the highest correlation coefficient of 19 test sample sets, the optimal penalty factor and kernel function parameter combination of the classification model is $C = 4096$ and $\sqrt{2}\sigma = 2$.

Figure 3. (a) The mean square errors of testing sample sets with different combinations of penalty factors and kernel function parameters, (b) the correlation coefficient of testing sample sets with different combinations of penalty factors and kernel function

parameters.

2.2 SVR regression model

The permeability and saturation of unconventional reservoirs are seriously affected by pore structure, and it is difficult to obtain these two parameters based on conventional logging curves. Therefore, the support vector regression (SVR) method is considered to construct the prediction model of reservoir permeability and saturation. The idea of using SVR to build a reservoir parameter prediction model is the same as the basic process of the SVM classification model, which is to first select the optimal dataset with high correlation to the prediction target value as the input. The relationship between permeability, saturation and logging curve is very complex. To determine the appropriate input training set, different logging data set combinations were used as the input training data, and the optimal input data set was selected by comparing the errors of the prediction model. The combination of different input logging data sets is shown in Table 2, including logging curves reflecting the reservoir lithology (Δ SP and Δ GR), logging curves reflecting reservoir physical properties (DEN, AC, CNL), logging curves reflecting reservoir electrical properties (RT), and reservoir porosity logging curve (POR) calculated by the core calibration logging curve method.

From the 16 wells in the study area, approximately 252 reliable and representative closed coring wells are selected to analyze the reservoir permeability and saturation, which are combined with the relevant logging curve data of the corresponding depth points to form the input training sample sets, 50 samples are randomly selected for back judgment, and the optimal input training sample set combination is selected according to the average relative error of the back judgment results. Fig. 4 shows the change in the average relative error of the regression permeability model and regression saturation model when using different input data sets. Combination 4 has the smallest (10.7%) average relative error to predict reservoir permeability, which reflects that reservoir permeability is jointly affected by porosity and shale content. Adding porosity data cannot improve the accuracy. Combination 5 has the smallest (2.1%) average relative error to predict reservoir saturation. From different average relative errors, the

average relative error of the saturation regression model changes little from combination 2 to combination 5, basically floating up and down by 2%, which illustrates that the porosity data calculated by conventional methods can improve the accuracy, but it is not obvious, which also shows that the reservoir saturation is mainly related to the electrical and comprehensive physical properties of the reservoir. Therefore, the optimal input training data set by the SVR regression permeability model is finally selected as combination 4, and the optimal input training data set by the SVR regression saturation model is combination 5.

Table 2. The combination of different input logging data sets

NO.	Input logging data sets	Prediction parameter
Combination 1	DEN, AC, CNL	Permeability
Combination 2	RT, DEN, AC, CNL	
Combination 3	DEN, AC, CNL, Δ GR	
Combination 4	DEN, AC, CNL, Δ SP	
Combination 5	DEN, AC, CNL, Δ SP,POR	
Combination 1	DEN, AC, CNL	Saturation
Combination 2	RT, DEN, AC, CNL	
Combination 3	RT, DEN, AC, CNL, Δ GR	
Combination 4	RT, DEN, AC, CNL, Δ SP	
Combination 5	RT, DEN, AC, CNL, Δ SP,POR

Figure 4. Characteristics of the average relative error by using different input data set combinations

3、Application effect analysis

3.1 SVM classification model

To evaluate the reliability of the SVM classification model for fluid recognition, the conventional fluid recognition method (cross plot of porosity and resistivity log) is introduced and compared with the commonly used artificial neural network algorithm (BP neural network). The input parameters of the BP neural network prediction model are the same as those of the SVM classification model. The optimal number of neuron layers is two layers, and the number of neurons in each layer is 12 and 14. The training function adopts the gradient descent adaptive learning rate function (traingdx function). Table 3 shows the comparison of fluid identification results of 19 new test sample sets by using the SVM classification model, conventional fluid recognition method and BP neural network model. The SVM classification model has the highest fluid identification accuracy (89.473%), followed by the BP neural network method (78.947%), and the conventional fluid recognition method has the lowest fluid identification accuracy (73.684%). This shows that using the SVM classification model to identify the low resistivity contrast oil layer is effective and feasible. Moreover, compared with the BP neural network model, the SVM classification model has certain advantages in solving the problem of small sample training, stronger generalization ability and better stability.

Table 3. Comparison of fluid identification results by different methods

NO.	Testing interval (m)	Oil testing results	SVM model	BP neural network	Conventional method
1	2502.7-2503.5	Oil layer	Oil layer	Oil layer	Oil layer
2	2516.4-2520.7	Oil layer	Oil layer	Oil layer	Oil layer
3	2565-2571.3	Oil layer	Oil layer	Oil layer	Oil layer
4	2356.1-2360.5	Oil layer	Oil layer	Oil-water layer	Oil layer
5	2531.6-2533	Oil layer	Oil layer	Oil layer	Oil layer
6	2590-2595.8	Oil layer	Oil layer	Oil-water layer	Oil layer
7	2397.6-2401.8	Oil layer	Oil-water layer	Oil layer	Oil-water
8	2469.4-2472.9	Oil-water layer	Oil-water layer	Oil layer	Oil-water

NO.	Testing interval (m)	Oil testing results	SVM model	BP neural network	Conventional method
9	2813.5-2816.2	Oil-water layer	Oil-water layer	Oil-water layer	Oil layer
10	2652.8-2656.8	Oil-water layer	Oil-water layer	Oil-water layer	Oil layer
11	2607.4-2609.8	Oil-water layer	Oil-water layer	Oil-water layer	Oil-water
12	2614.2-2618	Oil-water layer	Oil-water layer	Oil-water layer	Oil-water
13	2544.3-2548.7	Oil-water layer	Oil-water layer	Oil layer	Oil layer
14	2602-2605.3	Oil-water layer	Oil-water layer	Oil-water layer	Oil-water
15	2696.4-2698.8	Water layer	Water layer	Water layer	Water layer
16	2595-2600.5	Water layer	Water layer	Water layer	Water layer
17	2665-2667.2	Water layer	Water layer	Water layer	Water layer
18	2819.1-2822	Water layer	Water layer	Water layer	Water layer
19	2527-2529.2	Dry layer	Oil layer	Dry layer	Oil layer
Accuracy			89.473%	78.947%	73.684%

3.2 SVR model

Fig. 5 is the log interpretation result of a well (M165) with a low contrast oil layer in our study area, in which the testing interval is 2590 m-2596.5 m, and the average resistivity is 12.6 Ω .m. The 8th and 9th tracks in Fig 5 are the calculation results of reservoir permeability and saturation, respectively. The blue solid line is the permeability curve calculated by the conventional model (core calibration logging curve method), and the yellow solid line is the permeability curve predicted by the SVR model. The 9th blue solid line is the saturation calculated by the Archie saturation model, and the yellow solid line is the saturation curve predicted by the SVR model. The reservoir parameters calculated by the SVR model are more consistent with the core analysis results.

In addition, the calculated permeability and saturation by using the SVR model and the conventional model are compared with the core analysis data of 129 sealed cores from 12 wells (Figure 6). The results show that the average relative error of permeability calculated by the conventional method is 0.385, and the permeability predicted by the SVR model is 0.259. The average relative error of water saturation calculated by the Archie model is 0.188, while the saturation predicted by the SVR model is 0.097. This further verifies that the constructed

SVR prediction model is feasible and effective.

Figure 5. Comparison of reservoir permeability and saturation calculated by the SVR regression model and conventional method (Well M165)

Figure 6. The comparison results of reservoir permeability (a) and saturation (b) calculated by the SVR regression model and conventional method, respectively.

4. Conclusions

(1) There is no obvious difference in physical and electrical properties between the low resistivity contrast oil layer and water layer in the tight sandstone reservoir of the Chang 8 member in the Huanxian area, Ordos Basin. It is difficult to effectively identify and

evaluate low resistivity contrast oil layers by using conventional logging data, which seriously restricts the exploration progress and development benefits of oil and gas resources in this area.

(2) This study analyzed the relationship between the logging response and pore fluid to optimize the input training dataset. The SVM learning method was used to construct the SVM classification model and SVR regression model for fluid identification and reservoir parameter prediction.

(3) The application results show that the SVM classification model has higher fluid identification accuracy than the BP neural network method and conventional fluid recognition method (cross plot of porosity and resistivity log). The reservoir permeability and saturation predicted by the SVR regression model are more consistent with the core analysis results, which proves that the SVM method is effective and feasible for low resistivity contrast oil reservoir interpretation.

References:

- [1] Pratama, E., Ismail, M. S., and Ridha, S. An integrated workflow to characterize and evaluate low resistivity pay and its phenomenon in a sandstone reservoir. *Journal of Geophysics and Engineering*. 2017, 14(3):513-519.
- [2] Bai, Z., Tan, M.J., Shi, Y.J., Li, G.R., and Clark, S.M. Reservoir characteristics and control mechanism of resistivity low-contrast oil pays in Chang 8 tight sandstone of Longdong West area, Ordos Basin. *Journal of Petroleum Exploration and Production Technology*. 2021, 11(6):2609-2620.
- [3] Mode, A.W., Anyiam, O.A., and Aghara, I.K. Identification and petrophysical evaluation of thinly bedded low-resistivity pay reservoir in the Niger Delta. *Arabian Journal of Geosciences*. 2015, 8(4):2217-2225.
- [4] Hakimov, N., Zolfaghari, A., Kalantari-Dahaghi, A. Negahban, S., and Gunter, G. Pore-scale network modeling of microporosity in low-resistivity pay zones of carbonate reservoir. *Journal of Natural Gas Science and Engineering*. 2019, 71: 10.1016/j.jngse.2019.103005.
- [5] Bai, Z., Tan, M.J., Li, G.R., and Shi, Y.J. Analysis of low-resistivity oil pay and

- fluid typing method of Chang 8(1) Member, Yanchang Formation in Huanxian area, Ordos Basin, China. *Journal of Petroleum Science and Engineering*. 2019, 175:1099-1111.
- [6] Shi, W.R., Zhang, Z.S., Huang, Z.S., Jiang, S., Shen, J.C., Feng, A.G., Zhao, H.Y., and Xing, J. Investigation of the Origin of Low Resistivity and Methods for the Calculation of Gas Saturation in Shale Gas Reservoirs in the Fuling Area. *Energy & Fuels*. 2021, 35(6):5181-5193.
- [7] Kadhim, F.S., Samsura, A., Idris, A.K., and Al-Dunainawi, Y. The use of artificial neural network to predict correlation of cementation factor to petrophysical properties in Yamamma formation. *International Journal of Oil Gas and Coal Technology*. 2017, 16(4):363-376.
- [8] Shahriari, M., Pardo, D., Picon, A., Galdran, A., Delser, J., and Torres-Verdin, C. A deep learning approach to the inversion of borehole resistivity measurements. *Computation Geosciences*. 2020, 24(3):971-994.
- [9] Wang, F., Yang, X.M., Zhang, Y.H., and Bian, H.Y. Application of multialgorithmic fusion methods to fluid identity in tight sand reservoir. *Progress in Geophysics*. 2015, 30(6):2785-2792.
- [10] Liu, D., Pan, B.Z., Zhou, Y.F., Wang, Y.B., and Huang, S.H. Fluid identification method based on high-resolution array induction logging using GA-SVM. *Progress in Geophysics*. 2017, 32(5):2051-2056.
- [11] Guo, Y.H., Pan, B.Z., Jiang, B.C., Liu, S.H., Fang, C.H., and Li, D. Tight sandstone reservoir evaluation by the combination of three-water model and mathematical method in Sulige Area. *Geophysical Prospecting for Petroleum*. 2015, 54(5):621-626.
- [12] Chen, M., and Peng, Z.L. Study on Low Resistivity Oil Layer of Qudi Oilfield Based on BP Neural Network Recognition. *Journal of Ningxia University(Natural Science Edition)*. 2020, 41(1):98-103.
- [13] Singh, Y., Nair, R.R., Singh, H., Datta, P., Jaiswal, P., Dewangan, P., and Ramaprasad, T. Prediction of gas hydrate saturation throughout the seismic section in Krishna Godavari basin using multivariate linear regression and multilayer feed

- forward neural network approach. *Arabian Journal of Geosciences*. 2016, 9(5): 10.1007/s12517-016-2434-6.
- [14] Miah, M.I., Zendehboudi, S., and Ahmed, S. Log data-driven model and feature ranking for water saturation prediction using machine learning approach. *Journal of Petroleum Science and Engineering*. 2020, 194: 10.1016/j.petrol.2020.107291.
- [15] Baneshi, M., Behzadijo, M., Schaffie, M., and Nezamabadi-Pour, H. Predicting Log Data by Using Artificial Neural Networks to Approximate Petrophysical Parameters of Formation. *Petroleum Science and Technology*. 2013, 31(12):1238-1248.
- [16] Roslin, A., and Esterle, J.S. Electrofacies analysis for coal lithotype profiling based on high resolution wireline log data. *Computers & Geosciences*. 2016, 91:1-10.
- [17] Yang, H., Li, S.X., and Liu, X.Y. Characteristics and resource prospects of tight oil and shale oil in Ordos Basin. *Acta Petrolei Sinica*. 2013, 34(1):1-11.
- [18] Fu, J.G., Dong, G.D., and Zhong X.P et al. Research progress of petroleum geology and exploration technology in Ordos Basin. *China Petroleum Exploration*. 2021, 26(3):19-40.
- [19] Zhao, J.L., Yan, B., Zhao, J.Z., Wang, Y.P., Xu, D.C., and Gao, X.L. Reservoir characteristics of Chang 8 and single-well productivity prediction of M wellblock in Huanxian oil district, Longdong area. *Oil & Gas Geology*. 2013, 5:694-699.
- [20] Chai, G.S., Shi, Y.M., and Du, S.H. et al. Sensitivity Evaluation and Influencing Factors Analysis of Tight Sandstone Reservoirs: A Case Study of the Chang-8 Reservoir in Yanchi Area of Ordos Basin. *Acta Scientiarum Naturalium Universitatis Pekinensis*. 2020, 56(2):253-261.
- [21] Vapnik, V.N. *Statistical learning theory*. New York: JWiley, 1998.
- [22] Hsu, C.W., Lin, C.J et al., A comparison of methods of multiclass support vector machines. *IEEE Transactions on Neural Networks*. 2002, 13(2):415-425.
- [23] Peng, T., and Zhang, X. Review on support vector machine and its applications in petroleum exploration and development. *Progress in Exploration Geophysics*. 2007, 30(2):91-95.