ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

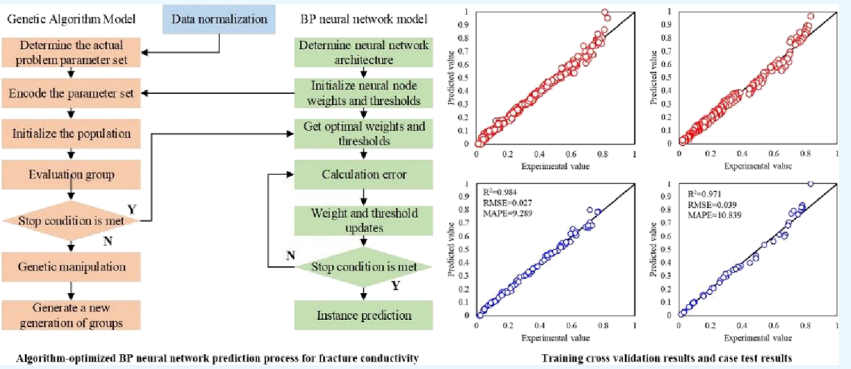
Fracture Conductivity Prediction Based on Machine Learning

Xiaopeng Wang, Binqi Zhang, Jianbo Du, Dongdong Liu, Qilong Zhang, and Xiaoqiang Liu*

Cite This: *ACS Omega* 2024, 9, 13469−13480

Read Online

ACCESS | Metrics & More | Article Recommendations



Algorithm-optimized BP neural network prediction process for fracture conductivity

Training cross validation results and case test results

ABSTRACT: Hydraulic fracturing technology is the main method to develop low-permeability reservoirs. Fracture conductivity is not only the basis of fracture optimization design but also one of the key parameters to determine the effect of hydraulic fracturing. However, current methods of calculating fracture conductivity require a lot of time and labor cost. This research proposes a fracture conductivity prediction model based on machine learning. The main controlling factors of fracture conductivity are determined using the Pearson coefficient method and gray correlation analysis. Example application shows that the R^2 values of the BP neural network model based on a genetic algorithm for predicting the fracture conductivity of block A and block B are 0.981 and 0.975, respectively, indicating that the machine learning model can accurately predict fracture conductivity.

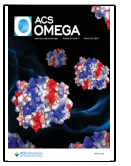
1. INTRODUCTION

Offshore low-permeability oilfields are an important block for increasing oil and gas production. Hydraulic fracturing technology is the main method to develop offshore low permeability reservoirs.[1−3] However, the cost of hydraulic fracturing for an offshore well is much higher than that of a terrestrial well. It is necessary to design better fracturing schemes to improve the fracturing effect, reduce the cost, and increase the benefit.[4,5]

Hydraulic fracture quality is the core of hydraulic fracturing quality, and fracture conductivity is particularly important.[6−8] The prediction of fracture conductivity is not only the basis of fracture optimization design but also one of the key parameters to determine the effect of hydraulic fracturing.[9−11] At present, the main research method for fracture conductivity is based on indoor testing methods. Xiao et al.[12] took carbonate rock as the research object and experimentally studied the effects of acid erosion, proppant embedding depth, and different proppant sizes and acid injection on fracture conductivity during hydraulic fracturing and acid fracturing; Guo et al.[13] conducted tests on steel plates and shale to understand the effects of flowback rate, fracturing fluid, and closure stress on proppant flowback and fracture conductivity; Sun et al.[14] used ceramic and quartz sand proppant obtained from Changqing Oilfield in China. For example, the fracture conductivity under different proppant mixing ratios was evaluated through experiments; Tariq et al.[15] studied the effects of different acid fracturing fluids, rock hardness, and surface roughness on the fracture conductivity of carbonate rocks through experiments. The conductivity data obtained through experiments are the most direct and accurate, but this method requires on-site coring and repeated experiments for different oil wells, and its cost consumption is very huge.

In response to the limitations of experimental methods, some scholars have proposed theoretical models for predicting fracture conductivity. Zhou et al.[16] based on the classic Biot poroelasticity theory to model porous media and capture fracture behavior through a phase field model and proposed a poroelastic medium fracture phase field model to calculate fracture conductivity. Rabczuk and Belytschko[17] proposed a new method for treating crack growth by particle methods to

Received: January 14, 2024
Revised: February 19, 2024
Accepted: February 21, 2024
Published: March 8, 2024



calculate fracture conductivity. Zhang et al.[18] analyzed the factors that may influence shale fracture conductivity, and a theoretical model of propped-fracture conductivity was established considering the effect of water damage on fracture conductivity. Based on Hertz theory of elastic contacts, Jia et al.[19] proposed mathematical models to calculate fracture conductivity, reduction in fracture aperture, proppant embedment, and deformation of rod-shaped proppants and further established models for calculating the porosity of the fracture with proppants in loose and close packing modes. Xu et al.[20] constructed the 3D tortuous fracture combining the normal distribution random function and the quartet structure generation set. This method included three main steps of fracture seeding, fracture growing, and geometric loft. The normal distribution random function was used in the seeding process to control the orientation and tortuosity of fractures, and the flowing pressure and velocity of fluid in fractures were calculated by the coupling model of elasticity and the lattice Boltzmann method. Zhang et al.[21] first used DEM to model the mechanical interaction between proppant pack and shale formation during fracture closing. Fracture conductivity after fracture closing was further calculated by DEM coupled with CFD. Su et al.[22] established a model for calculating fracture conductivity considering proppant embedment and fracture shape, and the influence of key parameters on fracture conductivity was analyzed. In general, numerical methods have some limitations: ① Numerical methods usually need to be calculated based on some simplified assumptions, considering only the fracture conductivity under specific conditions. These assumptions may not fully reflect the actual situation. ② Results of numerical models: the quality usually depends on the accuracy of the model parameters. The distortion of the parameters may lead to deviations in the calculation results, and some parameters are often difficult to obtain on site. ③ Numerical models usually need to consider fluid−structure interaction calculations, and their mathematical models are very complex, which requires a lot of computational resources and time, especially for complex geological structures, greatly limiting the feasibility of the model in practical engineering applications.

Some scholars carry out regression fitting based on limited experimental test data. Shi et al.[23] deduced fracture conductivity based on the Kozeny formula and the theory of elasticity considering the influence of proppant strength, particle size, sand concentration, closing pressure, proppant embedding, crushing, and proppant and fracture wall deformation. Kainer et al.[24] employed multiple linear regression to generate empirical correlations based on experimental data for different shale types. However, the fitting effect was poor. Zhu et al.[25] put forward a new method for testing shale branch fracture conductivity, and branching fracture conductivity test was carried out. Chen et al.[26] conducted conductivity tests on actual coal rock fractures to assess the effect of various particle size ratios on the conductivity of complex fractures in CBM reservoir and identified an optimized proppant blending approach that was suitable for hydraulic fracturing in coal seams.

With the development of artificial intelligence technology, machine learning technology has shown increasingly powerful performance in image processing, parameter prediction, etc. At present, with the production and development of oil fields, a large amount of conductivity test data have been accumulated, which provides a strong data foundation for the application of machine learning methods. At present, researchers have made preliminary explorations using artificial neural networks (ANN) to predict fracture conductivity. In 2021, Desouky et al.[27] collected approximately 350 data points from experiments in several important shale formations (Marcellus, Barnett, Fayetteville, and Eagle Ford) and used machine learning methods (ANN) to predict fracture conductivity for the first time. Preliminary exploratory studies have been conducted; however, up to now, no complete study on the application of machine learning methods for full-process prediction of fracture conductivity has been seen. Research in this field is still in its preliminary stages and requires in-depth excavation and systematic research to establish a complete set of methods to effectively predict the fracture conductivity.

There are many factors affecting fracture conductivity, and there is a strong nonlinear relationship between them. In order to take into account both calculation efficiency and prediction accuracy, this paper chooses the classical BP neural network model with a strong nonlinear fitting ability to predict fracture conductivity. The BP neural network has the following advantages:[28]

① BPNN has a good ability of self-learning and self-adaptation. Through the continuous learning of the sample data, the network can train the data, summarize the characteristics of the sample data, extend the trained BP neural network to the untrained data (that is, test data), and use the generalization ability of the network to realize the mapping from input to output so as to achieve the effect of prediction.

② BPNN has a good nonlinear mapping ability. A three-layer BP neural network can realize a complex nonlinear relationship within the system after training. In this process, it does not require a regular requirement for the input variable data itself, nor does it need a good understanding of the network processing mechanism, but only the processing of the input data can obtain the corresponding mapping.

Although the BP neural network has good self-learning and adaptive ability and can realize nonlinear mapping between variables, it also has some shortcomings:

① The convergence rate is slow. The traditional BP neural network needs a relatively constant learning rate and a low range to achieve stable learning of the network, so the training generalization process converges slowly.

② It is easy to fall into the local minimum in the process of training. The weights and thresholds of the BP neural network are randomly unstable, while the BP neural network depends on the weights and thresholds, which affect the prediction accuracy of the network to some extent. At the same time, although the BP neural network can realize the nonlinear mapping between variables, the error surface of nonlinear network is much more complex than that of the linear network, so there will be many local optimal functions, which easily fall into the local optimal solution.

Therefore, it is necessary to introduce an optimization algorithm to improve the performance of the BPNN model. In this research, the main controlling factors of fracture conductivity are determined by Pearson coefficient method and gray correlation analysis (PCM-GCA). A prediction model of fracture conductivity is established by using a BP neural network optimized by a genetic algorithm. Compared with the

https://doi.org/10.1021/acsomega.4c00448
ACS Omega 2024, 9, 13469−13480

theoretical model and empirical formula, the machine learning model does not need complex mathematical formulas and realizes a fast, cheap, and accurate method to predict fracture conductivity. This method is easy to apply on site, can effectively guide on-site production, is of great significance for improving oilfield development efficiency and economic benefits, and has a strong engineering application value.

2. FRACTURE CONDUCTIVITY PREDICTION MODEL

2.1. BP Neural Network Algorithm. The BP neural network is a self-learning nonlinear fitting modeling method, which can automatically adapt and determine the connection weight of each neuron according to the input training samples. After many times of training through the neural network system, the weights of each layer of the neural network will store fitting information, which is extracted from the sample data set. Finally, the desired predicted value can be obtained through the operation of input data and weights.[29] The calculation process of the BP neural network model includes information forward propagation and error back-propagation.

2.1.1. Forward Propagation. The training data are provided to the neural network and transferred from the input and hidden layer to the output layer, which is a forward propagation process. There are n, q, and m nodes in the input layer, hidden layer, and output layer, respectively. The weight between the input layer and the hidden layer is v_{ik}, and that between the output layer of the hidden layer is ω_{kj}. The output of hidden layer node and output layer node are expressed as the following, respectively:

$$y_k = f_1\left(\sum_{i=1}^{n} v_{ik}x_i\right) k = 1, 2, L, q \tag{1}$$

$$o_j = f_2\left(\sum_{k=1}^{q} \omega_{kj}y_k\right) \tag{2}$$

where yk and o_j are the hidden layer node and output layer node output, respectively. f_1 and f_2 are the activation functions. v_{ik} is the weight between the input layer and hidden layer. ω_{kj} is the weight between the hidden layer and output layer. x_i is the ith input parameter. n and q are the numbers of input layer and hidden layer nodes, respectively.

2.1.2. Back-Propagation. After completing a forward propagation, if the output result of the output layer does not match the expected value, that is, when the error range exceeds the limit value, the BP neural network enters the process of back-propagation. In the back-propagation process, the error is passed layer by layer as an adjustment signal and the weight and threshold of each layer are continuously adjusted to reduce the error value. After repeated learning, the error is reduced to an acceptable level.

P is input to learn samples (x_1, x_2, ..., and x_p), and t_j^p is the desired output. After the Pth sample is input to the network, the output y_j^p is obtained. According to the square error function, the error E_p of the P^{th} sample can be solved.

$$E_p = \frac{1}{2}\sum_{j=1}^{m}(t_j^p - y_j^p)^2 \tag{3}$$

The cumulative error BP algorithm is used to adjust ω_{jk} and make the global error E smaller, which can be expressed as follows:

$$\omega_{jk} = -\eta\frac{\partial E}{\partial \omega_{jk}} = -\eta\frac{\partial\left(\sum_{p=1}^{p} E_p\right)}{\partial \omega_{jk}} = \sum_{p=1}^{p}\left(-\eta\frac{\partial E}{\partial \omega_{jk}}\right) \tag{4}$$

where η is the learning rate.

Therefore, the weight adjustment formula of each neuron in the output layer is expressed in eq 5. Similarly, the weight adjustment formula of each neuron in the hidden layer is expressed in eq 6.

$$\omega_{jk} = \sum_{p=1}^{p}\sum_{j=1}^{m}\eta(t_j^p - y_j^p)f_2(S_j)z_k \tag{5}$$

$$v_{ki} = \sum_{p=1}^{p}\sum_{j=1}^{m}\eta(t_j^p - y_j^p)f_2(S_j)\omega_{jk}f_1(S_k)x_i \tag{6}$$

A typical three-layer BP neural network architecture consists of an input layer, a hidden layer, and an output layer, as shown in Figure 1.

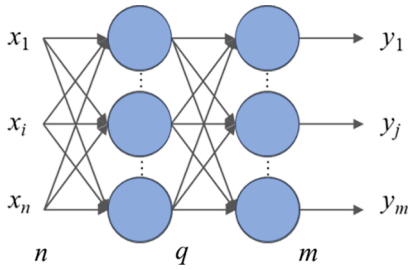


Figure 1. BP neural network structure diagram.

2.2. Genetic Algorithm. The initialization of the network weight and threshold determines which point on the error plane the network training starts from. Therefore, optimizing the initial weight and threshold of the BP neural network through genetic algorithms is crucial to obtaining the training results of the BP neural network.

The genetic algorithm is an adaptive heuristic global search algorithm. By simulation of the evolutionary mechanism of natural organisms, it uses selection, crossover, and mutation operations to search for the optimal solution in the problem space and perform global optimization. In the genetic algorithm, the crossover operator is used as the main operator because of its global search ability. The mutation operator is used as the auxiliary operator because of its local search ability.[30] Compared with the traditional optimization algorithm, the genetic algorithm has the following advantages:

① It has the characteristics of simplicity and maneuverability. The processing object of the genetic algorithm is not the parameters themselves but to encode the parameters to get gene individuals, and then there are genes to form chromosomes for genetic operation, which is operable.

② It has the characteristics of a group and global search. The advantage of the genetic algorithm is that it can simultaneously carry out a global search for multiple individuals in the population and evaluate multiple search results in the global space, so as to select a better high-quality solution and effectively avoid falling into local optimization.

2.2.1. Genetic Algorithm Parameter Initialization. The process of genetic algorithm parameter initialization is as follows: First, determine the parameter set of actual problem,

https://doi.org/10.1021/acsomega.4c00448
ACS Omega 2024, 9, 13469−13480

encode the parameter set, and determine the chromosome length and the range of coding values. The chromosome length (L) is the total number of weight and threshold of the BP neural network, which is determined by eq 7. The coding range is between −5 and 5.

$$L = n \times q + q + m \times q + m \tag{7}$$

2.2.2. Calculate Population Fitness. After initializing the population according to the chromosome coding value of each individual, sample data are used to test the BP neural network model under the condition of threshold and weight and calculate the difference between expected value and predicted value. The reciprocal of the sum of squares is used to evaluate the fitness of each individual (eq 8).

$$F = 1/\sum_{i=1}^{N} (d_i - D_i)^2 \tag{8}$$

where F is the individual fitness. N is the total number of samples. d_i is the predicted value of the ith sample. D_i is the actual value of the ith sample.

2.2.3. Compute Population Evolution Iterations. Optimized individuals are directly inherited to the next generation through selection, or new individuals are generated through pairwise crossover and then passed on to the next generation. The group size is 400, and the fitness of individual k is F. Then, the probability of individual i being selected is calculated based on the following:

$$p_k = F_i/\sum_{i=1}^{400} F_i \tag{9}$$

After calculation of the selection probability of each individual in the group, multiple rounds of selection are required. Each round generates a uniform random number between 0 and 1, and this random number is used as a pointer to determine the selected individual.

Crossover and mutation: the traditional genetic algorithm generally takes the probability of crossover and mutation as a fixed value, which restricts the diversity of the population in the iterative process and cannot effectively balance the stability of the population in the early and late stages of evolution. If the probability of crossover is too large, it is easy to make the population tend to be unitary, while if the probability of crossover P_c is too small, the speed of new individuals will be too slow; the effect of mutation probability P_m on the population is just opposite to that of crossover operation. Therefore, in order to ensure the diversity of the population in the genetic process and take into account the proportion of good individuals and new individuals in the whole population, this paper chooses a probability of adaptive change of crossover and mutation. The value of chromosome crossover variation in the genetic process is constantly changed according to the fitness, and the numerical adjustment formula is as follows:

$$c = \begin{cases} P_{cmax} - \dfrac{(P_{cmax} - P_{cmin})(f_{avg} - f')}{f_{avg} - f_{min}}, f' \le f_{avg} \\ P_{cmax}, f' > f_{avg} \end{cases} \tag{10}$$

$$P_m = \begin{cases} P_{mmax} - \dfrac{(P_{mmax} - P_{mmin})(f_{avg} - f')}{f_{avg} - f_{min}}, f' \le f_{avg} \\ P_{mmax}, f' > f_{avg} \end{cases} \tag{11}$$

where P_{cmax} and P_{mmax} are the maximum values of the crossover and mutation, respectively. P_{cmin} and P_{mmin} are the minimum values of crossover and mutation, respectively. f' is the current individual fitness. f_{avg} is the average fitness of the population.

2.3. Fracture Conductivity Prediction Process. The process for fracture conductivity prediction based on a BP neural network optimized by genetic algorithm is shown in Figure 2.

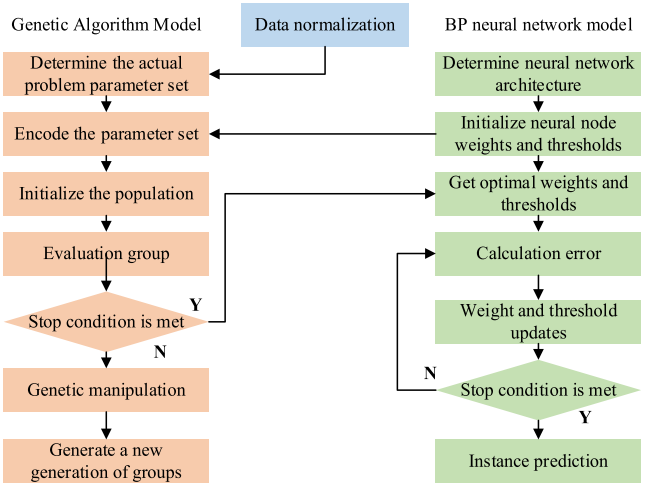


Figure 2. Algorithm-optimized BP neural network prediction process for fracture conductivity.

3. SAMPLE LIBRARY CONSTRUCTION AND DATA SET DIVISION

3.1. Data Cleaning. Due to equipment abnormalities and personnel errors, data quality problems, such as abnormal data or invalid data, are inevitable in fracture conductivity experimental data. In order to improve data quality, the boxplot method is used to clean the collected sample data. The advantage of this method is that it is not affected by outliers and can accurately and stably depict the discrete distribution of data, which is conducive to data cleaning. The inner limit of the box plot is used to detect and eliminate outliers and null values.[31] The difference between the upper quartile (Q3) and the lower quartile (Q1) of the data is called the interquartile range (IQR). The inner limit is calculated as follows:

$$\text{upper limit} = Q_3 + 1.5 \times \text{IQR} \tag{12}$$

$$\text{lower limit} = Q_1 - 1.5 \times \text{IQR} \tag{13}$$

The identification standard of abnormal points is to mark points outside the interval as abnormal points.

3.2. Sample Library Construction. The experimental data of fracture conductivity in this research come from two blocks. Block A is the shale outcrop of the Longmaxi Formation in Sichuan, which has a high clay mineral content of 31.0%, a high quartz content of 43.3%, and a low carbonate rock content of 15.3%. Block B is the outcrop of the Xujiahe

https://doi.org/10.1021/acsomega.4c00448
ACS Omega 2024, 9, 13469−13480

Formation in the Sichuan Basin, which has a high clay mineral content of 39.9%, a quartz content of 41.3%, and a low carbonate rock content of 13.1%. After data cleaning, the data box diagram of the conductivity of the two blocks is shown in Figure 3. A total of 502 sets of shale fracture conductivity experimental data in block A and 250 sets of shale fracture conductivity experimental data in block B are obtained. A data set is established based on the totaling 752 sets of experimental data.

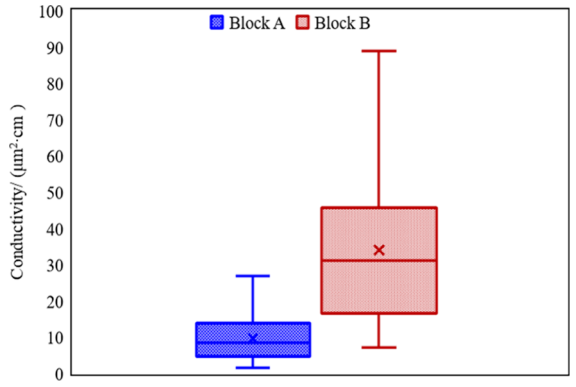


Figure 3. Conductivity data box diagram.

3.3. Exploratory Data Analysis. Exploratory data analysis (EDA) is the process of investigating the relationships, patterns, and collinearity among the input parameters. The experimental data include investigations of Young's modulus, sand concentration, temperature, Poisson's ratio, proppant particle size, closure pressure, and measured propped-fracture conductivity.

3.3.1. Statistical Analysis. The data were obtained from experiments on different shale formations. A complete statistical description of the data used for training is given in Table 1.

According to Table 1, it can be obtained that the Young's modulus of shale from block A is 31.46−42.25 GPa, and the Poisson's ratio is 0.208−0.271; the Young's modulus of shale from block B is 12.65−22.88 GPa, and the Poisson's ratio is 0.208−0.271. The average ratio is 0.147−0.219; overall, the Young's modulus of the shale used in the experiment is 12.65−42.25, and the Poisson's ratio is 0.147−0.271; the data distribution range is wide and can represent the rock mechanical properties of the shale reservoir in this oil field.

The parameter range used in the experimental data is sand concentration 2.5−10 (kg/m^2), average proppant particle size 30−105 (mesh), closure pressure 20−70 (MPa), and temperature 25−90 (°C); the data features are diverse enough to cover most fracking scenarios.

Table 1. Statistical Analysis of the Data Utilized for Modeling

parameters	conductivity/(μm^2·cm)	sand concentration/(kg·m^2)	proppant particle size/mesh	closure pressure/MPa	temperature/°C	Young's modulus/GPa	Poisson's ratio
			Overall data				
count	752.000	752.000	752.000	752.000	752.000	752.000	752.000
maximum	88.960	10.000	105.000	70.000	90.000	42.250	0.271
minimum	1.430	2.500	30.000	20.000	25.000	12.650	0.147
mean	17.597	4.574	68.238	48.112	48.138	30.669	0.221
standard deviation	17.122	2.992	31.025	17.383	23.222	9.474	0.033
median	12.240	3.000	55.000	50.000	32.000	34.497	0.224
variance	293.149	8.952	962.570	302.155	539.272	89.748	0.001
kurtosis	3.336	−0.474	−1.704	−1.374	−0.860	−1.223	−0.707
skewness	1.909	1.158	0.202	−0.143	0.763	−0.599	−0.472
			Block A data				
count	502.000	502.000	502.000	502.000	502.000	502.000	502.000
maximum	26.660	5.000	105.000	70.000	90.000	42.250	0.271
minimum	1.430	2.500	40.000	40.000	32.000	31.460	0.208
mean	9.476	2.790	84.124	57.251	56.566	36.996	0.240
standard deviation	5.498	0.530	25.535	11.838	23.302	3.077	0.018
median	8.290	2.500	105.000	60.000	60.000	37.038	0.240
variance	30.231	0.281	652.025	140.132	542.977	9.467	0.000
kurtosis	−0.634	10.514	−1.733	−1.421	−1.381	−1.131	−1.196
skewness	0.577	3.127	−0.438	−0.310	0.306	−0.067	0.003
			Block B data				
count	250.000	250.000	250.000	250.000	250.000	250.000	250.000
maximum	88.960	10.000	55.000	60.000	60.000	22.880	0.219
minimum	7.120	3.000	30.000	20.000	25.000	12.650	0.147
mean	33.904	8.156	36.340	29.760	31.216	17.965	0.183
standard deviation	20.576	2.671	7.807	11.083	10.163	2.981	0.020
median	30.935	10.000	30.000	30.000	25.000	18.192	0.183
variance	423.363	7.136	60.948	122.834	103.286	8.885	0.000
kurtosis	−0.435	−1.086	0.570	2.864	3.655	−1.175	−1.158
skewness	0.709	−0.848	1.142	1.832	2.181	−0.127	−0.004

https://doi.org/10.1021/acsomega.4c00448
ACS Omega 2024, 9, 13469−13480

3.3.2. Frequency Analysis. Frequency analysis is the statistics of the frequency of different values of a group of data or the frequency of data falling into a specified area, which can reflect the distribution status and distribution characteristics of all units in the population among groups. The conductivity data are grouped according to the quartile, and the number of individuals in each group is calculated, respectively. The conductivity data frequency analysis results of overall, block A and block B are shown in Table 2.

Table 2. Results of Frequency Analysis

data set	interval	frequency	percentage (%)	cumulative percentage (%)
overall	[1.43, 23.312)	595	79.122	79.122
	[23.312, 45.195)	93	12.367	91.489
	[45.195, 67.078)	37	4.92	96.41
	[67.078, 88.96]	27	3.59	100
total		752	100.000	100.000
block A	[1.43, 7.737)	235	46.813	46.813
	[7.737, 14.045)	151	30.08	76.892
	[14.045, 20.352)	97	19.323	96.215
	[20.352, 26.66]	19	3.785	100
total		502	100.000	100.000
block B	[7.12, 27.58)	114	45.6	45.6
	[27.58, 48.04)	77	30.8	76.4
	[48.04, 68.5)	32	12.8	89.2
	[68.5, 88.96]	27	10.8	100
total		250	100.000	100.000

According to the results of frequency analysis, there is an order of magnitude difference between the conductivity data of block A and block B, and the conductivity data of block A located in the interval [1.43, 14.045] account for 78.892%. The conductivity data located in the interval [14.045, 26.66] accounted for 23.108%. The conductivity data of block B located in the interval [7.12, 48.04] accounts for 76.4%, and the conductivity data located in the interval [48.04, 88.96] account for 23.6%. It can be seen that there is a certain imbalance in the sample data. From the overall data point of view, the proportions of diversion capacity data located in [1.43, 45.195] and [45.195, 88.96] are 91.489 and 8.511%, respectively. The data distribution is extremely uneven. Using all data sets to train the model may lead to weight imbalance of the model, deviation in the prediction results, and reduction of generalization ability. In order to reduce the influence of order of magnitude and unbalanced data distribution on model training, it is necessary to normalize the data.

3.4. Data Normalization. Due to the different orders of magnitude and dimensions of different parameters, the value ranges greatly, which affects the accuracy and solution speed of machine learning modeling. Parameter is processed using min−max normalization. This method can map the value range of each parameter to 0 and 1 and retains the relationship between the original data. The calculation method is shown in eq 14.

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{14}$$

where x is the sample data. $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the sample data, respectively. x' is the result of the data normalization.

3.5. Controlling Factors of Fracture Conductivity. In the process of machine learning, more input parameters can contain more information, but the calculation speed and accuracy decrease as the number of features increases. Therefore, the Pearson coefficient method is used to analyze the correlation between influencing factors and fracture conductivity. This method is not affected by the scale change of variables and can intuitively determine the linear correlation between data. The gray correlation method is used to extract the main controlling factors of fracture conductivity, select the most useful features, improve the quality of features, and improve the calculation speed and accuracy of the model.

3.5.1. Pearson Coefficient Method. The Pearson correlation coefficient is a dimensionless statistical index that can express the degree and direction of linear correlation between two variables.[32] The correlation coefficient is represented by the symbol r, and its value range is $-1 \le r \le 1$. If the correlation coefficient is less than 0, it is a negative correlation. If it is greater than 0, it is a positive correlation. If it is equal to 0, it means that there is no correlation. The calculation formula of r is shown in eq 15.

$$r = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \tag{15}$$

3.5.2. Gray Correlation Analysis. Gray correlation analysis quantitatively describes and compares the development and change situation of a system.[33] The specific calculation method is as follows:

(a) Analysis sequence determination. Determine the comparison sequence $x_i'(k)(i = 1,2,3, \cdots, n)$, and reference sequence $x_0'(k)(k = 1,2,3, \ldots, m)$. n is the number of independent variables, and m is the number of elements in each sequence.

(b) Dimensionless sequence. The data in the original sequence will affect the accuracy of calculation results due to differences in magnitude and dimension. The mean method is used for sequence processing:

$$x_i(k) = \frac{x_i'(k)}{\frac{1}{m}\sum_{k=1}^{m} x_i'(k)}, \quad (i = 0, 1, 2, \ldots, n;\ k = 1, 2, \ldots, m) \tag{16}$$

(c) Correlation coefficient calculation.

$$r_{0i}(x_0(k), x_i(k)) = \frac{\min\limits_{i=1}^{n}\min\limits_{k=1}^{m}|x_i(k) - x_0(k)| + \rho \times \max\limits_{i=1}^{n}\max\limits_{k=1}^{m}|x_i(k) - x_0(k)|}{|x_i(k) - x_0(k)| + \rho \times \max\limits_{i=1}^{n}\max\limits_{k=1}^{m}|x_i(k) - x_0(k)|} \tag{17}$$

where $r_{0i}(x_0(k), x_i(k))$ is the correlation coefficient of the kth point, dimensionless.$\min\limits_{i=1}^{n}\min\limits_{k=1}^{m}|x_i(k) - x_0(k)|$ and $\max\limits_{i=1}^{n}\max\limits_{k=1}^{m}|x_i(k) - x_0(k)|$ are the minimum and maximum differences between the two levels, respectively. ρ is the resolution coefficient, dimensionless; $|x_i(k) - x_0(k)|$ is the absolute difference sequence between the ith comparison sequence and the reference sequence.

(d) Correlation degree calculation. The correlation degree of two sequences can be calculated by eq 18.

https://doi.org/10.1021/acsomega.4c00448
ACS Omega 2024, 9, 13469−13480

Closure pressure	-0.714	-0.479	-0.562	0.639	0.4	-0.627	1
Proppant particle size	0.691	0.479	0.562	-0.639	-0.4	1	-0.627
Poisson's ratio	-0.259	-0.38	-0.196	0.032	1	-0.4	0.4
Temperature	-0.473	0.071	-0.096	1	0.032	-0.639	0.639
Sand concentration	0.666	0.123	1	-0.096	-0.196	0.562	-0.562
Young's modulus	0.269	1	0.123	0.071	-0.38	0.479	-0.479
Conductivity	1	0.269	0.566	-0.473	-0.259	0.691	-0.714
	Conductivity	Young's modulus	Sand concentr-ation	Temperature	Poisson's ratio	Proppant particle size	Closure pressure

Figure 4. Pearson coefficient heat map.

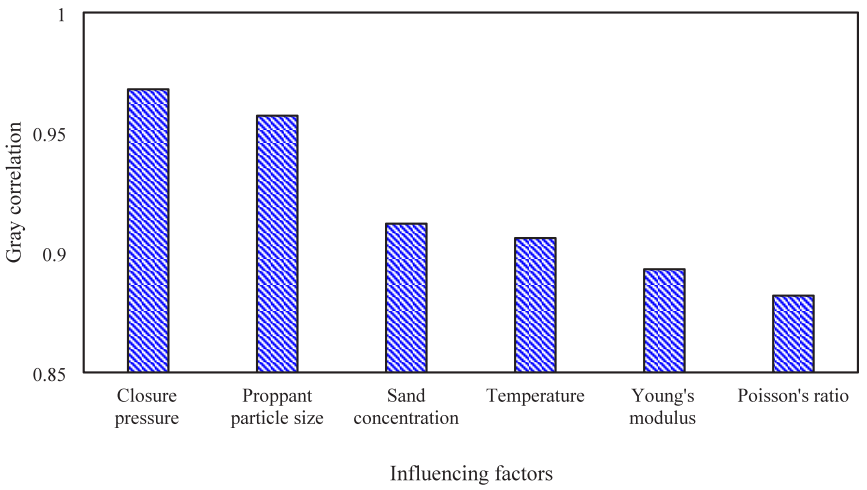


Figure 5. Gray correlation degree calculation results.

$$r_{0i}(x_0, x_i) = \sum_{k=1}^{n} \omega_k r(x_0(k), x_i(k)) \quad (18)$$

where $r_{0i}(x_0, x_i)$ is the degree of correlation between compared sequence and reference sequence, dimensionless. ω_k is the weight of k points, usually $\omega_k = \frac{1}{n}$, dimensionless.

(e) Relevance sorting. Relevance is sorted by size after the results.

3.5.3. Pearson Coefficient Calculation Results. The calculation results of the Pearson coefficient between influencing factors and fracture conductivity are shown in Figure 4.

According to the results, it can be seen that the fracture conductivity is negatively correlated with closure pressure, temperature, and Poisson's ratio and is positively correlated with proppant particle size, sand concentration, and Young's modulus.

3.5.4. Controlling Factor Analysis. Taking fracture conductivity taken as a reference sequence and each influencing factor as a comparison sequence, the correlation ranking between each influencing factor and fracture conductivity is calculated through gray correlation analysis. The result is shown in Figure 5. It can be found that the gray correlation between closure pressure and fracture conductivity is the highest, reaching 0.968, which is consistent with the calculation results of the Pearson coefficient. Taking a correlation degree of 0.85 as the judgment standard, influencing factors with correlation degree greater than 0.85 are regarded as the main controlling factors. Therefore, closing pressure, proppant particle size, sand dressing concentration,

https://doi.org/10.1021/acsomega.4c00448
ACS Omega 2024, 9, 13469−13480

temperature, Young's modulus, and Poisson's ratio are all related to fracture conductivity, which should be selected during model application.

3.6. Data Set Partitioning. Taking closure pressure, proppant particle size, sand dressing concentration, temperature, Young's modulus, and Poisson's ratio as input features, and fracture conductivity as the output target, the collected data are processed into a training set and a testing set. 80% of the total data is randomly selected as training set, and the remaining 20% is selected as testing set.

The key to the prediction effect of the model is to train and test the model on different data segments. Given the small size of the data set, a single data segmentation method (80% training and 20% testing) may not be able to evaluate the effectiveness of the model. Therefore, the fivefold cross validation technique is used to train the model on 80% of the training set.

4. MODEL APPLICATION

4.1. Model Structure Optimization. Hornik et al.[34] used functional analysis theory to prove that under a wide range of conditions, a three-layer neural network can approximate any function and its derivatives of all orders with arbitrary accuracy. Therefore, this paper uses the neural network structure of a single hidden layer to build the calculation model. The number of neurons in the hidden layer is an important superparameter in the neural network structure, which directly affects the capacity and learning ability of the network. In order to improve the prediction effect of the model, this parameter is optimized.

Multiple BP neural networks with input layer 5, hidden layers 5−30, and output layer 1 are built. 80% of the total data is randomly selected as training set, and the remaining 20% is selected as validation set, and the Early Stopping callback function is used to stop training when the model performance is no longer improved. At the same time, the initial weight and bias super parameters of the above neural network are optimized by the GA algorithm. Using the mean square error as the loss function, after three times of repeated sampling to calculate the average value, the loss function of the number of neurons in different hidden layers is shown in Figure 6.

It can be seen that when the number of neurons in the hidden layer is too small, it is difficult for the neural network to capture complex data relations, resulting in underfitting of the model, and the loss of both the training set and validation set is higher. However, when there are too many neurons, the loss of the validation set is higher than that of the training set, the model is overfitted, and the generalization decreases. When the number of neurons in the hidden layer is 20, the GA-BP network shows better results in the training and validation sets. Therefore, the GA-BP neural network model of 5−20−1 is selected to calculate the conductivity and the loss curve of the training set and the validation set is shown in Figure 7.

Figure 6. Loss curves in different hidden layer sizes.

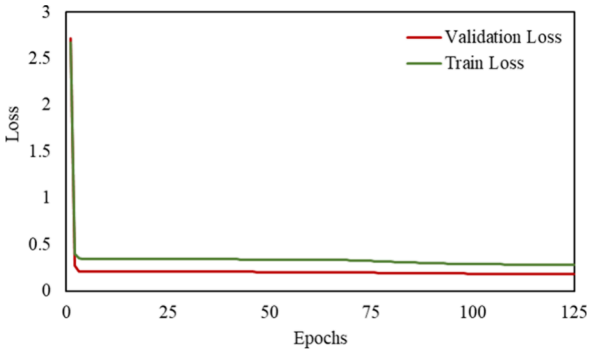


Figure 7. Fitness curve of GA.

4.2. Model Training. The hyperparameters of the proposed GA-BP neural network model after tuning are shown in Table 3.

Table 3. Optimum Values for the Model

GA	
parameters	values
number of population	400
P_{cmax}	0.5
P_{mmin}	0.1
P_{cmax}	0.05
P_{mmin}	0.01
BPNN	
parameters	values
inputs	6
hidden layer	1
neurons in the hidden layer	20
outputs	1
learning rate	0.1
activation function of the middle layer	Relu
activation function of the outer layer	linear

Two schemes are used to train the machine learning model. In the first scenario, 80% of all experimental data are randomly selected and the fivefold cross validation method is applied to train the model (fourfold is used for training, and onefold is used for validation). The model can test all of the training data. The results are shown in Figure 8.

The average results of cross validation show that the R^2 value of the validation set is 0.927, the root mean square error (RMSE) is 0.047, and the average absolute percentage error (MAPE) is 21.084%. However, it can be found from Figure 8 that due to the nonuniformity of data distribution in the whole data set, the prediction effect of high conductivity data is poor and deviates greatly from the 45° line. It shows that data normalization cannot completely eliminate the impact of uneven data distribution on the model. Therefore, considering the second training scheme, using 80% of the experimental

https://doi.org/10.1021/acsomega.4c00448
ACS Omega 2024, 9, 13469−13480

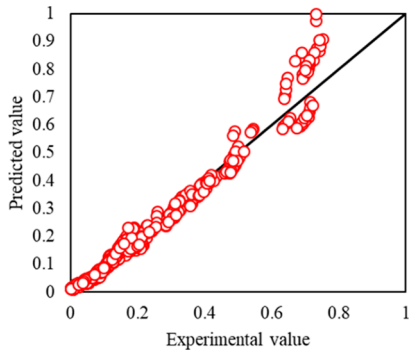


Figure 8. Training cross validation results of the overall data set.

data of shale in block A and block B, respectively, the model is trained by a fivefold cross validation method. The results are shown in Figures 9 and 10.

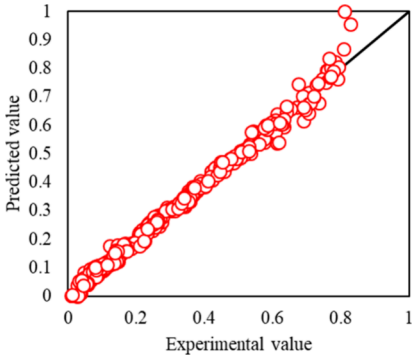


Figure 9. Training cross validation results of the block A data set.

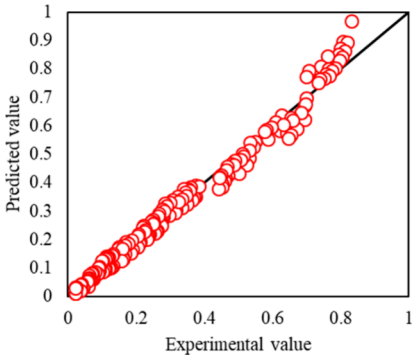


Figure 10. Training cross validation results of the block B data set.

For block A, the average results of cross validation show that the R^2 of the validation set is 0.975, the RMSE is 0.032, and the MAPE is 9.482%. For block B, the average results of cross validation show that the R^2 of the validation set is 0.962, the RMSE is 0.044, and the MAPE is 11.968%. Compared with the overall data set, the data distribution of blocks An and B is more uniform, so distinguishing block data to train the model can improve the performance of the model. It can be seen from the chart that although some of the predicted results of high conductivity are still quite different from the real values, the predicted values of conductivity of block A and block B are relatively close to the 45° line as a whole and the average relative error of the validation set is controlled within 15%, which shows that the generalization ability of the model is good.

4.3. Example Application. The remaining 20% of the data are predicted by using the trained model, and the prediction result of diversion capacity is shown in Figure 11. According to the results, the RMSE of the whole data set, block A, and block B are 0.048, 0.027, and 0.039, MAPE is 18.709, 9.289, and 10.839. The prediction effect is good, indicating that the model has a certain generalization ability.

In order to verify the prediction performance of the model, the repeated sampling technique is used to randomly divide the data set according to the Section 3.6 method and the model is trained according to the Section 4.2 Combined with the results of the first sampling, the prediction effect of the model is shown in Table 4.

According to Table 4, the average MAPE of the model for the overall data, block A, and block B is 17.392, 9.722, and 10.38%, respectively, indicating that the training model for distinguishing block data can get better prediction results and the GA-BP model has higher accuracy.

5. DISCUSSION

In this paper, the BP neural network model optimized by the genetic algorithm is used to predict fracture conductivity for the first time, which can better adapt to the complex fracture conductivity process and realize a fast, cheap, and accurate method to predict fracture conductivity. This method is convenient for field application, can effectively guide field production, is of great significance for improving oilfield development efficiency and economic benefits, and has strong engineering application value.

Compared with Desouky et al.'s[27] research contents, this paper contains a total of 752 sets of data; the advantage of the amount of data is significant. This makes the model in this paper able to more comprehensively learn the patterns and relationships in the process of fracture diversion, which provides a more sufficient and reliable data basis for the whole process prediction of fracture conductivity and strengthens the science and practicability of the research. In this paper, a genetic algorithm is used to optimize the BP neural network to avoid the dilemma of the model falling into the local optimal solution. Therefore, the meticulous work in model design and optimization in this paper makes the model more reliable and generalized than previous studies. This systematic model advantage not only improves the accuracy of fracture conductivity but also provides reliable tools and methods for in-depth research in this field.

According to the results, the uneven distribution of data will reduce the training effect of the model and distinguishing blocks to train the model can improve the performance of the model. Therefore, in the process of practical application, it is suggested that one establish a separate model for each block.

This study is based on a large number of experimental data, and the research results will provide guidance for the field. Considering the general equipment conditions in the oil field and the limited computing resources, the proposed method is purely data-driven. Because the pure data-driven model is usually lightweight while ensuring the accuracy of the results, the computational cost of training and reasoning is relatively low. It is worth noting that some scholars have proposed a neural network model that includes physical processes. Guo et al.[35] present a stochastic deep collocation method (DCM) based on neural architecture search (NAS) and transfer learning for heterogeneous porous media; Samaniego et al.[36] used DNNs to solve boundary value problems, proving that it

https://doi.org/10.1021/acsomega.4c00448
ACS Omega 2024, 9, 13469−13480

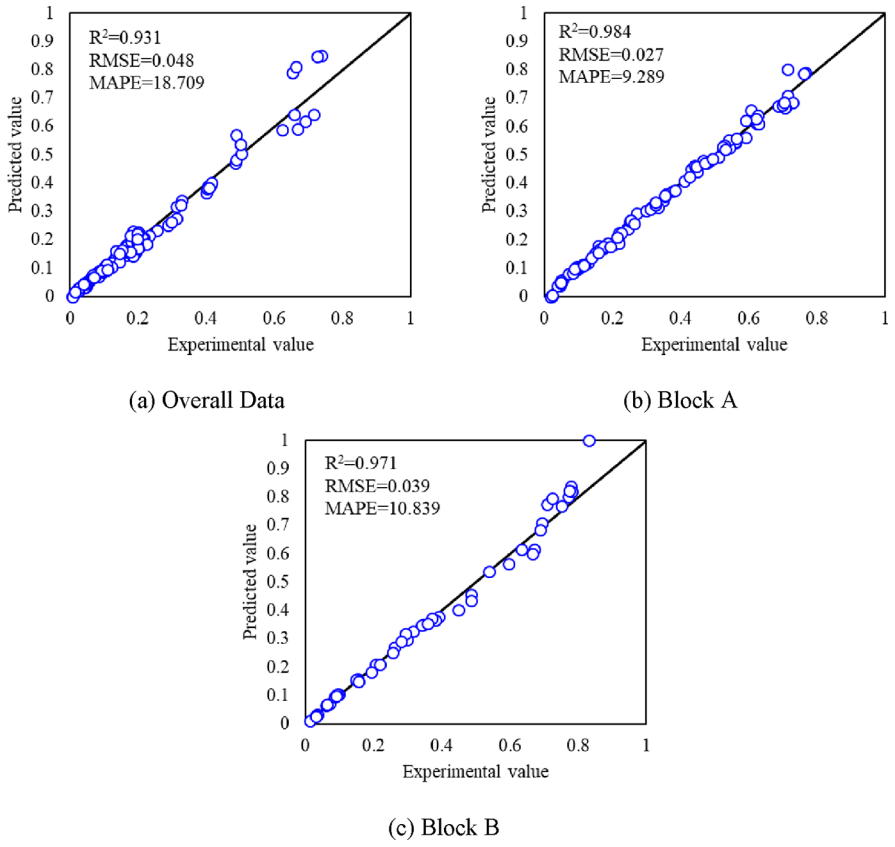


Figure 11. Case test results.

Table 4. Comparison of Model Accuracy for the Two Scenarios

	overall data			block A			block B		
parameters	sample 1	sample 2	mean	sample 1	sample 2	mean	sample 1	sample 2	mean
R^2	0.931	0.942	0.937	0.984	0.977	0.981	0.971	0.979	0.975
RMSE	0.048	0.039	0.044	0.027	0.034	0.031	0.039	0.035	0.037
MAPE/%	18.709	16.074	17.392	9.289	10.155	9.722	10.839	9.921	10.38

is possible to tackle the solution of very relevant BVPs using concepts and tools coming from deep learning. By adding physical information to the neural network model, we can provide additional prior knowledge to the model and help the model understand the basic laws of the system. The model may need less data for training and to ensure computational accuracy. This is especially beneficial in areas where data are scarce or expensive. Models that contain physical processes are usually more explanatory, which helps to explain the decision-making process and the results of the model. Moreover, the integration of physical processes can increase the robustness of the model to disturbance; even if there are noise or outliers in the input data, the model can better adapt and produce reliable output. Compared with these models, the data-driven prediction model established in this paper has limitations. Therefore, using a neural network with physical information to model and predict fracture conductivity will be the direction and focus of future research.

6. CONCLUSIONS

(1) The main controlling factors of fracture conductivity are determined by the Pearson coefficient method and gray correlation analysis, which are in the following order of importance: closure pressure, particle size of proppant, sand concentration, temperature, Young's modulus, and Poisson's ratio.

(2) The results of data analysis show that the Young's modulus of the shale used in the experiment is 12.65−42.25, and the Poisson's ratio is 0.147−0.271; the data distribution range is wide and can represent the rock mechanical properties of the shale reservoir in this oil field. However, compared with the data of each block, the data distribution of the conductivity of the whole data set is uneven. Using all data to train the model may lead to the deviation in the prediction results, and the use of data normalization method cannot eliminate the impact of uneven data distribution on the model.

(3) The GA algorithm is used to optimize the BP neural network, and the prediction model of GA-BP conductivity is established. The fivefold cross validation technique is used to train the model, and the model predicts the conductivity capacity on all data, block A, and block B. The average MAPE results of the two samples are 17.392, 9.722, and 10.38%. It shows that the uneven distribution of data will reduce the training effect of the model, and distinguishing blocks to train the model can improve the performance of the model.

https://doi.org/10.1021/acsomega.4c00448
ACS Omega 2024, 9, 13469−13480

Therefore, in the process of practical application, it is suggested to establish a separate model for each block.

(4) The GA-BP model can accurately predict the fracture conductivity. This method avoids complex mathematical calculation, has good adaptability and practicability, and can provide technical support on the spot.

AUTHOR INFORMATION

Corresponding Author

Xiaoqiang Liu − *School of Energy Resources, China University of Geosciences Beijing, Beijing 100083, China; School of Earth and Spacing Sciences, Peking University, Beijing 100871, China;* orcid.org/0000-0001-5694-5976; Email: liuxiaoqiang0535@126.com

Authors

Xiaopeng Wang − *State Key Laboratory of Offshore Oil Exploitation, Beijing 100028, China; Tianjin Branch of CNOOC Ltd., Tianjin 300450, China*

Binqi Zhang − *State Key Laboratory of Offshore Oil Exploitation, Beijing 100028, China; Tianjin Branch of CNOOC Ltd., Tianjin 300450, China*

Jianbo Du − *State Key Laboratory of Offshore Oil Exploitation, Beijing 100028, China; Tianjin Branch of CNOOC Ltd., Tianjin 300450, China*

Dongdong Liu − *State Key Laboratory of Offshore Oil Exploitation, Beijing 100028, China; Tianjin Branch of CNOOC Ltd., Tianjin 300450, China*

Qilong Zhang − *State Key Laboratory of Offshore Oil Exploitation, Beijing 100028, China; Tianjin Branch of CNOOC Ltd., Tianjin 300450, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.4c00448

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This study was sponsored by the National Natural Science Foundation of China (Grant No. 52204024), CNPC Innovation Found (Grant No. 2021DQ02-1006), and State and Key Laboratory of Offshore Oil Exploitation.

REFERENCES

(1) Abdelaziz, A.; Ha, J.; Li, M.; Magsipoc, E.; Sun, L.; Grasselli, G. Understanding hydraulic fracture mechanisms: From the laboratory to numerical modelling. *Advances in Geo-energy Research* **2023**, *7* (1), 66−68.

(2) Liu, X.; Qu, Z.; Guo, T.; Tian, Q.; Lv, W.; Xie, Z.; Chu, C. An innovative technology of directional propagation of hydraulic fracture guided by radial holes in fossil hydrogen energy development. *Int. J. Hydrogen Energy* **2019**, *44* (11), 5286−5302.

(3) Duenckel, R. J.; Barree, R. D.; Drylie, S; et al. proppants what 30 years of study has taught us. *SPE Annual Technical Conference and Exhibition*, SPE: San Antonio, 2017 DOI: 10.2118/187451-MS.

(4) Hui, G.; Chen, Z. X.; Chen, S. N.; Gu, F. Hydraulic fracturing-induced seismicity characterization through coupled modeling of stress and fracture-fault systems. *Advances in Geo-energy Research* **2022**, *6* (3), 269−270.

(5) Hou, B.; Chen, M.; Wang, Z.; Yuan, J. B.; Liu, M. Hydraulic fracture initiation theory for a horizontal well in a coal seam. *Petroleum Science* **2013**, *10* (2), 219−225.

(6) Qi, J.; Zhang, L. M.; Zhang, K.; Li, L. X.; Sun, J. J. The application of improved differential evolution algorithm in electro-magnetic fracture monitoring. *Advances in Geo-energy Research* **2020**, *4* (3), 233−246.

(7) Lu, Y. H.; Chen, K. P.; Jin, Y.; Li, H. D.; Xie, Q. An approximate analytical solution for transient gas flows in a vertically fractured well of finite fracture conductivity. *Petroleum Science* **2022**, *19* (6), 3059−3067.

(8) Qu, H. Y.; Zhang, J. L.; Zhou, F. J.; Peng, Y.; Pan, Z. J.; Wu, X. Y. Evaluation of hydraulic fracturing of horizontal wells in tight reservoirs based on the deep neural network with physical constraints. *Petroleum Science* **2023**, *20* (2), 1129−1141.

(9) Wang, Y.; Ren, Y.; Zhang, B.; et al. Study on the influencing factors of proppant embedment in hydraulic fracturing of shale reservoir. *Drill. Prod. Technol.* **2020**, *43* (4), 129−132.

(10) Wu, S.; Liu, H.; Li, X.; et al. Test and application of subdivision fracture control fracturing for tight oil horizontal wells in Ordos Basin. *Drill. Prod. Technol.* **2020**, *43* (3), 53−55.

(11) He, Y.; Zhao, S.; Lun, Z.; et al. Analysis of rheological and filtration properties of supercritical CO2. *Drill. Prod. Technol.* **2020**, *43* (3), 38−41.

(12) Xiao, H.; Xia, X.; Zhang, L.; et al. Comparative Experimental Study of Fracture Conductivity of Carbonate Rocks under Different Stimulation Types. *ACS Omega* **2023**, 49175.

(13) Guo, S.; Wang, B.; Li, Y.; et al. Impacts of Proppant flowback on Fracture Conductivity in Different Fracturing Fluids and flowback Conditions[J]. *ACS omega* **2022**, *7* (8), 6682−6690.

(14) Sun, H.; He, B.; Xu, H.; et al. Experimental investigation on the fracture conductivity behavior of quartz sand and ceramic mixed proppants[J]. *ACS omega* **2022**, *7* (12), 10243−10254.

(15) Tariq, Z.; Hassan, A.; Al-Abdrabalnabi, R.; et al. Comparative study of fracture conductivity in various carbonate rocks treated with GLDA chelating agent and HCl acid[J]. *Energy Fuels* **2021**, *35* (23), 19641−19654.

(16) Zhou, S.; Zhuang, X.; Rabczuk, T. A phase-field modeling approach of fracture propagation in poroelastic media[J]. *Engineering Geology* **2018**, *240*, 189−203.

(17) Rabczuk, T.; Belytschko, T. Cracking particles: a simplified meshfree method for arbitrary evolving cracks[J]. *International journal for numerical methods in engineering* **2004**, *61* (13), 2316−2343.

(18) Zhang, J.; Zhu, D.; Hill, A. D. Water-induced damage to propped-fracture conductivity in shale formations. *SPE Production & Operations* **2016**, *31* (02), 147−156.

(19) Jia, L.; Li, K.; Zhou, J.; et al. A mathematical model for calculating rod-shaped proppant conductivity under the combined effect of compaction and embedment. *J. Pet. Sci. Eng.* **2019**, *180*, 11.

(20) Xu, J.; Ding, Y.; Yang, L.; et al. Conductivity analysis of tortuous fractures filled with non-spherical proppants. *J. Pet. Sci. Eng.* **2021**, *198*, No. 108235.

(21) Zhang, F.; Zhu, H.; Zhou, H.; et al. Discrete-element-method/computational-fluid-dynamics coupling simulation of proppant embedment and fracture conductivity after hydraulic fracturing. *SPE Journal* **2017**, *22* (02), 632−644.

(22) Su, Y.; Wu, Z.; Cui, C.; et al. Calculation and analysis of the propped fracture conductivity created by hydraulic fracturing. *Pet. Geol. Oilfield Dev. Daqing* **2021**, *40* (6), 62−71.

(23) Shi, J. F.; Yu, S. S.; Zhang, L.; et al. Prediction model of fracture conductivity in tight oil reservoirs. *Drill. Prod. Technol.* **2021**, *44* (1), 82−86.

(24) Kainer, C.; Guerra, D.; Zhu, D.; et al. A comparative analysis of rock properties and fracture conductivity in shale plays. *SPE Hydraulic Fracturing Technology Conference and Exhibition*. SPE **2017**, D011S002R008 DOI:

(25) Zhu, H. Y.; Liu, Y. J.; Wang, X. Y.; et al. Modeling on conductivity of branched fractures of shale gas reservoir considering proppant fragmentation. *J. China Univ. Pet., Ed. Nat. Sci.* **2022**, *46* (1), 72−79.

(26) Chen, Q. D.; Zhou, J. Y.; Chen, W. Y.; et al. Experimental study on the change law of conductivity of proppant combinations with different particle sizes. *Liaoning Chem. Ind.* **2022**, *51* (5), 593−595.

(27) Desouky, M.; Tariq, Z.; Aljawad, M. S.; et al. Machine learning-based propped fracture conductivity correlations of several shale formations[J]. *ACS omega* **2021**, *6* (29), 18782−18792.
(28) Hua, X.; Zhang, G.; Yang, J.; Li, Z. Theory Study and Application of the BP-ANN Method for Power Grid Short-Term Load Forecasting. *ZTE Commun.* **2015**, *13* (3), 2−5.
(29) Liu, Y. Y.; Liu, Y. S.; Zheng, J. W. Intelligent rapid prediction method of urban flooding based on BP neural network and numerical simulation model. *J. Hydraul. Eng.* **2022**, *53*, 284−295.
(30) Zhang, Y. L.; Jin, H.; Gao, W. W.; et al. Research on soil water content inversion of digital images based on BP neural network optimized by genetic algorithm. *Water Saving Irrigation* **2022**, *12*, 74−80.
(31) Hongyi, S.; Fangfang, X.; Xinmin, W. Research on cleaning and repairing methods of civil building data on resources saving and environment protection. *Beijing Da Xue Xue Bao* **2020**, *56* (5), 785−795.
(32) Wang, J.; Wu, X. M.; Wang, A. F. Apply Pearson correlation coefficient algorithm to find abnormal energy meter users. *Power Demand Side Manage.* **2014**, *16* (02), 52−54.
(33) Sun, J.; Dang, Y.; Zhu, X.; et al. A grey spatiotemporal incidence model with application to factors causing air pollution. *Sci. Total Environ.* **2021**, *759*, No. 143576.
(34) Hornik, K.; Stinchcombe, M.; White, H. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural networks* **1990**, *3* (5), 551−560.
(35) Guo, H.; Zhuang, X.; Chen, P.; Alajlan, N.; Rabczuk, T. Stochastic deep collocation method based on neural architecture search and transfer learning for heterogeneous porous media. *Engineering with Computers* **2022**, *38* (6), 5173−5198.
(36) Samaniego, E.; Anitescu, C.; Goswami, S.; Nguyem-Thanh, V. M.; Guo, H.; Hamdia, K.; Zhuang, X.; Rabczuk, T. An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications. *Comput. Methods Appl. Mech. Eng.* **2020**, *362*, No. 112790.

https://doi.org/10.1021/acsomega.4c00448
ACS Omega 2024, 9, 13469−13480