Journal of Computer Science and Artificial Intelligence
ISSN: 3078-8242 (Print), ISSN: 3078-2309 (Online) | Vol. 3, No. 1, 2025

Application of PSO-LSTM Neural Network Based on Improved Wavelet Denoising to Predict Drilling Rate of Penetration

Xinjie Fang [1], Zhongzhi Hu [1], Yan Zhou [2], Wei Song [2], Junying Pan [2]

[1] College of Mechanical Engineering, Sichuan University of Science and Engineering, Yibin 644000, China
[2] Jidong Oilfield Company, Petrochina Company Limited, Tangshan 063000, China

Abstract: Accurate prediction of the rate of penetration (ROP) is one of the important factors to improve the drilling efficiency of oil fields. At present, most of the established ROP prediction models have low accuracy, because the traditional noise reduction methods are generally used to process the drilling data. At the same time, the influence of wellbore friction and well deviation on WOB is not considered, and only the ground WOB is qualitatively used as the input parameter of the ROP prediction model. In this paper, seven deep directional wells in the central part of Bohai Bay, China are selected as the data source. Firstly, the drilling data are cleaned and screened, and the improved wavelet denoising method is adopted to effectively eliminate the noise interference while ensuring the data quality. Then the ground drilling pressure is corrected to improve the rationality of the data. The random forest feature selection method is used to analyze the correlation of various factors affecting the penetration rate and reduce the redundancy of the model. Finally, the particle swarm optimization algorithm is used to optimize the LSTM neural network mechanical drilling speed prediction model. The influence of traditional wavelet noise reduction method, ground drilling pressure and LSTM model hyperparameters on the error of mechanical drilling speed prediction model is discussed. The results show that the average error of the prediction model of the rate of penetration is reduced after the noise reduction method and the ground drilling pressure are modified. The optimized LSYM prediction model of the rate of penetration has higher prediction accuracy than other models, and the prediction accuracy reaches 87.3 %, which further verifies the superiority and effectiveness of the prediction model of the rate of penetration established in this paper.

Keywords: Wavelet noise reduction; ROP prediction; LSTM neural network; Particle swarm optimization.

1. Introduction

In order to better develop deep oil and gas resources, it is the general trend to shorten the drilling cycle, reduce drilling costs and improve drilling efficiency. The most direct and effective way to solve these problems is to increase the rate of penetration. The rate of penetration is an important index in drilling engineering, which is directly related to the mining cost and mining time. If the rate of penetration can be accurately predicted, the drilling time can be estimated, the allocation of resources can be optimized, and the efficiency of personnel and equipment can be improved, to achieve the purpose of reducing the mining cost [1]. With the advent of big data, artificial intelligence technology has developed rapidly. Applying intelligent technology to the intelligent equipment for drilling speed prediction in new blocks, we can get the adjustment value of various drilling equipment parameters under the best prediction of drilling speed in advance, to effectively guide drilling engineering. At the same time, in the drilling process, the drilling parameters can be adjusted according to the optimal prediction of the rate of penetration in real-time drilling, and the drilling efficiency can be improved.

In 1969, Young [2] made the first attempt of real-time drilling optimization, in which constant WOB and rotation speed were assumed in the drilling application of PDC bit and rock bit. The proposed model used a combination equation of ROP, bit bearing wear, tooth or tool wear and cost. In 1974, Bourgoyne and Young [3] established a ROP prediction numerical model, which is one of the earliest regression models and one of the most comprehensive experimental models recognized by the industry. It has been used as a standard and reliable method to improve drilling efficiency [4-7]. However, this ROP prediction model has many non-linear factors that have not been considered, such as lithology, gamma coefficient, acoustic time difference, riser pressure, etc., relying on the traditional empirical formula model to predict the ROP has not been able to meet the needs of current drilling engineering.

By imitating the mechanism of human brain learning prediction and decision-making [8], the machine learning method can carry out deep non-linear mapping of the laws between data, and is considered by researchers as a suitable alternative method for predicting the rate of penetration [9 ~ 10]. Table 1 shows the research status of foreign and domestic scholars using artificial intelligence methods to predict the rate of penetration.

In order to solve these problems, this paper uses the following methods to model the drilling speed prediction model. Firstly, the data is cleaned and normalized by the method of data preprocessing, and then the improved wavelet denoising method is adopted to effectively eliminate the noise interference while ensuring the data quality. At the same time, the ground drilling pressure is corrected to improve the rationality of the data. The random forest feature selection method is used to analyze the correlation of various factors affecting the penetration rate and reduce the redundancy of the model. Finally, the particle swarm optimization algorithm is proposed to optimize the LSTM neural network mechanical drilling speed prediction model.

Table 1. Research status at home and abroad

author	time	title	model technique	defect
Chao Gan [11]	2019	Prediction of drilling rate of penetration (ROP) using hybrid support vector regression: A case study on the Shen Nong jia area, Central China	Hybrid bat algorithm optimizes support vector machine algorithm	The problem of downhole drilling pressure is not considered
Liqi [12~13]	2021	Drilling rate prediction model based on PSO-BP	Particle swarm optimization BP neural network model	There is no preprocessing of the missing value and discrete value of the data, the data denoising method loses too much information, and the real drilling pressure is directly replaced by the drilling pressure measured on the ground, and the model error is large.
		Based on BAS-BP, the drilling rate prediction model of drilling machinery is established.	Beetle whisker algorithm to optimize BP neural network	
Hongtao Liu [14]	2022	Rate of Penetration Prediction Method for Ultra-Deep Wells Based on LSTM–FNN	LSTM-FNN algorithm	No noise treatment is carried out and the problem of downhole drilling pressure is not considered.
Ehsan Brenjkar [15]	2022	Computational predictionof the drilling rate of penetration (ROP): A comparison of various machine learning approaches and traditional models	Four methods of multi-layer perceptron neural network, radial basis function neural network, adaptive neural fuzzy inference system and support vector regression are combined with heuristic algorithm.	Directly taking the ground drilling pressure as the model input, there is a problem that the downhole drilling pressure is not considered.
Shi Xiangchao [16]	2022	Drilling rate prediction method based on LSTM recurrent neural network model	LSTM	No noise processing, no drilling pressure correction, but a simple qualitative analysis of the data.
Zhang Ligang [17]	2022	Based on MEA-BP neural network, the prediction of drilling speed of drilling machinery is carried out.	Mind evolutionary algorithm optimizes BP neural network model	The hard threshold method of wavelet de-noising is selected in the de-noising method, which makes the drilling data oscillate and experiment in a vertical well, which is not practical enough.
Liu Ye [18]	2023	Time series feature characterization and prediction method of ROP based on Attention-LSTM	LSTM	There is no noise processing and no consideration of the real downhole drilling pressure, and the error is large.
Mu huayan [19]	2023	Optimization of drilling rate prediction model based on mechanical specific energy	KNN	There is no noise processing, no consideration of downhole drilling pressure.
Ji Hui [20]	2023	A prediction method of penetration rate based on particle swarm optimization LSTM neural network model	PSO-LSTM	The Pearson correlation coefficient method is used for feature selection, ignoring the non-linear relationship between the factors affecting the rate of penetration, no noise reduction processing, and directly replacing the real drilling pressure measured on the ground with the real drilling pressure in the well. The model error is large.

2. Principle and Method

2.1. Wavelet De-noising

Due to the change of drilling tool structure, drill string vibration and other factors, the collected drilling data contains large noise when the data is collected at the drilling site. Wavelet denoising is an excellent signal denoising method. It has good time-frequency characteristics and can well describe the unbalanced parts of the original signal data, such as spikes and breakpoints. Wavelet denoising includes modulus maxima method and threshold method [21 ~ 22]. Using wavelet modulus maxima method to denoise, the requirement of wavelet decomposition scale will be very high, because the noise has a great influence on the wavelet coefficients on the small scale, which may produce many pseudo extreme points, resulting in more information loss. The wavelet threshold denoising method was proposed by American scholar Donoghue [23]. Compared with the modulus maxima method, this method is simple in calculation, can largely suppress noise, and can retain the information of the original signal well. It is a simple and effective method and has been applied in many fields [24 ~ 27]. Therefore, this paper uses wavelet threshold denoising method to process the noise of drilling data. The quality of wavelet threshold denoising method depends on two decisive factors, one is the threshold, and the other is the selection of threshold function [28 ~ 29]. In the related research on the prediction of penetration rate, most of them use fixed threshold to measure the noise [31,30,20,19,18,14], as shown in Formula 1.

$$\lambda = \sigma_n \sqrt{2\ln N} \tag{1}$$

In the formula: λ is the threshold; σ_n is the standard deviation of noise data; N is the length of noise data.

Using the same threshold to measure the noise of wavelet coefficients of different decomposition scales will inevitably reduce the denoising effect. In view of this, the modified threshold determination method [32] is introduced, as shown in Equation 2.

$$\lambda = \sigma_n \sqrt{2\ln N} / \log_2(j+1) \tag{2}$$

In the formula: j is the decomposition scale.

The corrected threshold λ will decrease with the increase of decomposition scale j, which is consistent with the propagation law of noise and ensures the effectiveness of denoising. The traditional wavelet threshold function is divided into hard threshold function and soft threshold function, as shown in Formula 3 and Formula 4 [33].

$$w^* = \begin{cases} w, & |w| \geq \lambda, \\ 0, & |w| < \lambda, \end{cases} \tag{3}$$

$$w^* = \begin{cases} [sign(w)](|w| - \lambda), & |w| \geq \lambda, \\ 0, & |w| < \lambda. \end{cases} \tag{4}$$

In the formula: w^* is the wavelet coefficients after wavelet processing; λ is the threshold value of the function ; sign ( ) is a sign function; w is the wavelet detail coefficient after wavelet decomposition.

These two function methods are widely used in the process of noise reduction of drilling data in the establishment of ROP prediction model [20, 19, 18, 14], but each has its own shortcomings. The hard threshold function sets the decomposition coefficients less than the threshold value in different scale spaces to zero, while retaining the decomposition coefficients greater than the threshold value, which will lead to non-continuity in time. This breakpoint problem will bring oscillation to the signal and ultimately reduce the denoising effect [34 ~ 36]. The soft threshold function has continuity in time, but because there will be a constant deviation between the estimated wavelet coefficients and the real wavelet coefficients in time, this method will lose singular points, and the noise reduction effect of the signal with low signal-to-noise ratio is not obvious [37]. In order to solve the above problems, a correction method is introduced to the threshold function [33], as shown in Equation 5.

$$w^* = \begin{cases} sign(w)\left\{|w| - \dfrac{\lambda}{\exp^3[\alpha(|w| - \lambda)/\lambda]}\right\}, & |w| \geq \lambda, \\ 0, & |w| < \lambda. \end{cases} \tag{5}$$

Signal-to-noise ratio (SNR) is an important basis for judging the denoising effect [38]. The greater the signal-to-noise ratio, the better the noise reduction effect [39]. Assuming that the signal data to be denoised is, the signal data after denoising is, the signal energy is, and the noise energy is, the calculation formula is as follows.

$$snr = 10\log_{10}\left(\frac{P_x}{P_y}\right) \tag{6}$$

$$P_x = \sum_{i=1}^{N} x_i^2 \tag{7}$$

$$P_y = \sum_{i=1, j=1}^{N} (x_i - y_j)^2 \tag{8}$$

2.2. PSO-LSTM Algorithm Model

Time series characteristics are hidden in drilling data, and long short-term neural network (LSTM) has great advantages in processing time series data. The long-term and short-term neural network (LSTM) is a variant of the recurrent neural network. It was proposed by Sepp Hochreiter and J¨ Sturgen Schmidhuber in 1997 [40]. It has advantages in nonlinear feature learning and can effectively avoid the gradient explosion problem of general recurrent neural networks. It has been widely used in the field of time series data prediction [41 ~ 45]. The LSTM neural network unit structure is shown in Figure 1.

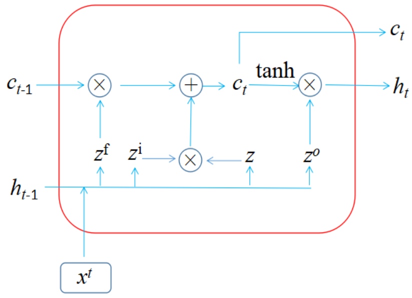

Figure 1. LSTM unit structure

In the figure, c_{t-1} and c_t represents the cell unit state at the previous moment and the current moment; h_{t-1} and h_t represents the hidden layer state of the previous moment and the current moment; x^t represents the input of the current moment; tanh represents the activation function; z^f, z^i and z^o represent the three gate structures of LSTM, which are forgetting gate, input gate and output gate, respectively. The expressions are shown in Formula 9, Formula 10 and Formula 11.

$$z^f = \sigma(w^f \times [h_{t-1}, x^t]) \tag{9}$$

$$z^i = \sigma(w^i \times [h_{t-1}, x^t]) \tag{10}$$

$$z^o = \sigma(w^o \times [h_{t-1}, x^t]) \tag{11}$$

The prediction performance of LSTM neural network is generally affected by hyperparameters, especially the number of hidden layer nodes. Too much will lead to overfitting of the model, and too little will lead to large prediction error of the model [46]. The particle swarm optimization algorithm [47] is good at finding the global optimal solution. Combining the particle swarm optimization algorithm with the number of neurons in the LSTM hidden layer can solve the problem that the LSTM neural network model cannot determine the number of hidden layer neurons, resulting in over-fitting or under-fitting. The specific steps are as follows:

(1) The position of each particle is used as the number of hidden layer neurons in the LSTM model, and the mapping result is calculated.

(2) Specify the number of network iterations, back propagation update error.

(3) The negative value-R2 of the network prediction fitting degree is taken as the fitness, and the calculation formula of

R2 is as follows.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(ROP_t - ROP_p)^2}{\sum_{i=1}^{n}(ROP_t - ROP_{t,mean})^2} \quad (12)$$

In the formula: ROP_t the true value representing the rate of penetration; ROP_p represents the predicted value of penetration rate; $ROP_{t,mean}$ represents the mean value of the true value of the penetration rate.

If the number of model iterations is reached, the solution is stopped, otherwise the position and velocity of each particle are updated, and step (1) is returned.

2.3. Box Diagram Data Processing

Due to the certain environmental interference in the drilling operation site, there are some abnormalities or missing in the acquired data. Using these data for modeling, the prediction results are biased. Cleaning the missing values and discrete values of the original drilling data is the key link to improve the reliability of the prediction results of the rate of penetration.

By checking the missing values, all the data can be converted to True and False, where the True value represents the missing value, and then the sum of all the True values is counted to obtain the number of missing values. The box plot can check the discrete values and reflect the main characteristics of the original data distribution, which is often used in various fields. The five factors in the statistical data of the box-type chart are the first quartile Q1 (the number 25 % after the value is arranged from small to large), the median Q2, the third quartile Q3 (the number 75 % after the value is arranged from small to large), the minimum value and the maximum value. The minimum value min and maximum value max are defined as follows.

$$min = Q1 - 1.5(Q3 - Q2) \quad (13)$$

$$max = Q3 + 1.5(Q3 - Q2) \quad (14)$$

The box plot regards the data less than min and greater than max as discrete values. After the statistics of missing values and discrete values, they can be processed by deletion, filling and so on.

2.4. Drilling Pressure Correction

Due to the influence of wellbore friction and well deviation, the downhole WOB is 30 % lower than the surface WOB on average [48]. If the drilling pressure measured on the ground is directly regarded as the real drilling pressure in the well to establish the prediction model, the prediction result of the drilling speed may have a large deviation. Xuyue Chen et al. [49] obtained the relationship between the bottom hole WOB and the measured WOB by analyzing the internal force of the drill string generated by the bottom hole WOB in each well section, as shown below.

$$WOB_b = WOB \bullet e^{-u\gamma_b} \quad (15)$$

In the formula: WOBb is the downhole WOB, WOB is the WOB measured on the ground, u is the friction coefficient of the wellbore, 0.25 ~ 0.4, generally 0.35 [50 ~ 51], γb is the inclination angle.

2.5. Optimization of Random Forest Feature Parameters

According to the drilling data records, the drilling parameters include hanging weight, drilling pressure, rotational speed, vertical pressure, deviation angle, flow rate, torque and well depth. If all become input parameters, it will not only increase the amount of calculation of the model, but also greatly reduce the prediction accuracy. At present, most of the drilling speed prediction models use correlation analysis method to screen drilling parameters. However, when this method is used to analyze the correlation of drilling parameters, some important factors with low correlation may be ignored due to data reasons. Therefore, it is necessary to use correlation analysis to find input parameters or delete some factors with low correlation analyzed [52].

The Random Forest (RF) algorithm is proposed by Breiman [53], which can give the importance evaluation of feature variables. It has the advantages of good effect, short training time, and is not easy to fall into the over-fitting problem. Therefore, it is widely used in various classification, prediction, feature selection and other related issues [54]. For the prediction of penetration rate, the conventional feature importance measurement method of RF mainly uses the distribution accuracy of out of bag data (OOB data) [55 ~ 56]. The importance of features can be calculated from these OOB data [57]. It is assumed that there are K training samples of drilling data, each sample has N features, and the random forest has L decision trees. According to the distribution accuracy of OOB data, the specific steps of feature importance ranking of drilling parameters are as follows:

(1) Initialize $l = 1$, create a decision tree T_l.

(2) Randomly select 2K / 3 samples to train T_l, and calculate the l decision tree OOB data distribution accuracy C_l.

(3) Randomly perturb the feature X_i, $i = 1, 2, 3.....N$ in the OOB dataset, and calculate C_{li} again.

(4) For $l = 2, 3, 4.....L$, repeat steps 2 ~ 3.

(5) The importance measure P_i of feature X_i is calculated by formula 18.

$$P_i = \frac{1}{L}\sum_{l=1}^{L}(C_l - C_{li}) \quad (16)$$

(6) The order of feature importance is obtained by sorting P_i in descending order.

In summary, the flow chart of the PSO-LSTM model established in this paper is as follows.

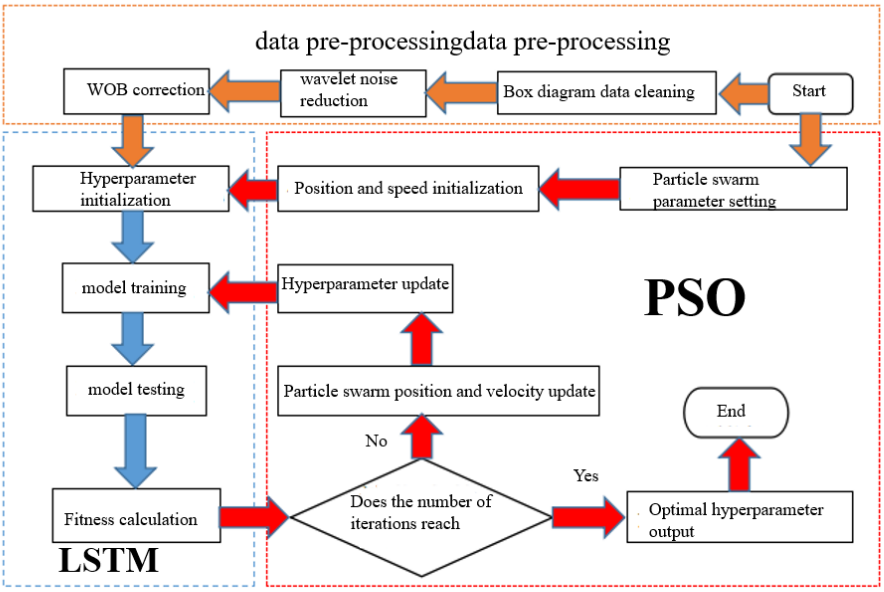

Figure 2. PSO-LSTM model establishment flow chart

3. Case Analysis

3.1. Case Well Situation

In this paper, seven deep directional wells in the central part of Bohai Bay, China are selected as data sources. The drilling data of six wells are used as the training data set of PSO-LSTM model, and the drilling data of another well are used as the test data set of PSO-LSTM model. Among them, the wellbore trajectory of the test well is type, the well depth is 3442 m, the vertical depth is 2818.85 m, the depth of the build-up point is 352.61 m, and the horizontal displacement of the bottom hole is 1802.03 m. The three-opening well structure is adopted: the first opening ϕ508.00 mm well hole is put into the ϕ660.40 mm surface casing 31 m; the ϕ444.50 mm technical casing 505 m was put into the ϕ339.70 mm borehole in the second section to seal the upper formation of basalt. The third section of ϕ139.70 mm hole is run into ϕ215.90 mm oil layer casing cementing completion.

According to the drilling record data, there are a total of 10052 sets of drilling data in seven wells. Among them, there are 8156 groups of drilling data wells 1 ~ 6 in training wells and 1896 groups of drilling data wells 7 in test wells. This paper will use this as a data source to establish a PSO-BP mechanical drilling speed prediction model.

3.2. Box Diagram Data Processing

It is found that there is no missing value in the data, but there are many discrete values in each well, as shown in Figure 3 (taking well 1 as an example).

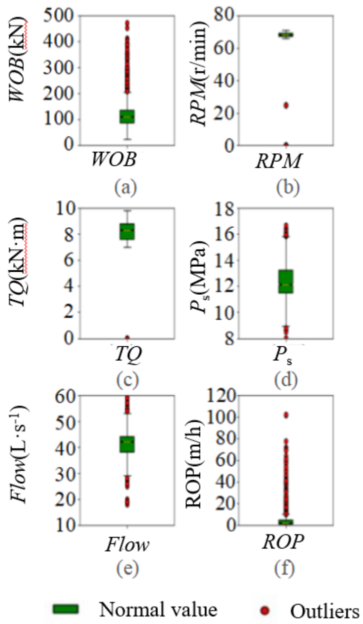

Figure 3. Discrete value analysis

In this paper, the processing of two types of outliers is processed by forward interpolation.

3.3. Wavelet De-noising

Here, the traditional soft threshold function, hard threshold function and improved noise reduction method are used to denoise the drilling data. The signal-to-noise ratio calculation results of each well are visualized, as shown in Fig.4.

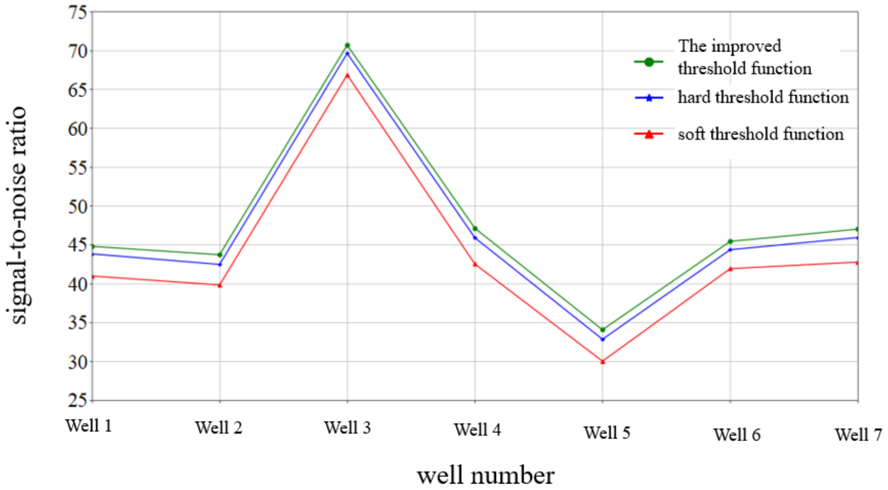

Figure 4. Signal-to-noise ratio calculation results after different threshold functions

It can be seen from Fig.4 that after using the improved noise reduction method, the signal-to-noise ratio results of each well are higher than those using the traditional threshold function method, which shows that the noise reduction effect has been improved.

3.4. Drilling Pressure Correction

The corrected WOB and correction amount are shown in Fig. 5 (taking well 1 as an example).

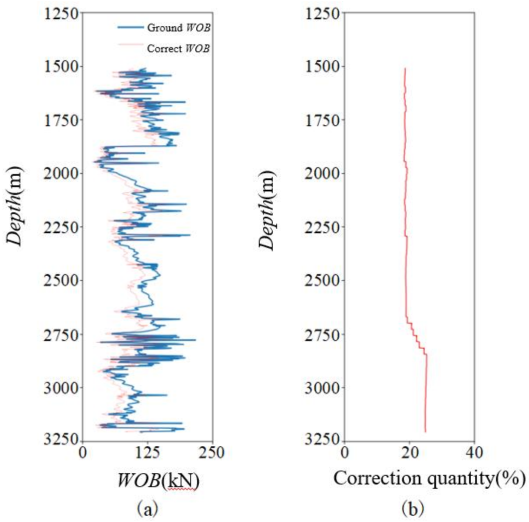

Figure 5. Drilling pressure correction and correction amount diagram

It can be seen from figure 5 (a) that the corrected WOB is generally smaller than the ground WOB; from figure 5 (b), the correction of drilling pressure is constantly changing, because the well section recorded by well 1 drilling data is an inclined section, and the deviation angle is constantly changing with the change of wellbore trajectory.

3.5. Optimization of Random Forest Feature Parameters

The importance of drilling parameters of each well is shown in figure 6.

In Figure 6, the horizontal axis represents the drilling parameters, the vertical axis represents the type of well, and the number in the table represents the degree of feature importance. The larger the value of the number, the higher the corresponding feature importance. The importance of drilling parameters was analyzed, and it was found that only the accuracy of out-of-bag data distribution of rotation speed in all wells was less than 0.1. Therefore, this paper selects well depth, suspension weight, WOB, torque, vertical pressure, flow rate and deviation angle as the input characteristic parameters of PSO-LSTM mechanical drilling speed prediction model.

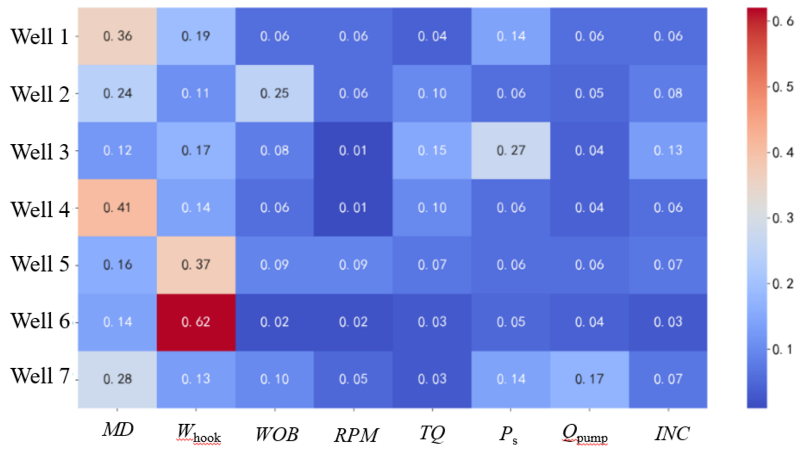

Figure 6. Importance analysis of drilling parameters

3.6. Establishment of PSO-LSTM Mechanical Drilling Speed Prediction Model

3.6.1. Modeling Parameters

The parameters of the PSO-LSTM model set in this paper are shown in table 2.

Table 2. Model building parameters

Dataset category	sample size	tier number	learning rate	number of particles	c_1	c_2	particle range	iteration times
training sets	8156	2	0.01	80	0.5	0.5	(1,50)	100
testing sets	1896	2	0.01	80	0.5	0.5	(1,50)	100

3.6.2. Experimental Results

The particle swarm search results are shown in Figure 7.

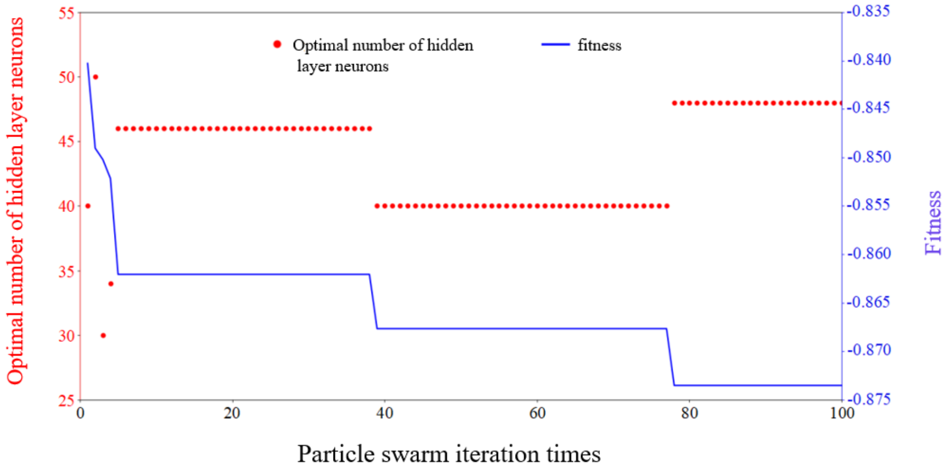

Figure 7. Particle swarm search results

It can be seen from Fig. 7 that with the increase of the number of iterations of the particle swarm, the optimal fitness of the overall particle swarm is constantly decreasing, and the corresponding optimal number of hidden layer neurons is also constantly updated. When the model reaches the lowest fitness, that is, when the R2 fitting degree is the highest currently, it is 0.873, and the corresponding optimal number of hidden layer neurons is 48.

Figure 8 shows the relative error and average error of the experimental results of the PSO-LSTM model. At the same time, the prediction results of the model without correction and the prediction results after the traditional wavelet noise reduction method are compared.

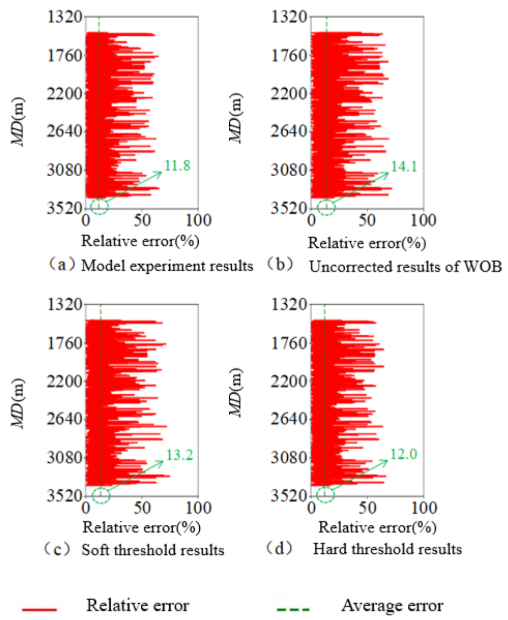

Figure 8. Relative error schematic diagram of mechanical drilling rate prediction

The average error of the model is reduced by 2.3 % after the drilling pressure is corrected. Compared with the traditional wavelet denoising method using soft threshold function and hard threshold function, the average error of the improved denoising method is reduced by 1.4 % and 0.2 % respectively.

4. Conclusion

(1) In recent years, wavelet de-noising method has been widely used in the field of ROP prediction, but most of them are qualitative, and the prediction accuracy is low. Based on the drilling field data, this paper introduces a correction method for wavelet noise reduction. It is found that the prediction accuracy of the model is improved after adjusting the threshold function and threshold determination. This is because the improved threshold function overcomes the discontinuity and deviation of the traditional wavelet threshold function.

(2) When domestic and foreign scholars model the prediction model of ROP, most of them directly model the ground data as real downhole data, and the rationality of the data will inevitably be reduced. In this paper, after modifying the surface WOB, it is found that the average error of the model is reduced by 2.3 %. This is because the model takes into account the influence of wellbore friction and well deviation, and does not directly regard the surface WOB as the real downhole WOB as the input of the model, which fully demonstrates the need to consider the relationship between downhole data and ground data.

(3) The traditional LSTM neural network is easy to fall into the problem of over-fitting or under-fitting due to the random selection of the number of hidden layer neurons. In this paper, the particle swarm optimization algorithm is used to optimize the number of hidden layer neurons in the LSTM network structure. From the experimental results, the optimized LSTM model has a higher fitting degree for the prediction of the rate of penetration than other machine learning models, and the prediction error is within an acceptable range, which can provide a certain guiding role for the engineering application of accurately predicting the rate of penetration in the drilling site.

References

[1] Liu Shengwa, Sun Junming, Gao Xiang, et al. Analysis and establishment of prediction model of ROP based on artificial neural network [J]. Computer Science, 2019, 46(S1): 605-608.

[2] Young, F. S. Computerized Drilling Control [J]. Journal of Petroleum Technology, 1969, 21(4): 483-496.

[3] Bourgoyne A T, Young F S. A Multiple Regression Approach to Optimal Drilling and Abnormal Pressure Detection [J]. Society of Petroleum Engineers Journal, 1974, 14(04):371-384.

[4] Bahari A, Baradaran Seyed A. Trust-Region Approach To Find Constants of Bourgoyne and Young Penetration Rate Model in Khangiran Iranian Gas Field [J]. Society of Petroleum Engineers, 2023.

[5] Hasan B M, Aboozar B, Hamidrezaidreza M. INTELLIGENT DRILLING RATE PREDICTOR [J]. International Journal of Innovative Computing Information and Control, 2011, 7(4): p.1511-1519.

[6] Irawan S, Rahman A M A, Tunio S Q. Optimization of Weight on Bit During Drilling Operation Based on Rate of Penetration Model [J]. Research Journal of Applied Sciences Engineering & Technology, 2012, 4(12): 1690-1695.

[7] A S S L, B K Y K, A J W S. Development of a trip time for bit exchange simulator for drilling time estimation [J]. Geothermics, 2018, 71:24-33.

[8] Carbonell J G, Michalski R S, Mitchell T M. AN OVERVIEW OF MACHINE LEARNING [J]. Machine Learning, 1983: 3-23.

[9] Amer M M, Dahab A S, El-Sayed A A H. An ROP Predictive Model in Nile Delta Area Using Artificial Neural Networks [J]. 2017.

[10] Kor K A G. Is Support Vector Regression method suitable for predicting rate of penetration? [J]. Journal of Petroleum Science & Engineering, 2020, 194(1).

[11] Chao Gan, Wei-Hua Cao, Min Wu, et al. Prediction of drilling rate of penetration (ROP) using hybrid support vector regression: A case study on the Shennongjia area, Central China [J]. Journal of Petroleum Science and Engineering, 2019, 181: 106200.

[12] Li Qi, Qu Fengtao, He Jingbin et al. Drilling rate prediction model based on BAS-BP [J]. Journal of Xi'an Shiyou University(Natural Science Edition), 2021, 36(6): 89-95.

[13] Li Qi, Qu Fengtao, He Jingbin, et al. Drilling rate prediction model based on PSO-BP [J]. Science Technology and Engineering, 2021, 21(19): 7984-7990.

[14] Hongtao L, Yan J, Xianzhi S, et al. Rate of Penetration Prediction Method for Ultra-Deep Wells Based on LSTM–FNN [J]. Applied Sciences, 2022, 12(15).

[15] Ehsan Brenjkar, Ebrahim Biniaz Delijani. Computational prediction of the drilling rate of penetration (ROP): A comparison of various machine learning approaches and traditional models [J]. Journal of Petroleum Science and Engineering, 2022, 210: 110033.

[16] Shi Xiangchao, Wang Yuming. Drilling rate prediction method based on LSTM recurrent neural network model. CN202210022755.6 [2022-04-15].

[17] Zhang Ligang, Miao Zhenhua, Huang Xiaogang, et al. Prediction of ROP based on MEA-BP neural network [J]. Automation & Instrumentation, 2022, 37(11): 87-92.

[18] Liu Ye, Zhang Fuqiang, Yang Shuopeng, et al. Time series feature characterization and prediction method of ROP based on Attention-LSTM. 202211438623[2023-09-27].

[19] Mu Huayan, Sun Jinsheng, Ding Yan, et al. Drilling rate prediction model optimization based on mechanical specific energy [J]. Drilling & Production Technology, 2023, 46(3): 16-21.

[20] Ji Hui, Zhu Liang, Lou Yishan, et al. A prediction method of penetration rate based on particle swarm optimization LSTM neural network model. CN202211353816.3 [2023-09-27].

[21] Qin Y, Mao Y, Tang B. Multicomponent decomposition by wavelet modulus maxima and synchronous detection [J]. Mechanical Systems and Signal Processing, 2017, 91(jul.): 57-80.

[22] Dai-fei, Guo, Wei-hong, et al. A Study of Wavelet Thresholding Denoising [C]// The 16th World Computer Congress in 2000. [2023-09-11].

[23] Johnstone D I M. Ideal Spatial Adaptation by Wavelet Shrinkage [J]. Biometrika, 1994, 81(3): 425-455.

[24] An Dong, Wei Yajing, Zhang Qian, et al. CEEMDAN-wavelet threshold denoising method for shafting contour reconstruction signal [J]. Tool Engineering, 2023, 57(08):153-159.

[25] Huang Lingzhi, Chen Siqi, Li Chengyu et al. Application of Gaussian process regression model based on wavelet denoising in settlement prediction of concrete face rockfill dam [J]. Journal of Water Resources and Water Engineering, 2023, 34(03): 144-150.

[26] Cai Banggui, Zhu Yunan, Wang Biao. EEMD combined with improved wavelet threshold is applied to ECG signal denoising research [J]. Modern Electronics Technique, 2023, 46(12): 137-140.

[27] Wu Xueli, Wang Zhuangzhuang, Xi Bingquan, etc. Ground penetrating radar denoising based on wavelet adaptive threshold method [J]. Science Technology and Engineering, 2023, 23(11): 4686-4692.

[28] He C, Xing J, Li J, et al. A New Wavelet Thresholding Function Based on Hyperbolic Tangent Function [J]. Mathematical Problems in Engineering, 2015, (2015-9-16), 2015, 2015 (PT.16): 528656.1-528656.10.

[29] Phinyomark A, Limsakul C, Phukpattaranont P. Optimal Wavelet Functions in Wavelet Denoising for Multifunction Myoelectric Control [J]. 2010.

[30] Zizhen Wang, Guanlin Chen, Rui Zhang, Weidong Zhou, Yitao Hu, Xunjie Zhao, Pan Wang, Early monitoring of gas kick in deepwater drilling based on ensemble learning method: A case study at South China Sea, Process Safety and Environmental Protection,Volume 169, 2023, Pages 504-514, ISSN 0957-5820.

[31] Xiao Wu. Rate of penetration prediction using hybrid Gaussian process regression model with ARD structure in geological drilling process.

[32] Jing-Yi L, Hong L, Dong Y, et al. A New Wavelet Threshold Function and Denoising Application [J]. Mathematical Problems in Engineering, 2016, (2016-5-9), 2016, 2016 (pt.5):1-8.

[33] Sun Kezhong, Lu Yingchun, Yang Xiaoxiao and so on. The application of improved wavelet threshold function in high power supply signal denoising [J]. Journal of Power Supply, 2023.

[34] Hai-Rong J, Xue-Ying Z, Jing B. A continuous differentiable wavelet threshold function for speech enhancement [J]. Journal of Central South University: English Version, 2013, 20(8): 7.

[35] Haltmeier M, Munk A. Extreme value analysis of empirical frame coefficients and implications for denoising by soft-thresholding [J]. Applied and Computational Harmonic Analysis, 2014(3).

[36] He C, Xing J C, Yang Q L. Optimal Wavelet Basis Selection for Wavelet Denoising of Structural Vibration Signal [J]. Applied Mechanics & Materials, 2014, 578-579:1059-1063.

[37] Yang Chengjin, Nie Chunyan, Wang Huiyu et al. Research and effect evaluation of EMG interference noise reduction based on wavelet improved threshold [J]. Electronic Measurement Technology, 2021, 44(22): 80-86.

[38] Tang J Y, Chen W T, Chen S Y, et al. Wavelet-based vibration signal denoising with a new adaptive thresholding function [J]. Journal of Vibration & Shock, 2009, 28(7):118-121.

[39] Zhu Yankang, Xu Gangqiang, Zhou Kangkang, etc. Optimization of camshaft casting process parameters based on signal-to-noise ratio and TOPSIS [J]. Foundry, 2023,72(07): 909-916.

[40] Sepp Hochreiter, Jürgen Schmidhuber. Long Short-Term Memory. *Neural Comput* 1997, 9 (8): 1735–1780.

[41] Zhifeng Ma, Hao Zhang, Jie Liu, MS-LSTM: Exploring spatiotemporal multiscale representations in video prediction domain, Applied Soft Computing, Volume 147, 2023, 110731, ISSN 1568-4946.

[42] Vamshikrishna Domala, Tae-wan Kim, Application of Empirical Mode Decomposition and Hodrick Prescot filter for the prediction single step and multistep significant wave height with LSTM, Ocean Engineering,Volume 285, Part 1, 2023, 115229, ISSN 0029-8018.

[43] Shaowei Pan, Bo Yang, Shukai Wang, Zhi Guo, Lin Wang, Jinhua Liu, Siyu Wu, Oil well production prediction based on CNN-LSTM model with self-attention mechanism, Energy, Volume 284, 2023, 128701, ISSN 0360-5442.

[44] Kai Wang, Yu Hua, Lianzhong Huang, Xin Guo, Xing Liu, Zhongmin Ma, Ranqi Ma, Xiaoli Jiang,A novel GA-LSTM-based prediction method of ship energy usage based on the characteristics analysis of operational data, Energy, 2023, 128910, ISSN 0360-5442.

[45] Qing Kang, Elton J. Chen, Zhong-Chao Li, Han-Bin Luo, Yong Liu, Attention-based LSTM predictive model for the attitude and position of shield machine in tunneling, Underground Space, Volume 13, 2023, Pages 335-350, ISSN 2467-9674.

[46] Liu Dong, Luo Qi, Tang Lei, etc. The multi-dimensional neutronics diffusion equation is solved based on PINN deep machine learning technology [J]. Nuclear Power Engineering, 2022, 43(02):1-8.

[47] Kennedy J, Eberhart R. Particle swarm optimization [C]// IEEE International Conference on Neural Networks, 2002.

[48] Kaveh Shaygan, Saeid Jamshidi. Prediction of rate of penetration in directional drilling using data mining techniques [J]. Geoenergy Science and Engineering, 2023, 221: 111293.

[49] Xuyue Chen, Honghai Fan, Boyun Guo, et al. Real-time prediction and optimization of drilling performance based on a new mechanical specific energy model [J]. Arabian Journal for Science and Engineering, 2014, 39: 1-11.

[50] Johansick, C.A.; Friesen, D.B.; Dawson, R.et al.: Torque anddrag in directional wells-prediction and measurement. J. PetroleumTechnol. 36(6), 987–992(1984), Li, J.et al.: Research on torque and drag in extended-reach hori-zontal wells and its application in Chenghai-1 area. Oil Drill. Prod. Technol. 31(3), 21–25(2009)

[51] Aadnoy, B.S.: Theory and Application of a New Generalized Modelfor Torque and Drag. IADC/SPE 114684(2008).

[52] Shi Xiangchao, Wang Yuming, Liu Yuehao, etc. Discussion on the application of artificial intelligence method to the prediction of drilling rate [J]. Oil Drilling & Production Technology, 2022, 44(01):105-111.

[53] Reiman L. Random forest [J]. Machine Learning, 2001, 45:5-32.

[54] Yao Dengju, Yang Jing, Zhan Xiaojuan. Feature selection algorithm based on random forest [J]. Journal of Jilin University (Engineering and Technology Edition), 2014, 44(1):137-141.

[55] ltmann André, Toloşi Laura, Oliver S, et al. Permutation importance: a corrected feature importance measure [J]. Bioinformatics(10): 1340 [2023-09-11].

[56] Strobl C, Boulesteix A L, Kneib T, et al. Conditional variable importance for random forests [J]. Bmc Bioinformatics, 2008, 9(1):307-307.

[57] Neave H R, Journal o S S S C, Chandler R, et al. Understanding Randomness [J]. Journal of the Royal Statistical Society Series C [2023-09-11].