J. COMBIN. MATH. COMBIN. COMPUT. 127a (2025) 6199--6213

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Research on the construction of a working fluid system model for ultra-high temperature dense pressurized leakage prevention and plugging combined with multivariate nonlinear regression and machine learning optimization

Siqi Chen[1, ✉]

[1] *Chongqing City Management College, Chongqing, 401331, China*

ABSTRACT

At present, drilling fluid leakage in oil and gas drilling engineering in complex formations is a worldwide technical problem. The study explains the mechanism of dense pressure-bearing plugging at the bottom of the fracture, explores the influencing factors of the pressure-bearing capacity of the leakage prevention and plugging working fluid, and establishes a mathematical model by using multivariate nonlinear regression analysis. Based on the machine learning technology, the support vector machine algorithm is selected as the prediction method of the particle size of the working fluid for leakage prevention and plugging, and the system model of the ultra-high-temperature dense pressurized leakage prevention and plugging working fluid is constructed. It is found that the established multivariate nonlinear regression analysis has good fit and accuracy, and the average relative error is only 2.9%, and the seam width (-0.694) and formation pressure (0.502) have the greatest influence on the pressure-bearing capacity of the working fluid for leakage prevention and plugging. The prediction accuracy of the support vector machine model for the working fluid particle size was 95.36%, and the prediction F1 values on multiple datasets were all greater than 0.9, showing excellent prediction results. The constructed mathematical model can be used to guide the field operation, which is conducive to the long-term stable plugging and scientific leakage prevention of fissure-based leakage.

Keywords: multiple nonlinear regression analysis, machine learning, support vector machine; prediction model, leakage prevention and plugging

✉ Corresponding author.
E-mail address: Eland0177@163.com (S. Chen).
Received 15 January 2024; Revised 10 February 2024; Accepted 30 November 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-341

1. Introduction

At present, drilling fluid leakage in oil and gas drilling engineering in complex formations is a worldwide technical challenge. Well leakage in fractured formations accounts for about 80% or more of the total number of well leakage, often with faster leakage rate, larger leakage loss, low success rate of plugging, and its loss accounts for about 90% or more of the total consumption of well leakage [1-4]. Therefore, it is of great significance to construct a model of ultra-high-temperature dense pressure-bearing leakage prevention and plugging working fluid system based on multivariate nonlinear regression and machine learning optimization.

With the increasing development of oilfield technology, pressurized leakage plugging technology has been increasingly emphasized and studied. Pressurized leak plugging technology is a key technology widely used in oilfield production, which can effectively solve the problems of wellhead leakage and tubing column damage in oilfield production, and improve the field production efficiency and wellhead safety [5-8]. Pressurized leakage plugging technology mainly refers to the use of certain technical means to stop leakage and seepage by pressurizing and sealing the pipeline or pore to achieve the effect of stopping leakage and seepage when encountering leakage and seepage in the process of wellhead construction, debugging and production [9-11]. Pressurized leakage plugging technology research is mainly focused on the following aspects: plugging material research. Leakage plugging material is the core of pressurized leakage plugging technology. At present, the commonly used plugging materials include glue, seaweed powder, polymer hydrogel and so on. When choosing the leak plugging materials, it is necessary to consider its leak plugging efficiency, pressure resistance and environmental protection and other factors [12-15]. Leak plugging tool research. Leak plugging tool is a carrier to realize pressure leakage plugging, and its performance is an important factor affecting the effect of leakage plugging. At present, the plugging tools are generally divided into two categories: internal and external plugging, in which the internal plugging tools are more complex, and the plugging effect can be monitored in real time by logging tools such as MFT and PDM [16-19]. Research on plugging technology. The research of pressurized plugging technology includes screening plugging scheme, plugging construction debugging, plugging effect monitoring and other aspects. At present, the plugging technology is mainly applied to oil wells, natural gas wells and water injection wells [20-23].

The study analyzes the dense pressure-bearing plugging mechanism of leakproof plugging materials in fractured formations, selects the influencing factors of the pressure-bearing capacity of the leakproof plugging working fluid from the formation factors, fluid factors, and process factors, collects the relevant data through indoor experiments, and establishes a model using the multivariate nonlinear regression method to explore and check the fitting results of the model. Then, the collected well leakage data were analyzed and studied using support vector machine algorithm to establish a prediction model for the parameters of the leak prevention and plugging working fluid, and the working fluid particle size was classified into four categories, and the prediction performance of this paper's model for the particle size of the leak prevention and plugging working fluid was evaluated by using the F1 score through the testing of different machine learning methods on the training set, testing set and validation set.

2. Mechanism of dense pressurized containment in fractured formations

The existence of fractures in the exploration and development of oil and gas resources can lead to the leakage of working fluids, which in turn induces serious reservoir damage. Ultra-high-temperature dense pressurized leakproof plugging materials are carried into the annular wellbore by drilling fluids, and firstly, they have to overcome the depositional effect and sealing effect, and enter into the fractures under the effect of differential pressure in the wellbore. The plugging layer is formed by bridging, stacking and filling in the fracture, which is a typical granular material system composed of different types of leakage prevention and plugging materials. The dense pressure-bearing performance of the plugging layer mainly depends on two aspects:

1) Interaction between blocking layer and crack wall surface. Leak prevention and plugging materials into the crack, large particle size particles first bridging, medium and small particle size particles for filling and blocking step by step, forming a dense pressure-bearing plugging layer. Bridging is the basis for the formation of the blocking layer, and there is interaction between the leakage plugging particles and the crack wall surface. When the crack closure stress is greater than the compressive strength of the bridging particles, the particles will be crushed by extrusion and lead to blocking instability, therefore, the large-size bridging particles should have high compressive strength to prevent crushing and destabilization.

2) Densification and structural strength of the internal blocking layer. Leakage prevention and plugging materials in the cracks undergo transport, bridging and accumulation of filling and other dynamic processes to form an effective plugging layer, the ability of the plugging layer to resist infiltration leakage destabilization and shear dislocation destabilization, determines its internal densification and structural strength. As a typical granular material system, the plugging layer has a variety of physical mechanisms and multi-scale structural characteristics with “force chain” as the core, and the internal force chain network structure and its stability affect the macroscopic densification and structural strength of the plugging layer. When a large number of stable force chain network structure is formed inside the plugging layer, it is conducive to improving the stability of the overall structure of the plugging layer and forming a dense pressure-bearing plugging layer with strengthened force chain network structure.

3. Working fluid pressure capacity and parameter prediction methods

For ultra-high-temperature dense pressure-bearing leak prevention and plugging materials, in order to understand the influencing factors of their pressure-bearing capacity, a mathematical model of the pressure-bearing capacity of the leak prevention and plugging working fluid was established by the multivariate nonlinear regression analysis method, and then a prediction model of the leak prevention and plugging working fluid was established by using the machine learning method, which can predict the particle size of the working fluid of the workover fluid in order to prevent and plug the wells from designing the workover operations.

3.1. Working fluid pressure capacity model

Regression analysis is a statistical method for determining the interdependence between two or more variables. Its advantage is that it is easy and convenient to operate when analyzing multi-factor models, and it can accurately analyze the degree of correlation between the factors and improve the accuracy of the prediction equation. Multiple regression analysis is used to fit the relationship between the pressure capacity of leak prevention and plugging working fluid and the experimental factors.

The general nonlinear regression model can be expressed as:

$$y = f(x,b) + \varepsilon \tag{1}$$

Here, x is an observable independent random variable, b is the parameter vector to be estimated, and y is an independently observed variable whose mean depends on x & b,ε is the random error.

3.1.1. Model calculations. In order to introduce the least squares approach to nonlinear regression, a simple one-parameter model is first considered:

$$Y_i = f(X_i,b) + \varepsilon_i = bX_{1i} + b^2 X_{2i} + \varepsilon_i, E(\varepsilon_i) = 0, Var(\varepsilon_i) = \delta^2 \tag{2}$$

Define the residual sum of squares:

$$S(b)=\sum_{i=1}^{n}\varepsilon_i^2=\sum_{i=1}^{n}[Y_i-f(X_i,b)]^2 \\ =\sum_{i=1}^{n}[Y_i-bX_{1i}-b^2X_{2i}]^2 \tag{3}$$

According to the least squares principle of regression, the sum of squares of the residuals should be minimized, and thus the derivation of $S(b)$ yields:

$$\frac{dS}{db}=2\sum_{i=1}^{n}[Y_i-f(X_i,b)]\left(-\frac{df(X_i,b)}{db}\right) \\ =2\sum_{i=1}^{n}[Y_i-bX_{1i}-b^2X_{2i}][-X_{1i}-2bX_{2i}]=0 \tag{4}$$

Organized:

$$2b^3\sum_{i=1}^{n}X_{2i}^2+3b^2\sum_{i=1}^{n}X_{1i}X_{2i}+b\left(\sum_{i=1}^{n}X_{1i}^2-2\sum_{i=1}^{n}X_{2i}Y_i\right)-\sum_{i=1}^{n}X_{1i}Y_i=0 \tag{5}$$

This is a cubic equation about b. b then has three possible solutions. Substituting each of these three solutions into $S(b)$ and taking the one that minimizes $S(b)$, $S(b)$ is the final solution to the regression model.

3.1.2. Model testing:

1) Hypothesis testing of regression equations. These partial regression coefficients $\hat{b}_j$ calculated from the sample are estimates of the overall partial regression coefficients b_j. If these overall partial regression coefficients are equal, the multiple regression equation is meaningless. Therefore, as with linear regression, it is necessary to test these partial regression coefficients after the equation has been constructed. Hypothesis testing of multiple regression equations can also be done using analysis of variance (ANOVA).

The sum of squared deviations from the mean of the dependent variable y is decomposed into two parts by regression analysis.

$$SS_T=\sum_{i=1}^{n}(y_i-\overline{y})^2=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2+\sum_{i=1}^{n}(\hat{y}_i-\overline{y})^2=SS_R+SS_E \tag{6}$$

Also, the degrees of freedom are decomposed into two parts. One of the regression degrees of freedom is the number of independent variables:

$$df_R=p,df_E=df_T-df_R=(n-1)-p=n-p-1 \tag{7}$$

From this, the mean square of the two components can be calculated separately:

$$MS_R=SS_R/df_R=\sum_{i=1}^{n}(\hat{y}_i-\overline{y})^2/p \tag{8}$$

$$MS_E=SS_E/df_E \tag{9}$$

The test hypothesis for ANOVA is $H_0:b_1=b_2=\cdots=b_p=0$, which implies that there is no regression relationship between dependent variable y and all independent variables x_j and that the multiple regression equation is not meaningful. The corresponding alternative hypothesis $H_1:b_1,b_2,\cdots,b_p$ is not all 0 and H_0 holds when there is:

$$F = \frac{MS_R}{MS_E} \sim F(p, n-p-1) \tag{10}$$

That is, F obeys a F distribution. This allows the regression equation to be tested for significance using the F statistic.

2) Hypothesis testing of regression coefficients. The fact that the multiple regression equation is statistically significant does not mean that every partial regression coefficient is significant, so it is necessary to test every partial regression coefficient. At $b_j = 0$, the partial regression coefficients $\hat{b}_j (j = 1, 2, \cdots, p)$ follow a normal distribution, so the partial regression coefficients can be tested with the t statistic. Test hypothesis $H_0 j : b_j = 0, H_{1j} : b_j \neq 0$. When H_{0j} holds, and $\hat{b} \sim N(b, \sigma^2 (X' X)^{-1})$, note $(X' X)^{-1} = (c_{ij})$. The constructed t statistic is:

$$t_j = \frac{\hat{b}_j - b_j}{s_{\hat{b}_j}} \quad j = 1, 2, \cdots, p \tag{11}$$

where $s_{\hat{\beta}_j}$ is the standard error of the j nd partial regression coefficient, the calculation of which is complicated. Then:

$$s_{\hat{b}_j} = \sqrt{c_{jj}} s_{y.x} \tag{12}$$

$$s_{y.x} = \sqrt{\frac{\sum_{i=1}^{n} \varepsilon^2}{n-p-1}} = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n-p-1}} = \sqrt{\frac{SS_E}{df_E}} = \sqrt{MS_E} \tag{13}$$

where $s_{y,x}$ is called the residual standard deviation or standard error of estimate and also reflects the degree of variation in the dependent variable y after subtracting the linear effects of the respective variables x. s_{yx} can be compared with the standard deviation s_y of y, which gives an indication of the magnitude of the linear effect of all independent variables x on y.

When the original hypothesis $H_{0j} : b_j = 0$ holds, the t statistic above obeys a t distribution with $n-p-1$ degrees of freedom. Given the level of significance α, find the critical value of the two-sided test $t_{1-\alpha/2}$. When $|t_j| \geq t_{1-\alpha/2}$, the null hypothesis $H_{0j} : b_f = 0$ is rejected, it is considered b_f significantly non-zero, and the linear effect of independent variable x_f on dependent variable y is significant. When $|t_j| < t_{1-\alpha/2}$, accept the null hypothesis $H_0 j : b_j = 0$, consider b_j to be zero and the linear effect of independent variable x_j on dependent variable y is not significant.

3.2. *Predictive modeling of working fluid parameters*

Mathematical methods are difficult to express and judge the nonlinear relationship of various complex situations in the drilling fluid leakage, thus highlighting the power of machine learning algorithms, which can classify and predict various leakage and plugging parameters in the drilling process through many types of models based on the principle of nonlinear functions. These machine learning algorithms have been shown to be effective in many complex well leakage accidents, and are of great significance in guiding the leakage prevention and plugging work during the drilling process. The built-in modules and libraries of Python are used to realize the prediction model of leakage prevention and plugging working fluid parameters, and this paper adopts the support vector machine algorithm to establish the prediction model.

Support vector machine (SVM) is a common binary classification model, and its basic idea is to map the data into a high-dimensional space and find a hyperplane in that space, so that the distance from various types of data points to the hyperplane is maximized.SVM algorithm has a wide range of applications in the fields of classification, regression and so on. Specifically, for a given training dataset, SVM will determine the best decision boundary by calculating the distance between each sample point and the hyperplane. In order to avoid overfitting and improve generalization performance, SVM also introduces a kernel function, which can map linearly indistinguishable data into a high-dimensional space, thus achieving nonlinear classification.

Functional Interval: for a given training dataset T and hyperplane (w,b), define the functional interval of hyperplane (w,b) with respect to sample point (x_i, y_i) as:

$$\hat{\gamma}_i = y_i(w \cdot x_i + b) \tag{14}$$

Define the minimum of the function interval of hyperplane (w,b) with respect to the training dataset T and hyperplane (w,b) with respect to all sample points (x_i, y_i) in T, i.e:

$$\hat{\gamma} = \min_{i=1,\cdots,N} \hat{\gamma}_i \tag{15}$$

Geometric Interval: for a given training dataset T and hyperplane (w,b), define the geometric interval of hyperplane (w,b) with respect to sample point (x_i, y_i) as:

$$\hat{\gamma}_i = y_i\left(\frac{w}{\|w\|} \cdot x_i + \frac{b}{\|b\|}\right) \tag{16}$$

Define hyperplane (w,b) with respect to the training dataset T and hyperplane (w,b) with respect to the minimum of the geometric interval between all sample points (x_i, y_i) in T, i.e:

$$\hat{\gamma} = \min_{i=1,\cdots,N} \hat{\gamma}_i \tag{17}$$

3.2.1. Linearly Differentiable Support Vector Machines. Suppose that a training dataset $T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$ on a feature space is given, where $x_i \in \chi = R^n, y_i \in Y = \{+1, -1\}, i = 1, 2, \cdots, N$. Linearly differentiable support vector machines are used to solve classification problems for linearly differentiable problems. Linearly differentiable support vector machines use interval maximization to find the optimal separating hyperplane, the so-called maximal interval separating hyperplane is to find the separating hyperplane with the largest geometric interval.

Given a linearly divisible training dataset, the separation hyperplane obtained by interval maximization or equivalently by solving the corresponding convex quadratic programming problem is learned as:

$$w^* \cdot x + b^* = 0 \tag{18}$$

And the corresponding classification decision function is:

$$f(x) = sign(w^* \cdot x + b^*) \tag{19}$$

Solving a maximally spaced hyperplane can be transformed into an optimization problem, and the linearly divisible support vector machine learning algorithm is given below: let the linearly divisible training set $T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$, where:

$$x_i \in \chi = R^n, y_i \in Y = \{+1, -1\}, i = 1, 2, \cdots, N \tag{20}$$

1) Construct and solve the constrained optimization problem:

$$\begin{aligned}&\min_{\alpha} \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^{N}\alpha_i \\ &s.t. \quad \sum_{i=1}^{N}\alpha_i\ y_i = 0 \\ &\alpha_i \geq 0, i = 1,2,\cdots,N\end{aligned} \tag{21}$$

Find the optimal solution $\alpha^* = (\alpha_1^*, \alpha_2^*, \cdots, \alpha_N^*)^T$.

2) Compute $w^* = \sum_{i}^{N}\alpha_i^* y_i x_i$ and choose a positive component $\alpha_j^* > 0$ of α^* and compute $b^* = y_j - \sum_{i=1}^{N}\alpha_i^* y_i (x_i \cdot x_j)$.

3) Find the separating hyperplane (18) and the classification decision function (19).

3.2.2. Linear support vector machines. Suppose that dataset $T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$ is defined on the feature space where $x_i \in \chi = R^n, y_i \in Y = \{+1, -1\}, i = 1,2,\cdots,N$. When the training dataset is not divisible, it means that some sample points (x_i, y_i) cannot satisfy the constraint that the function interval is greater than or equal to 1. Whereas the datasets encountered in reality are indivisible, to overcome this problem, a slack variable $\xi_i \geq 0$ can be introduced for each sample point (x_i, y_i) such that the function interval plus the slack variable is greater than or equal to 1. Thus the constraint becomes:

Sufficient dimensionality reduction based on principal support vector machine:

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i \tag{22}$$

The objective function becomes:

$$\frac{1}{2}\| w \|^2 + C\sum_{i=1}^{N}\xi_i \tag{23}$$

where $C > 0$ is called the penalty parameter. The learning problem of a linearly indivisible linear support vector machine thus becomes the following convex quadratic programming problem:

$$\begin{aligned}&\min_{\alpha} \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^{N}\alpha_i \\ &s.t. \quad \sum_{i=1}^{N}\alpha_i\ y_i = 0 \\ &0 \leq \alpha_i \leq C, i = 1,2,\cdots,N\end{aligned} \tag{24}$$

4. Experimental results and analysis

4.1. Factors affecting the pressure-bearing capacity of the working fluid

4.1.1. Pre-laboratory preparation:

1) Setting of influencing factors. The influencing factors of the pressure-bearing capacity of the working fluid for leakage prevention and plugging are selected from three aspects: stratigraphic factors, fluid factors and process factors. Stratigraphic factors: stratum lithology X, stratum fracture S, stratum pressure p, stratum temperature T. Fluid factors: density D and viscosity AV. process factors: injection speed Q.

2) Experimental core. For sandstone, the core plunger is directly drilled by rotary grinding with a

plunger driller, and then a single penetrating seam parallel to the axial direction of the plunger is obtained after applying a certain pressure to the core plunger, which is used to simulate the fracturing process, and then artificial ceramic grains of different grain sizes are used to support to achieve the purpose of different seam widths. For carbonate rock, the core taken out during drilling has developed cracks, and if the standard plunger is disturbed too much during drilling, the cracks are easy to be enlarged, and the core is easy to be ruptured and fall out, which can not accurately simulate the underground fracture state. To solve this problem, the core plungers were prepared using CNC wireline cutting technology, which retained the natural cracks in the core well, so that each core plunger had a single natural crack parallel to the axial direction. In order to be closer to the real situation of the formation after acidizing during well workover, especially the fracture surface, the prepared carbonate core plungers were subjected to indoor simulation of acidizing, and the fracture after acidizing was processed to obtain the core plungers with different fracture widths.

3) Experimental equipment. The main equipment used in this experiment is the unconventional natural gas production simulation system. The equipment mainly includes core holder, sealing and heating box, fluid pump, pressure gauge and other components.

4.1.2. Experimental data collection. An indoor test was conducted to evaluate the influence of different factors on the pressure-bearing capacity of the plunger leakage prevention and plugging working fluid in fractured rock samples. The experimental data are not measured with the same kind of instruments, standardized on the same scale and operated in the same way, so there is bound to be an error in each set of data mainly due to the scale factor. In order to eliminate such errors as much as possible, each type of data was standardized and each factor was corrected to a level. The Z-SCORE method was used to standardize the field data.

4.1.3. Analysis of model results. Organize the indoor experimental data of the pressure-bearing capacity of leak-proof and leak-plugging working fluid, establish the mathematical model of the pressure-bearing capacity of leak-proof and leak-plugging working fluid through the method of multivariate nonlinear regression analysis, and analyze it to get the pressure-bearing capacity (P) influencing factors and laws.

SPSS software was used to open the normalized data and carry out multivariate nonlinear regression to obtain the data model equation as follows: $P=-0.189X-0.694S+0.031T+0.502p-0.005D+0.099AV-0.181Q+0.562$.

Sig=0.008<0.05, indicating that the model is applicable. In order to verify the accuracy of the mathematical model, the experimental data were fitted to the experimental data of the pressure-bearing capacity data, and the comparison of the fitted value of the pressure-bearing capacity with the actual value is shown in Figure 1. The fitted curve of pressure capacity of leakage prevention and plugging working fluid is similar to the actual curve of experiment, and the correlation coefficient of model fitting is $R^2=0.922$, which is good, so it can be used to analyze the influencing factors and laws of pressure capacity of working fluid for leakage prevention and plugging in fractured stratum.

It can be seen from the equations of the mathematical model of pressure-bearing capacity and the fitting results:

1) The coefficients of formation temperature (0.031), formation pressure (0.502) and apparent viscosity (0.099) are positive, so the pressure-bearing capacity of working fluid for leakage prevention and plugging is positively correlated with the formation temperature, formation pressure and apparent viscosity, and the coefficients of formation lithology (-0.189), seam width (-0.694), density (-0.005) and injection speed (-0.181) are negative. Therefore, the pressure-bearing capacity of leakage prevention and plugging working fluid is negatively correlated with the formation lithology, seam width, density and injection speed.

2) Among the stratum factors, the absolute value of the coefficient of seam width is the largest, so the

seam width has the largest influence on the pressure-bearing capacity, and the absolute value of the coefficient of stratum pressure and stratum lithology is second only to the seam width, so the stratum pressure and lithology also have a great influence on the pressure-bearing capacity, and the temperature has a very small influence on the pressure-bearing capacity.

3) Among the fluid factors, the larger the apparent viscosity is, the higher the pressure-bearing capacity of the leak prevention and plugging working fluid is, because the coefficient of density is very small, and the coefficient is close to 0, so the density has a small influence on the pressure-bearing capacity, and the coefficient of apparent viscosity is medium, so the apparent viscosity has an influence on the pressure-bearing capacity.

4) Among the process factors, the larger the injection speed is, the smaller the pressure-bearing capacity of the working fluid for leak prevention and plugging is, and the absolute value of the coefficient is medium, so the injection speed has a certain effect on the pressure-bearing capacity.

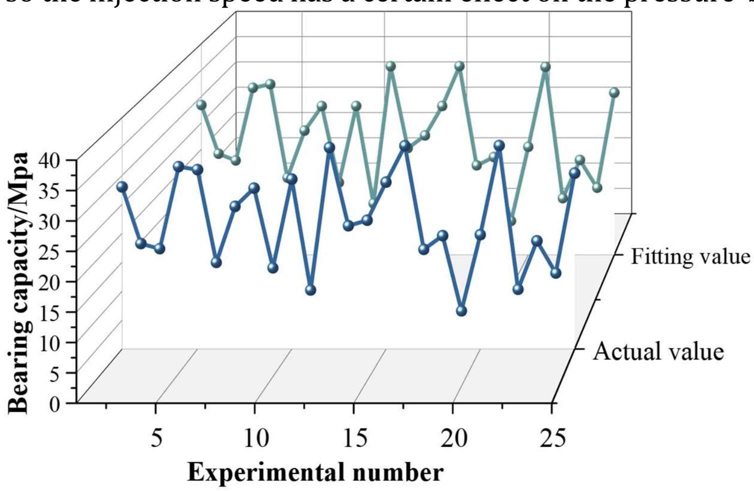


Fig. 1. Comparison of pressure-bearing capacity fitting value and actual value

The relative errors between the actual and fitted values of pressure capacity are shown in Figure 2. The relative errors between the fitted and actual values of pressure capacity range from 0.3% to 5.5%, with an average relative error of only 2.9%.

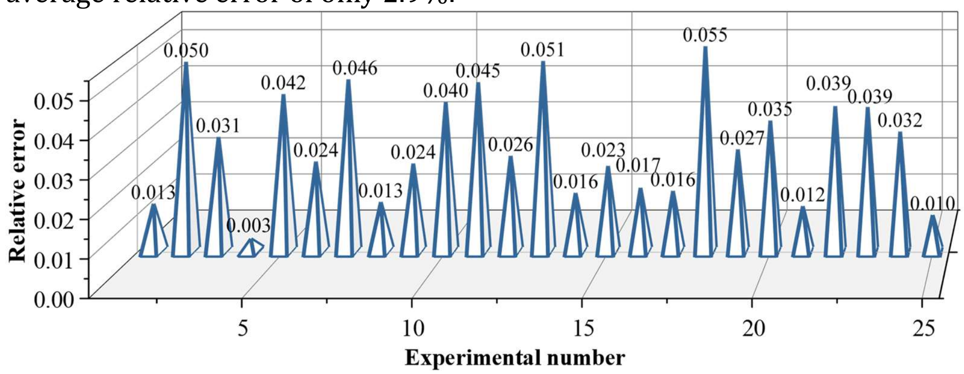


Fig. 2. Relative error between actual value and fitting value of pressure-bearing capacity

Figure 3 shows the analysis of the residuals from the fitting of the pressure capacity.The basal distribution of the data points of the P-P plot is near the fitted line. The calculated residual of 0.028 indicates that the regression model can fit the original data better without singularities. Therefore the mathematical model based on multivariate nonlinear regression analysis is suitable for the analysis of the factors and laws affecting the pressure-bearing capacity of working fluid for leakage prevention and plugging in fractured formations.

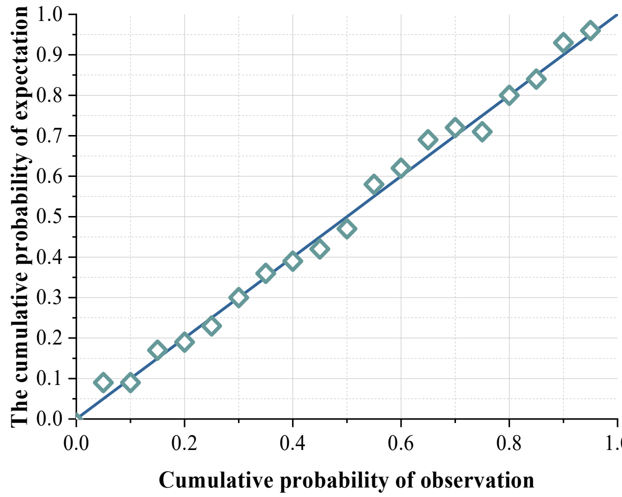


Fig. 3. Fitted residual analysis P-P plot of pressure-bearing capacity

4.2. *Predictive analysis of working fluid parameters*

For the prediction model based on support vector machine, the relationship between input and output is calculated according to the model classification by using the data such as leakage parameter given in the training set to make an accurate classification prediction between the working fluid parameter and the output.

4.2.1. Data collection and cleansing. At present, six major categories of data, including well leakage, drilling working fluid, logging curve, logging interpretation and stratification and formation pressure data, are collected and organized for 115 wells in an oil mine in a certain area to construct the experimental dataset. There are many missing data in the logging interpretation results, and from the data, it is mainly because some well leakage sections are in non-reservoir sections, so the interpretation results do not contain data such as pore seepage saturation in the well leakage sections.

Characterization data are processed as follows:

1) Firstly, empty rows and columns are removed when reading the file.

2) For the logging curve and interpretation results data, the anomalies are also processed as empty values, which is convenient for the subsequent unified processing.

3) Round up the well depths and keep the whole number before merging, otherwise there will be a lot of missing data because of the difference after the decimal point.

4) If the logging curve and the interpretation results have renamed columns when merging, the logging curve column will be removed and retained due to more missing interpretation results.

5) In addition to the numerical items, there are text items such as directional, 91+screw, screw+58 and so on, which are uniformly processed as the average value.

6) For leakage related data such as whether leakage, leakage speed, grain size, layer position, pump pressure, displacement, etc., no leakage is added as 0.

7) The format of leakage speed column is not uniform, and the regular expression is used to extract the value.

8) Pump pressure and displacement columns also use regular expressions to extract the numerical items.

9) Outliers in the data are removed.

10) The layer data are strings, which cannot be recognized by the program, and are replaced by numbers.

11) Logging data are filled with interpolation method.

12) Layer position, layer top depth, pore pressure, rupture pressure, collapse pressure, and closure

pressure are populated with priors.
13) The remaining characteristics are populated with median values.

4.2.2. Division of raw data. When training a well leakage prediction model, the processed well leakage data is usually divided into three parts, i.e., the training set, the test set, and the validation set, in order to evaluate the model's performance and generalization ability. When performing model training and evaluation, it is necessary to ensure that there is no overlap between the training, validation and test sets, so as not to lead to model overfitting that has an impact on the accuracy of the training results. Eight wells were randomly taken out from the block as the validation set, and the remaining data were divided into the training and test sets in the ratio of 8:2.

4.2.3. Forecast analysis. Particle size distribution is an important parameter in the process of designing bridging plugging formulations, which are selected on the basis of particle size distribution for plugging operations to minimize fluid loss. The design of bridging particle size is crucial if the material is to enter the crack smoothly and form a solid bridging.

In this study, we address the problem of predicting the particle size of D90, the largest particle of the leakage prevention and plugging working fluid, whose particle size distribution range is shown in Fig. 4, which is categorized into four classes (0, 500], (500, 1500], (1500, 3000], and (3000, +∞), (unit: μm). The training set was inputted into different classification models for training, and model tuning, etc., to derive the optimal model and parameters.The D90 particle size distribution ranges were mostly concentrated in the two ranges of (500, 1500] and (1500, 3000], (0, 500] was the second largest, and the distribution was the least in (3000, +∞].

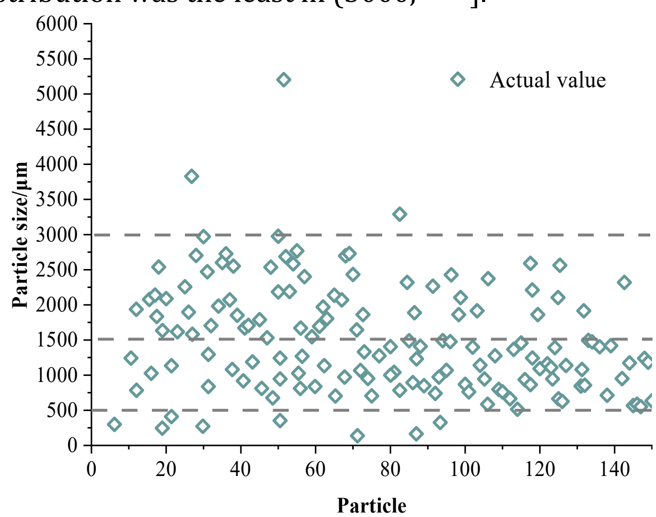
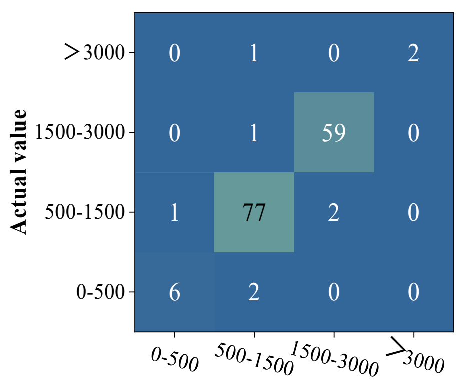
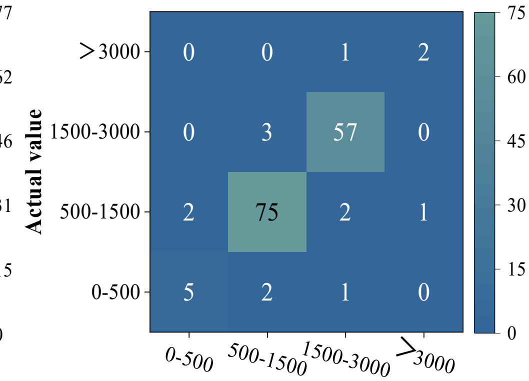


Fig. 4. Plugging formula D90 particle size distribution range

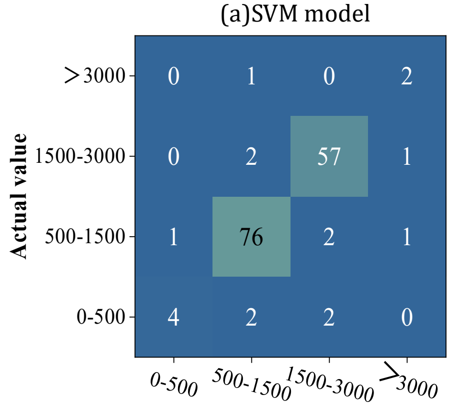
The k-nearest neighbor, neural network, GA-SVC, decision tree, random forest, extreme random tree, stochastic gradient descent, AdaBoost and Bagging meta-estimation are selected as the comparison methods, and the confusion matrix of the validation set for grain size classification is shown in Fig. 5, and (a) to (j) are the corresponding prediction results of each algorithm. Most of the particle sizes in the validation set are in the range of 500 to 1500 μm as well as 1500 to 3000 μm. In terms of prediction, the support vector machine algorithm correctly predicted 77 cases in the range of 500 to 1500 μm, 59 cases in the range of 1500 to 3000 μm were all correctly predicted, the support vector machine algorithm also correctly predicted the most of the 6 samples in the range of 0 to 500 μm, and also correctly predicted 2 out of the 3 cases in the range of >3000 μm. Comprehensively the support vector machine algorithm performed the best, with only 7 prediction errors within 151 points, and the overall prediction accuracy of the particle size of the leakage prevention and plugging working fluid was 95.36%.



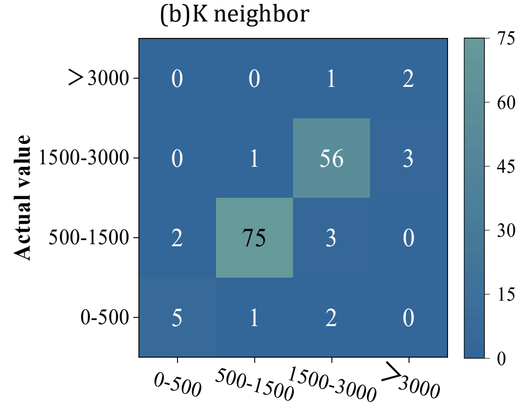
(a)SVM model



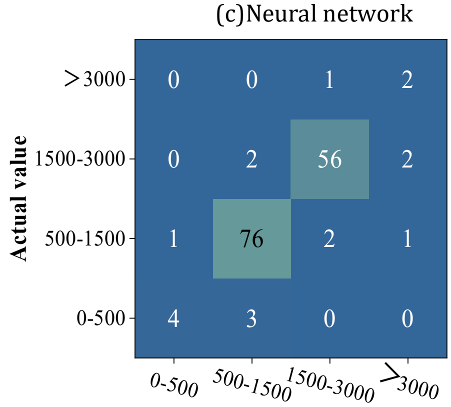
(b)K neighbor



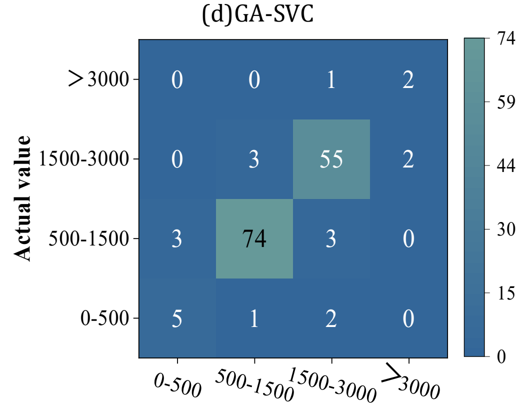
(c)Neural network



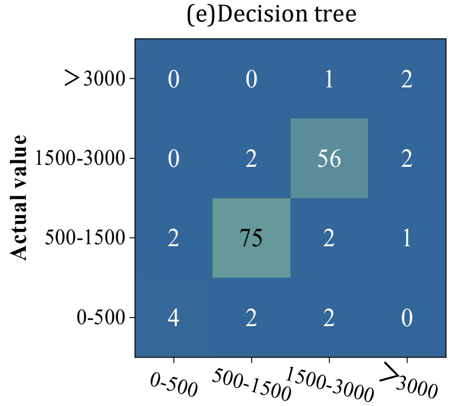
(d)GA-SVC



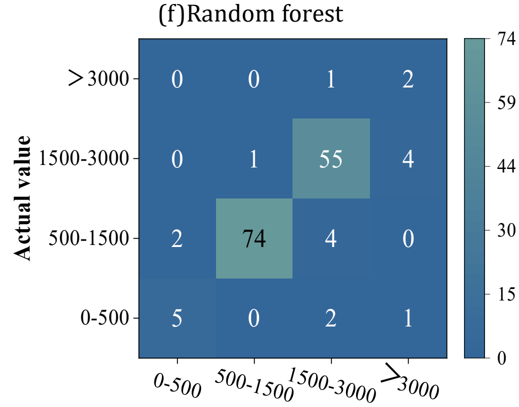
(e)Decision tree



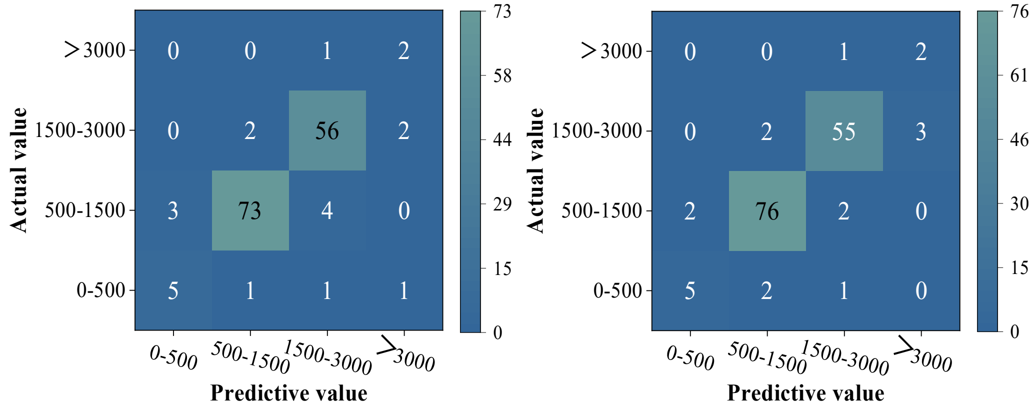
(f)Random forest



(g)Extreme random tree



(h)Random gradient descent



(i)AdaBoost (j)Bagging meta-estimate

Fig. 5. Validation set confusion matrix (particle size classification)

The F1 scores of each model data set are shown in Figure 6. It can be seen that the stochastic support vector machine prediction model has a better performance in the training set, test set and validation set, with scores of 0.9173, 0.9081 and 0.9320, respectively, i.e., this paper's prediction model for the parameters of the leakage prevention and plugging working fluids based on the support vector machine has a high prediction performance.

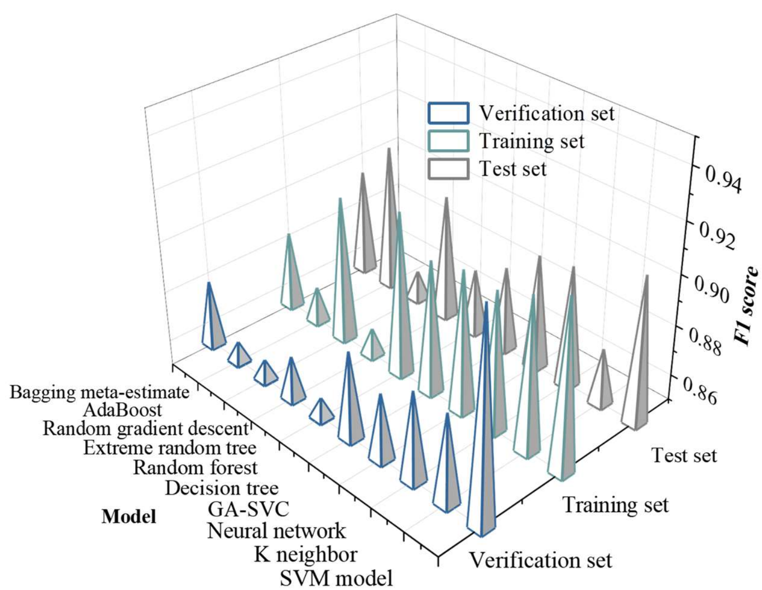


Fig. 6. F1 score for each model dataset (extended training set)

5. Conclusion

Leakage is one of the common downhole complex problems in the drilling process, and there is an urgent need to carry out an in-depth study on the materials of ultra-high-temperature dense pressure-bearing leakage prevention and plugging working fluids and other materials. In this study, we select the factors affecting the pressure-bearing capacity of leakage prevention and plugging working fluids, establish a mathematical model of pressure-bearing capacity through multivariate nonlinear regression analysis, and establish a prediction model for the parameters of leakage prevention and plugging working fluids based on the support vector machine by using machine learning algorithms, and then analyze and test the two mathematical models. The main results are as follows:

1) The established mathematical model of pressure-bearing capacity of leakage prevention and plugging working fluid in fractured formations has a fitting correlation coefficient of R^2=0.922, which

is a good fit. The model shows that the factors that have the greatest influence on the pressure-bearing capacity of the leakproof plugging working fluid are seam width (-0.694) and formation pressure (0.502). The model has high accuracy and small fitting error, and the average relative error is only 2.9%, which can be applied to analyze the influencing factors and rules of the pressure capacity of working fluid for leakage control and plugging in fractured formations.

2) Most of the D90 particle sizes of leakproof plugging working fluid are distributed in (500, 1500] and (1500, 3000]. For the prediction of the maximum particle size of D90 of the leakage plugging material, the support vector machine model in this paper has the best performance, and its F1 value for the training set, test set and validation set is above 0.9, and the prediction accuracy in the validation set is 95.36%.

Based on the above study, the model of this paper can provide certain reference for the design of leakage prevention and plugging scheme and construction parameters in the field, which can be used to guide the temporary plugging and repairing operation of the leakage prevention and plugging working fluid in fractured formations, which is of great significance in realizing the long-term and stable plugging in the field.

About the Author

Siqi Chen works at Chongqing City Management College and is engaged in drilling fluid technology research.

References

[1] Yang, J., Sun, J., Bai, Y., Lv, K., Zhang, G., & Li, Y. (2022). Status and prospect of drilling fluid loss and lost circulation control technology in fractured formation. Gels, 8(5), 260.

[2] Kang, Y., Xu, C., You, L., Yu, H., & Zhang, B. (2014). Comprehensive evaluation of formation damage induced by working fluid loss in fractured tight gas reservoir. Journal of Natural Gas Science and Engineering, 18, 353-359.

[3] Lichun, J. I. A., Mian, C. H. E. N., Bing, H. O. U., Zhen, S., & Yan, J. (2014). Drilling fluid loss model and loss dynamic behavior in fractured formations. Petroleum exploration and development, 41(1), 105-112.

[4] Zhang, Z., Wei, Y., Xiong, Y., Peng, G., Wang, G., Lu, J., ... & Wang, J. (2022). Influence of the location of drilling fluid loss on wellbore temperature distribution during drilling. Energy, 244, 123031.

[5] Zhu, X., Cheng, F., Shi, C., & Chen, K. (2020). Mechanical plugging—solid expandable tubular refracturing technology. Journal of Mechanical Science and Technology, 34, 2357-2364.

[6] Jinsheng, S. U. N., Yingrui, B. A. I., Cheng, R., Kaihe, L. Y. U., Fan, L. I. U., Jie, F. E. N. G., ... & Huijun, H. A. O. (2021). Research progress and prospect of plugging technologies for fractured formation with severe lost circulation. Petroleum Exploration and Development, 48(3), 732-743.

[7] Wang, X., Liao, L., Bai, R., Liu, F., Zheng, W., & Wang, S. (2022). Study on Plugging Technology in Oil Test. Open Journal of Yangtze Oil and Gas, 7(4), 240-252.

[8] Hu, A., Li, K., Feng, Y., Fu, H., & Zhong, Y. (2022). Laboratory Evaluation and Field Application of a Gas-Soluble Plugging Agent: Development of Bottom Water Plugging Fracturing Technology. Energies, 15(18), 6761

[9] Liu, J., Kou, X., Liang, X., Gao, X., Ye, Z., & Jin, Y. (2024, October). Research Progress onLeak Prevention and Plugging Technology in Deepwater Drilling. In Journal of Physics: Conference Series (Vol. 2834, No. 1, p. 012025). IOP Publishing.

[10] Chen, X., Lu, X., Liu, P., Du, J., Liang, C., Huang, Q., ... & Liu, F. (2024). A critical review of key points in

temporary plugging fracturing: Materials, injection, temporary plugging, and design. Geoenergy Science and Engineering, 240, 212981.

[11] Deng, B., Jia, Z., Liu, W., Liu, X., Gu, J., Ye, W., ... & Qu, Z. (2021). Selective water plugging technology for horizontal well with screen completion. Energies, 14(4), 791.

[12] Pan, Y., Cui, X., Wang, H., Lou, X., Yang, S., & Oluwabusuyi, F. F. (2023). Research Progress of Intelligent Polymer Plugging Materials. Molecules, 28(7), 2975.

[13] Xiangwei, K., Xuanji, E., Tianjun, Q., Qing, C., Yong, R., Subing, W., ... & Yu, L. (2023). Study on Mathematical Model and Influencing Factors of Composite Temporary Plugging Pump Pressure in Shale Gas Wells. Special Oil & Gas Reservoirs, 30(4), 156.

[14] Rong, W., Lishan, Y., Yao, L., Xuda, Y., Lü, B., & Jiaqi, C. (2022). High Temperature Plugging Mechanism and Experimental Evaluation of Temporary Plugging Agent. Journal of Petrochemical Universities, 35(2), 62.

[15] Bai, Y., Liu, Y., Sun, J., & Lv, K. (2024). Plugging Mechanism of Rigid and Flexible Composite Plugging Materials for Millimeter-Scale Fractures. SPE Journal, 29(04), 1786-1801.

[16] Xu, B., Wang, X., Liu, X. B., Wang, Q. G., & Chen, W. J. (2021). New and Successful Field Application of Water Plugging in Thick Oil Layer Using Positioning Balanced Fracturing Technology. In Proceedings of the International Field Exploration and Development Conference 2020 (pp. 2843-2863). Springer Singapore.

[17] Zhongchao, L., Jiang, S., Xuyang, Z., Xiaohui, H., & Ming, H. (2021, August). Study on Cross Layer Plugging Technology for Injection Well. In International Petroleum and Petrochemical Technology Conference (pp. 400-408). Singapore: Springer Singapore.

[18] Wu, H., Huang, K., Liu, J., Zhang, B., Liu, J., Kong, C., ... & Li, Z. (2021, March). Research and Application of Compound Plugging Removal Technology in HTHP Condensate Gas Well. In International Petroleum Technology Conference (p. D012S045R095). IPTC.

[19] Zou, Y., Luo, C., Wang, B., Zhang, H., & Zhang, C. (2021). Development and Application of TL-1 New High Temperature and High Pressure Plugging Instrument. Open Journal of Yangtze Oil and Gas, 6(3), 107-115

[20] Li, Z., Zhou, Y., Qu, L., & Zhang, X. (2023). Smart plugging and low-damage foam drilling fluid technology for CBM drilling. International Journal of Oil, Gas and Coal Technology, 32(2), 103-125.

[21] Li, S., Niu, C., Zhang, Y., Ge, J., & He, W. (2022). Application of flexible colloid-fiber compound plugging technology in oil and gas wells with large differential pressure in the East China Sea. In MATEC Web of Conferences (Vol. 356, p. 01011). EDP Sciences.

[22] Qing, C., Zhang, X., Liu, C., Liu, Q., & Gong, H. (2024, October). Research and Application of Plugging Drilling and Completion Technology for Difficult Old Well in Changchun Gas Storage. In Journal of Physics: Conference Series (Vol. 2834, No. 1, p. 012074). IOP Publishing.

[23] Deng, J., Zhang, A., Ji, Y., Sun, X., & Su, J. (2022). Research on Gel Plugging Technology of Coalbed Methanein Deposit. Chemistry and Technology of Fuels and Oils, 58(5), 895-901.