SPE-214059-MS

Research on Pore Pressure Prediction Technology of HTHP Wells in South China Sea Based on Machine Learning

Dongsheng Xu, Jin Yang, Yuhang Zhao, and Jianchun Fan, China University of Petroleum, Beijing; Yanjun Li, China University of Petroleum, Beijing & CNOOC China Limited-Hainan; Xun Liu, CNOOC China Limited-Tianjin; Kejin Chen, Zehua Song, Xun Zhang, and Hong Zhu, China University of Petroleum, Beijing

 DOI 10.2118/214059-MS

This paper was prepared for presentation at the Gas & Oil Technology Showcase and Conference held in Dubai, UAE, 13 - 15 March, 2023.

Abstract

The Yingqiong Basin in the South China Sea is located at the intersection of the Eurasian and Indo-Chinese plates, with complex geology and often accompanied by abnormally high pressure. In this paper, we analyze the causes of anomalous high pressure in the South China Sea and analyze the commonly used machine learning methods, support vector machine and BP neural network, and use both methods to predict a block in Yingqiong Basin. The field application was carried out using this method, and the application showed that the prediction accuracy exceeded 95%, the complexity was reduced by 42%, and the drilling efficiency was improved by more than 53%, which played a good guiding effect to the field.

Keywords: offshore drilling, formation pressure prediction, support vector machine, high temperature and pressure, BP neural network

Introduction

The South China Sea is one of the three largest HTHPs in the world, with a formation pore pressure of up to 2.35. The cause of the anomalous pressure is still unclear, severely limiting the exploration and development of the Yinggehai and Qiongdongnan basins[1]. HTHP natural gas resources are expected to account for 2/3 of the overall resources, with huge resource potential. The complex geological conditions and the development of bottom and watercourse sands pose a great challenge for pre-drill formation pressure prediction[2]. Due to the poor understanding of the formation pressure mechanism in this area, the pressure prediction error is large. In the case of formation overpressure, the pressure window is extremely narrow, and downhole complications such as well leaks and well surges are often frequent, often leading to frequent downhole complications. There are several methods commonly used in formation pressure prediction, such as Fillippone method, empirical coefficient method, Eaton model, Bowers model, equivalent depth method, etc. Also the percentage of different formation pressure prediction methods can be adjusted by the influence factor. In this paper, based on the previous research, in order to predict the formation pore pressure more accurately, we put the machine learning method for intelligent prediction of formation pore pressure, and the prediction results show that the prediction accuracy is more accurate[3].

Mechanism and Identification of Anomalous HTHP Formation

The formation of a certain anomalous high pressure body in underground mainly depends on two factors: firstly, it has the inducing mechanism and conditions for the formation of anomalous high pressure; secondly, it has a good confining cap for storing high pressure. Among them, the genesis mechanism is the prerequisite for the formation of anomalous high pressure, while a good confining cap is necessary for the storage of the generated anomalous high pressure. The study of anomalous high pressure formation mechanism can help to understand and deepen the understanding of anomalous HTHP phenomenon and establish a reasonable model of formation pore pressure[4].

Mechanism of Anomalous HTHP Formation

Anomalously high pressure may be caused by a variety of overlapping factors, including geological, physical, chemical, and kinetic factors. However, for a particular pressure body, the formation may be caused by one factor dominated by others[5-7].

Unbalanced Compaction. During the burial and deposition of a formation, pore water is discharged from the sediment under the force and the formation is compacted. The unbalanced compaction action to produce anomalous high pressure should have the following conditions: (i) rapid sediment compaction of the formation; (ii) huge thickness of sediment; (iii) good cover layer to restrict the discharge of pore water and store the anomalous high pressure. The size of the anomalous high pressure generated by unbalanced compaction is related to the deposition rate of the formation, and the faster the deposition rate under the same confining conditions, the larger the anomalous high pressure will be generated.

Tectonic Extrusion. In the area of intense tectonic compression, the strata will be strongly compressed by the tectonic stress, and the sedimentary compaction of the strata will be controlled by the stresses in three directions (one vertical direction and two horizontal directions) simultaneously, and the strata will be in a three-dimensional compaction state. When the drainage rate of the formation is smaller than the compaction rate of the formation, it will cause the increase of the pore pressure of the formation and form the abnormal high pressure. Tectonic compression is similar to the mechanism of unbalanced compaction, but the difference is that unbalanced compaction is only controlled by vertical stress, while tectonic compression is controlled by the stress in three directions simultaneously.

Hydrothermal Compaction. With the increase of the burial depth of the stratum, the temperature of the stratum is also gradually rising, because the coefficient of thermal expansion of water is greater than the coefficient of thermal expansion of rock, so the expansion of pore water than rock expansion is more obvious volume. The increase in pore volume will reduce the effective stress, then the pore fluid needs to bear a greater part of the overlying rock pressure, and then the formation of abnormal high pressure. For the hydrothermal pressurization mechanism is still some controversy, some scholars believe that the hydrothermal pressurization mechanism is not enough to make the formation produce anomalous high pressure.

Abnormal HTHP Identification Steps

According to the different response characteristics of acoustic velocity and density when the formation is unloaded, a discriminative method based on the acoustic velocity-density rendezvous diagram can be established to distinguish the loading class and unloading class pressure-forming mechanism, and the basic operation steps to realize this discriminative method are as follows[8-9].

1. Logging data collection and processing: collect logging data such as acoustic velocity, density and natural gamma of the target well, and correct these logging data for processing. This includes range checking, correction of environmental influences, stratification of logging curves, filtering and smoothing, etc. The purpose is to remove artifacts and obtain true and accurate logging data.
2. Mudstone logging parameters extraction: Stratify the natural gamma data, obtain the stratified mud content, and extract the acoustic velocity and density data of mudstone according to the calculated stratified mud content. The mudstone velocity and density data are interpolated to obtain depth-aligned and equally spaced mudstone logging data.
3. Preliminary judgment of stratigraphic unloading: Draw a graph of each data under the same depth coordinates, and make a preliminary judgment of whether the sonic and density are reversed according to the trend of the curve, and determine the stratigraphic section where unloading may occur.
4. Intersection map drawing: extract the most reliable mudstone acoustic velocity and density data near the reversal point, and then draw the acoustic velocity-density intersection map. In order to facilitate the drawing and discrimination of the rendezvous map, it is better to convert the units of acoustic velocity and density data to km/s and g/cm3, respectively.
5. Discriminate between loading and unloading mechanisms: judge whether the stratum is in loading or unloading state according to the rendezvous map, and then determine the formation mechanism of anomalous high pressure by combining with geological conditions. When judging the unloading of the formation, pay attention to exclude the possibility of reversing the logging data due to lithological changes.

Causes of Anomalous HTHP in Yingqiong Basin

Due to the wide variety of anomalous high pressure causes and the small amount of available information in the target block, it is impossible to analyze each cause individually to determine whether it has an impact on the anomalous high pressure in the target block[10]. Therefore, in this section, based on the general understanding of the field experts and the limited data available in the target area, we analyze the pressure transfer, stratigraphic undercompaction, and tectonic extrusion causes separately.

The Miocene Sanya and Meishan formations of the Yinggehai Basin in the South China Sea are dominated by shallow marine mudstones, which are the main hydrocarbon source rocks in the basin. The bottom incipient zone is located in the center of the main hydrocarbon producing depression and has a strong gas source base. The hydrocarbon source rocks of Meishan Formation in this area are distributed in large area and thickness, and are in favorable gas window. The target is located at the bottom of Huangliu Formation, underlain by Miocene Sanya Formation and Miocene Meishan Formation, which has rich gas source base. The target is close to the up-dip part of the bottom, and some faults are clearly visible in the seismic section, which have been broken down into the hydrocarbon source rocks of Meishan Formation-Sanya Formation, and broken up into a sand body of Huangliu Formation and ended in the overlying large set of mudstone, effectively communicating with the hydrocarbon source rocks of Meishan Formation-Sanya Formation. The structure is adjacent to the main hydrocarbon source rock layer of the underlying Meishan Formation and Sanya Formation[11]. The mature hydrocarbon fluids are driven by the overpressure in the source rock, and are transported from the lower high potential area to the upper low potential area by the vertical microfractures in the bottom-opening activity wave area as the transport system, while gathering in the dominant transport direction under the control of the tectonic ridge background, and finally enriching in the watercourse sand reservoir trap to form an upward-dipping sharp extinction type rocky gas reservoir. The overlying thick layer of shallow marine high-pressure mudstone cover provides the necessary conditions for late gas preservation.

Predictive Model

There are many machine learning algorithms, and in this paper, we choose two commonly used and effective algorithms, support vector machines and BP neural networks, and use data from actual drilling to train the models of the two methods, and then compare the computational effect and stability.

Introduction to the Support Vector Machine Method

Support vector machine is a supervised learning model that can be used for the analysis of pattern recognition, classification and regression problems. The model is based on statistical learning theory and the principle of structural risk minimization. The main idea is to build a linear separation hyperplane that maps nonlinear inputs into a high-dimensional feature space so that the separation edge between positive and negative examples is maximized[12-13]. The principle of support vector machine is shown in Fig. 1, x_1 and x_2 represent the input parameter features, w is the kernel function, b is the bias, and I_2 represents the best solution for data classification.

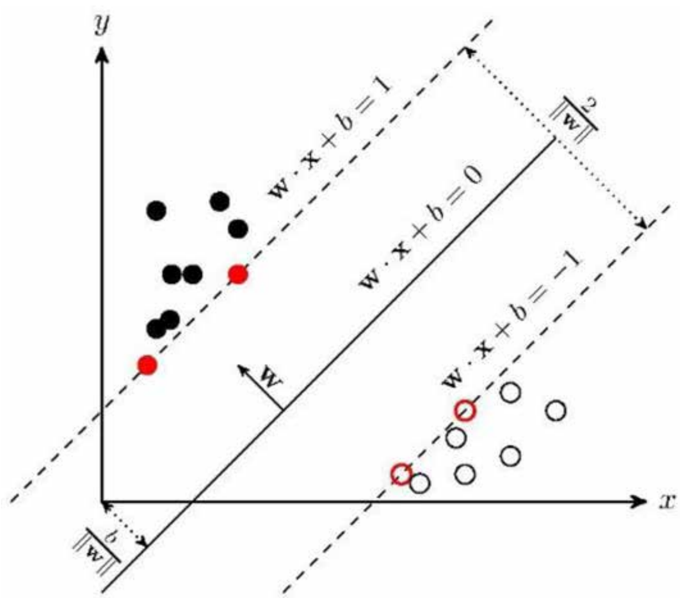

Figure 1—Support vector machine principle

The sample set of the deep well formation stress field prediction problem is: $\{(x_1, y_1), (x_2, y_2)\ldots(x_i, y_i)\ldots(x_l, y_l)\}$,where $x_i \in R^n$, an n-dimensional attribute vector, is the spatial coordinate value: $y_i \in R$, the property of the attribute vector, is the stress field value.

Using the ε-insensitive loss function, the problem of minimizing the risk to the ε-insensitive loss function is transformed into an optimization problem of its dyadic form, and the fitting function f(x) = (w-x) + b is constructed through the training set, the support vector machine is

$$\min\varphi = \frac{\|w\|^2}{2} + c\cdot\sum_{i=1}^{l}(\xi_i + \xi_i^*) \tag{1}$$

where: w is the weight of x; c is the penalty factor, which indicates the degree of penalty for samples exceeding the error ε; ξ_i, ξi^* are the relaxation variables.

Find equation (1) to minimize the objective function pairwise problem can get the fitting function, for the nonlinear problem, transform the data set into a linear regression problem in high-dimensional Hilbert feature space by mapping, perform linear regression in the high-dimensional feature space, use the kernel function to replace the inner product of 2 training points in the high-dimensional Hilbert space, and calculate the nonlinear regression function (decision function):

Where: K(x, xi) is the kernel function; α_i, αi*, b are the optimal solutions obtained by quadratic programming, see equation (3).

$$\max\delta(a, a^*) = \frac{1}{2}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(\alpha_i \cdot \alpha_j) + \sum_{i=1}^{l} y_i(\alpha_i - \alpha_i^*) -)\sum_{i=1}^{l} y_i(\alpha_i + \alpha_i^*) \tag{2}$$

The constraints are.

$$s.t.\begin{cases} \sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0, i = 1, 2, \cdots, l \\ 0 \le \alpha_i, \alpha_i^* \le c, i = 1, 2, \cdots, l \end{cases} \tag{3}$$

There are 4 types of kernel functions in common use

Linear kernel function (Linear)

$$k(x, x') = x \cdot x' \tag{4}$$

Polynomial kernel function (Poly)

$$k(x, x') = (x \cdot x' + m)^p \tag{5}$$

where: m and p are kernel function parameters, $m \ge 0$, and p is a positive integer.

Gaussian radial basis function (RBF):

$$k(x, x') = \exp(-\frac{\| x - x' \|}{\sigma^2}) \tag{6}$$

where: σ is the width of the kernel function, $\sigma > 0$.

Sigmoid kernel function (Sigmoid)

$$k(x, x') = \tanh[v(x \bullet x') + \theta] \tag{7}$$

where: v, θ为 is the kernel function parameter, >0, >0.

BP Neural Network

BP neural network is an artificial neural network based on error back propagation algorithm, which has many advantages. Therefore, BP neural network is widely used in oil, transportation, chemical industry, etc[14-17]. The structure of BP neural network generally consists of three layers: input layer, hidden layer and output layer. Its topology is shown in Fig.2.

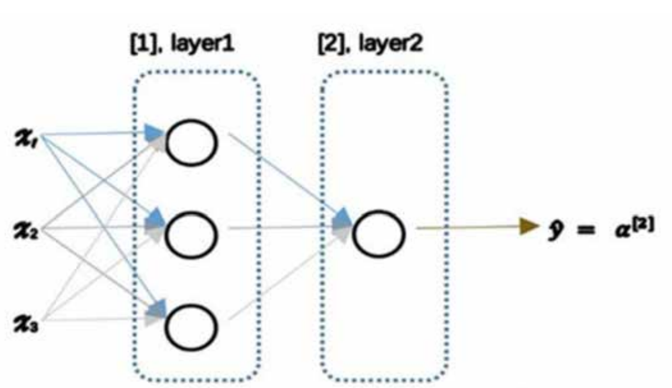

Figure 2—Principle of BP neural network

The model training process of BP neural network has two main parts: forward propagation of the signal and backward propagation of the error. The process is roughly divided into four steps: (1) forward propagation of the model, the input data is first fed into the BP neural network from the input layer, and then passed into the hidden layer after weighting calculation, and the information in the hidden layer is passed to the output layer after weighting calculation, and the model calculation results are generated in the output layer. (2) Error calculation, the results of the BP network calculation with the real value of the data, if the error is within the accuracy range can end the cycle, otherwise enter the third step. (3) Back propagation, according to the activation function selected by the network and the error value backward layer by layer to calculate the propagation value of the error. (4) Update the weights for the next iteration of the calculation.

Evaluation Method

In order to evaluate the stability and prediction accuracy of the model, two evaluation indexes, linear regression coefficient of determination (R^2) and mean square error (*mse*), are used. The linear regression coefficient of determination indicates the goodness of fit, which can reflect the goodness of fit results of the training set.

$$R^2 = SSR/SST - SSE/SST \tag{8}$$

$$SST = SSR + SSE \tag{9}$$

Where: *SST* is the sum of total squares; *SSR* is the sum of regression squares; *SSE* is the sum of squared residuals.

The *mse* denotes the expected value of the squared difference between the predicted and true values of the parameter, and the smaller its value, the more it indicates that the model has better accuracy.

$$mse = \sum_{i=1}^{n} \frac{1}{n}[f(x_i - y_i)] \tag{10}$$

Machine Learning Calculation and Analysis

The Yingqiong basin in the South China Sea is tectonically complex, influenced by the intersection of geological structures and fractures, and the coupling and interlocking influence of multiple genesis mechanisms, which results in complex formation high pressure genesis. Traditional pressure prediction methods have large errors and are very likely to cause downhole accidents and borehole scrap rates as high as 30%, so the development of machine learning formation pressure prediction technology can solve the calculation of formation pressure under the interlocking influence of multiple factors.

Algorithm Comparison Analysis

The algorithm is implemented using Python software. First, the sample dataset is imported, and the initial value of the number of training sets is set to 600 data from 6 wells and the number of test sets is 100, and the training and test sets are randomly generated in the sample dataset by learning training. A support vector machine (SVM) prediction model is created, where the loss function p-value is set to 0.1 and the kernel function type is set to RBF kernel function.

A BP neural network model is created, where the learning rate is set to 0.1, the maximum number of training is set to 1,000, the period of displaying intermediate results is set to 10, and the global minimum error is set to 0.0001.

The training sets of the above 2 models are learned and trained separately, and the models are validated on the test set after training, and the values of the 2 evaluation parameters, mean square error and linear regression coefficient of determination, are returned. Different number of training sets will have certain influence on the model prediction accuracy, therefore, the number of training sets needs to be adjusted continuously, and the number of training sets is incremented from 600 to 650 for model training one by one, and the returned *mse* and R^2 values are compared by statistical analysis, and the specific results are shown in Fig.3.

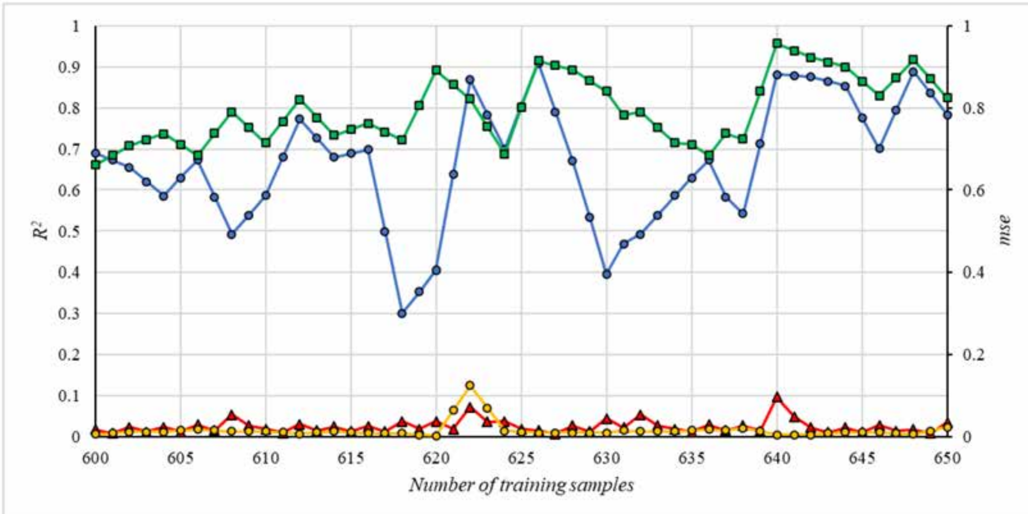

Figure 3—Performance of support vector machine and BP neural network with different training samples

As can be seen from Fig. 3, most of the R^2 data points of the support vector machine model are above the BP neural network model, indicating that the support vector machine training fit is higher and more stable. The linear regression coefficient of determination of the support vector machine model reaches the highest point when the number of training samples is 117, and the R^2 maximum value is 0.958; the linear regression coefficient of determination of the BP neural network model reaches the highest point when the number of training samples is 640, and the R^2 maximum value is 0.909.

In terms of prediction error, the *mse* value of the support vector machine model is basically in a low range, and in contrast, the *mse* value of the BP neural network fluctuates more and has a higher value. The *mse* value of the support vector machine model is the lowest when the number of training samples is between 642 and 649, and the minimum value of *mse* is 0.001. The *mse* value of the BP neural network model is the lowest when the number of training samples is 643 and 649, and the minimum value of *mse* is 0.003.

Comparing the curves of the 2 evaluation parameters in Fig. 1, it can be seen that the R^2 value of the support vector machine model is basically above that of the BP neural network model, while the opposite value of *mse*, so the stability and accuracy of the support vector machine model are better than that of the BP neural network model. Considering the evaluation indexes of *mse* and R^2, the model with smaller *mse* value and larger R^2 value is selected as the optimal model.

Case validation

Using the preferred support vector machine based formation pressure prediction method, the formation pressure prediction results match well with the actual field measurements, and the maximum error is not more than 5% calculated by the relative error rate, and some well sections achieve an error rate of less than 1%, and the complexity is reduced by 42% and the drilling efficiency is improved by more than 53% during the operation of the exploratory well[18].

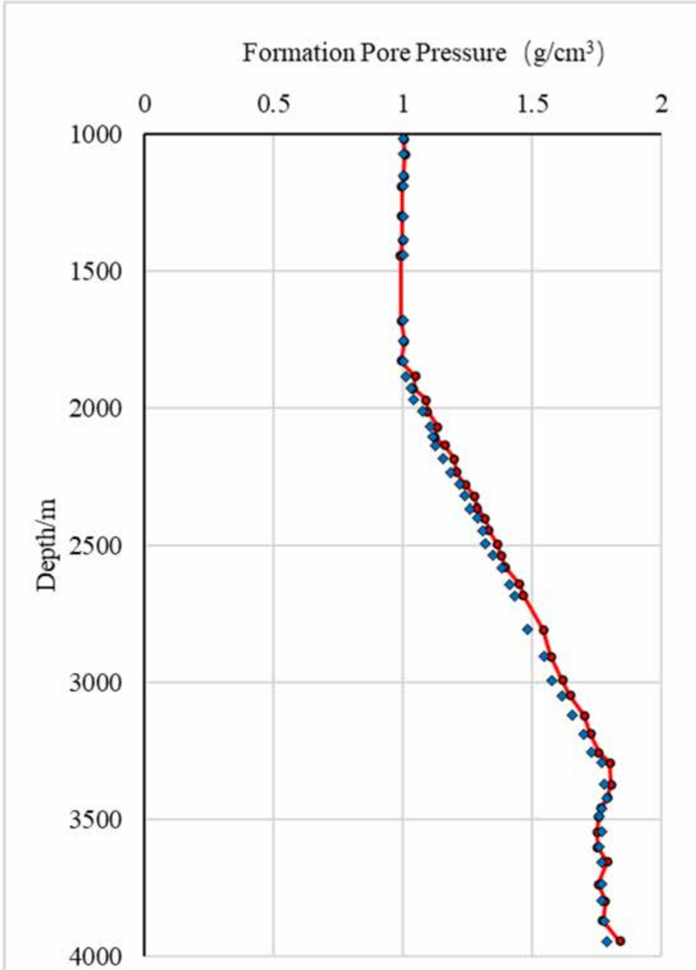

Figure 4—Comparison of support vector machine formation pressure prediction and actual measurement results for a HTHP well in Yingqiong basin(1000m-4000m)

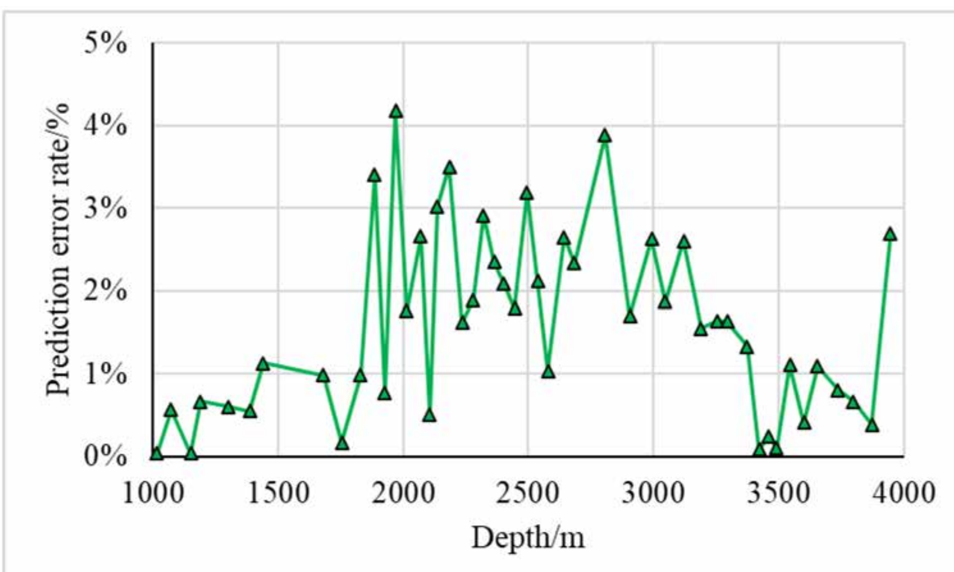

Figure 5—Error analysis of support vector machine formation pressure prediction and actual measurement results in a HTHP well in Yingqiong basin (1000m-4000m)

Through a high-precision formation pressure prediction method, combined with information from 12 exploratory wells, the data body of formation pressure in this block was mapped by finite element differencing, providing a reference for formation pressure data for wells that have not yet been developed.

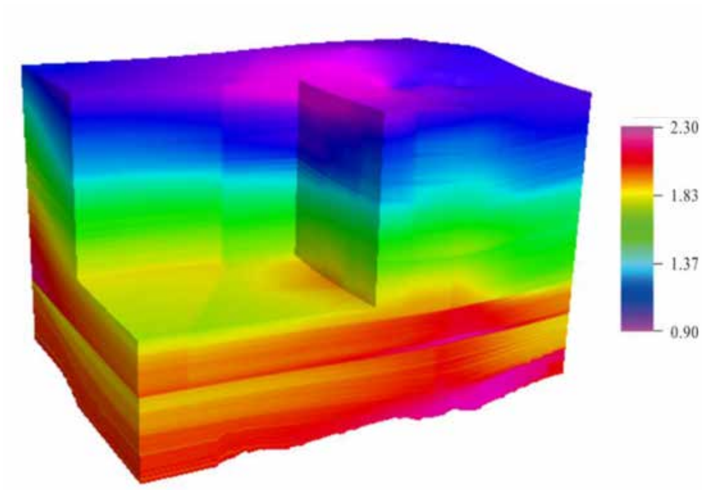

Figure 6—A HTHP certain block stratigraphic pressure data body in Yingqiong basin

Conclusion

1. The geological structure of Yingqiong basin in the South China Sea is complex, and the anomalous high-pressure genesis mechanism varies significantly. A scientific and reasonable selection of formation pressure prediction methods can effectively improve the accuracy of formation pressure prediction in the whole block and improve drilling efficiency.
2. Two machine learning methods, support vector machine and BP neural network, were compared, and the results showed that the prediction accuracy and stability of the support vector machine method were significantly better than the BP neural network.
3. After comparing the preferred support vector machine method, the application was carried out to the actual situation of the block. The application results showed that the prediction accuracy of this method is high, and the maximum error does not exceed 5%. The use of this method reduces the drilling complexity by 42% and improves the drilling efficiency by more than 53%, and the formation pressure data body of the whole block established by this method.

References

[1] Tong Xiao. Research on dynamic early warning system for oilfield development based on data mining [D]. Xi'an University of Petroleum, 2016.

[2] Han Shan, Che Mingguang, Su Wang, Xiao Yuxiang, Wu Zhongbao, Chen Jianyang, Wang Libin. Single-well production prediction method and application for shale gas in the Weiyuan block of Sichuan Basin[J]. *Special Oil and Gas Reservoirs*, 2022, **29**(06):141–149.

[3] Liu Shengwa, Cao Xianghua. Intelligent identification method of drilling conditions based on decision tree[J]. *New Industrialization*, 2022, **12**(01):28–30.

[4] Wang H. F. Differences in the tectonic evolution of Cenozoic basins in the South China Sea and its genesis analysis[D]. Xi'an University of Petroleum, 2018.

[5] Yang P, Xia B, Cai Zorong, Wan Zhifeng, Huang Qiangtai, Zhang Yong. Genesis mechanism of the Yingge Sea Basin in the northern South China Sea: Insights from comparative tectonic analysis with the Weihe Basin[J]. *Marine Geology and Quaternary Geology*, 2017, **37**(06):65–75.

[6] Yan Haiyuan, Shi Chenghui, Feng Yongcun, Wei Baohua, Li Bin, Zhang Ming, Deng Jingen. Study on the time-varying upper limit pressure of shear rupture in a high-temperature and high-pressure well in South China Sea[J]. *Drilling Technology*, 2021, **44**(05):11–15.
[7] Xie Jing, Wu Huimei, Lou Yishan, Zhai Xiaopeng. Prediction model of high-temperature and high-pressure formation rupture pressure in deepwater South China Sea[J]. *Fracture block oil and gas field*, 2021, **28**(03):378–382.
[8] Liu Aijun, Fan Caiwei, Wu Yunpeng, Li Fang. Research on pressure prediction method of it source based on transfer model in high temperature and high pressure field in South China Sea[J]. *China Offshore Oil and Gas*, 2021, **33**(01):50–55.
[9] Li Wentuo, Li Zhong, Liu Bing, Peng Hailong, Deng Wenbiao. Current status and outlook of formation pressure prediction technology in Yingqiong Basin, South China Sea[J]. *Offshore Petroleum*, 2020, **40**(03):8-12+18.
[10] Gao Hongfang, Zhong Hexian, Sun Meijing, Nie Xin, Jiang Tao, Huang Wenkai, Du Wenbo, Chen Jiale. Tectonic control of large deep-water turbidite fan system and its genesis in the southeastern South China Sea basin[J]. *Geology of China*, 2020, **47**(05):1395–1406.
[11] Li Huayang, Wu Huimei, Nan Chong, Zhao Baowen, Yan Ke, Wang Quanhang. Research and application of formation pressure prediction technology in L-1 well area in deepwater South China Sea[J]. *Petroleum Geology and Engineering*, 2020, **34**(03):113–117.
[12] Sun Ting, Zhao Ying, Yang Jin, Yin Qishuai, Wang Wenxing, Chen Yuan. A real-time intelligent identification method for drilling conditions based on support vector machine[J]. *Petroleum Drilling Technology*, 2019, **47**(05):28–33.
[13] Wang X, Zhang QZ. Improved sparrow search algorithm to optimize support vector machine for well leak prediction[J]. *Science Technology and Engineering*, 2022, **22**(34):15115–15122.
[14] Zhang LG, Miao ZH, Huang SG, Yuan SB. Drilling speed prediction of drilling machinery based on MEA-BP neural network[J]. *Automation and Instrumentation*, 2022, **37**(11):87–92.
[15] Bai Tingting. Research on oil formation identification and reservoir parameter prediction based on MPSO-BP neural network[D]. *Northeastern Petroleum University*, 2022.
[16] Song Prophet, Zhu Shuo, Li Gensheng, Zeng Yijin, Guo Huijuan, Hu Zhijian. Large hook load and turntable torque prediction based on BP-LSTM dual input network[J]. *Journal of China University of Petroleum (Natural Science Edition)*, 2022, **46**(02):76–84.
[17] Song Prophet, Yao Xuezhe, Li Gensheng, Xiao Lizhi, Zhu Zhaopeng. An LSTM-BP neural network-based method for pore pressure calculation in formations[J]. *Petroleum Science Bulletin*, 2022, **7**(01):12–23.
[18] Xiao P, Han C. Research and practice of integrated speed-up technology for high-temperature and high-pressure exploratory wells in the South China Sea Dongfang 13-X[J]. *Chemical Engineering and Equipment*, 2021,(11):51–53.