
Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 5, pp. 4218–4227, DOI 10.12694/scpe.v25i5.3070

PREDICTION METHOD OF RATE OF PENETRATION BASED ON FUZZY SUPPORT VECTOR REGRESSION

LI YANG[*], LISHEN WANG[†], LILI BAI[‡], AND WENFENG SUN[§]

Abstract. Predicting the rate of penetration (ROP) is important for optimizing drilling parameters, improving drilling efficiency, and optimizing economic benefits throughout the drilling process. The current prediction model of ROP based on machine learning algorithms does not consider the interference of outliers. Therefore, in this study, we propose a method to predict ROP based on fuzzy support vector regression (FSVR). First, appropriate input parameters were selected from the controllable parameters. Second, based on the local outlier factor, a fuzzy membership degree was assigned to each sample. Finally, the sample with the fuzzy membership value was input into the model for ROP prediction. The results demonstrated that the goodness of fit (R2) of the improved FSVR model is 0.9634, and the mean absolute error is 0.1974. Compared with standard SVR and other models, the improved FSVR model has a stronger anti-interference ability, smaller prediction error for normal samples, and higher accuracy.

Key words: drilling; rate of penetration; support vector regression; local outlier factor; fuzzy membership.

1. Introduction. The rate of penetration (ROP) is an important factor affecting drilling duration and drilling cost. Accurate prediction of ROP can provide support for drilling parameter optimization, thus enhancing drilling efficiency while minimizing drilling expenses. There are three types of ROP prediction models [1]:traditional mathematical equations, which consider a few small numbers of ROP influencing parameters [2], statistical models represented by multiple regression, which can have increased solving accuracy for the traditional ROP coefficient [3]; and models employing machine learning algorithms. Wu et al [4] established a linear prediction model using principal component analysis (PCA); however, in some cases, the element obtained by this method is not optimal. Su et al. [5] applied a gradient boosting decision tree algorithm to predict ROP, which is more accurate than the existing method of predicting ROP; however, the characteristics involved were limited. Kahraman et al. [6] proposed an ROP prediction model based on backpropagation (BP) neural networks; however, BP neural networks have low convergence speed and can easily fall into local extrema. Yang et al. [7] established a fuzzy neural network model by combining fuzzy control with a BP neural network, and the level of prediction accuracy was higher than that of ordinary BP neural networks. Zhao et al. [8] proposed an ROP prediction model based on extreme learning machines, achieving better generalization performance than the common BP neural network. Li et al. [9] combined the beetle antennae search (BAS) algorithm with a BP neural network to establish the BAS-BP model. Compared with standard BP and genetic-algorithm BP models, the error is lower and the prediction accuracy is higher; however, the generalization ability of the model should be improved. Xu et al. [10] established an ROP prediction model based on integrated learning, utilizing a variety of machine learning algorithms, and the results show that the prediction accuracy of the integrated model is higher than that of the single model; however, the model is more complex. Sabah et al. [11] used a multilayer perceptron particle swarm optimization (PSO-MLP) model for ROP prediction, achieving favourable performance. As a representative of traditional machine learning, support vector regression (SVR) has been demonstrated to have fitting performance compared to alternative machine learning methods(e.g., k-nearest neighbours, linear regression, polynomial regression, and decision trees) in experiments involving a drilling process that is characterized by high nonlinearity and complexity [12, 13]. The above two models are

[*]College of Electrical Information Engineering, Northeast Petroleum University, Daqing 163318, China
[†]College of Electrical Information Engineering, Northeast Petroleum University, Daqing 163318, China
[‡]Shanghai Technical Institute of Electronics & Information, STIEI Shanghai,201411, China
[§]Shanghai Technical Institute of Electronics & Information, STIEI Shanghai,201411, China (Corresponding author, 380780857@qq.com)

compared with the proposed model. Hamid et al. [14] used an imperialist competitive algorithm to optimize SVR. They established the model, compared three kernel functions, and found that the Gaussian radial basis kernel function had the best effect. Therefore, the Gaussian radial basis kernel function was also applied in this study.

Recently, entropy-based improved support vector machine models have been proposed to solve the classification problem [15, 16, 17]. Asadolahi et al. [18] and Xue et al. [19] improved the loss function by combining the fuzzy response with robust SVR to improve the accuracy of the fuzzy regression model. Chakravarty et al. [20] and Liu et al. [21] improved the fuzzy function and the robustness of the model against outliers and verified its effectiveness. Fuzzy control has been continuously applied to complex system processing [22]. Successful applications have been demonstrated in wind turbine control [23], concrete component analysis [24], and life prediction of slewing bearings [25]. Nevertheless, applications in drilling engineering remain scarce.

During drilling operations, signals are disturbed to produce outliers or slip drilling and other situations occasionally occur in drilling may lead to sudden increase in the values of certain parameters,which interferes with the training of the model and consequently affects its performance. A method of elimination has been adopted to process outlier data; however, it may delete misjudged data. When processing new data sets, there will be outliers of new data sets, which should be eliminated again. Therefore, we focused on reducing the influence of abnormal data on model prediction without deleting primary data. Current ROP prediction models based on machine learning algorithms do not consider interference from outlier samples; however, the fuzzy support vector regression (FSVR) model proposed in this study considers such interference. According to the outlier detection method based on the local outlier factor (LOF) algorithm, the LOF values of all data points are calculated, and then the membership values of all data points are calculated through the LOF values. A normal sample has a large membership value, and an outlier sample has a small membership value. The fuzzy training set with membership values is input into the training model. This can reduce the interference of outlier samples in the model, and the normal samples with a large "contribution" to the training model can play the greatest role. We demonstrated how this model has stronger anti-interference ability and higher prediction accuracy than other models.

2. Method.

2.1. Calculation of local outlier factor values. Outlier detection based on LOF is a density-based outlier detection method; it works on the principle that each data point is assigned a LOF that depends on neighbourhood density. For any data point in a particular dataset, if the number of surrounding points is high and its local neighbourhood is dense, the data point is considered as a normal data point. Contrarily, an outlier is a data point that is far from the nearest neighbour of a normal data point. Thus, LOF can be understood as the outlier degree of a sample relative to its neighbours.

Among the k nearest neighbours to P, the distance between the farthest neighbour and P is the k-distance of P. The k-distance neighbour of sample P is the neighbour set whose distance from P is not greater than the k-distance of P. It is defined as the k-distance neighbour set $N_k(P)$, provided $N_k(P)$ by equation (1), where P' is the neighbour of P and D is the sample set. $d(P, P')$ is the distance between P' and P, and $d_k(P)$ is the k-distance of P. Furthermore, $d_k(P, O)$ is the k-th reachable distance from sample O to sample P. It corresponds to the largest distances among the k-distance from O and the distance from P to O, and it is provided by the expression in equation (2), where $d_k(O)$ is the k-distance of O and $d(P, O)$ is the distance between P and O.

$$N_k(P) = \{P' | P' \in D \backslash \{P\} | d(P, P') \leq d_k(P)\} \tag{1}$$

$$d_k(P, O) = max\{d_k(O), d(P, O)\} \tag{2}$$

The neighbours within a sample and the density relationship between the sample and its neighbours can be described by the concept of local reachable density. The local reachable density of a sample P is the number of elements in the k-nearest neighbour set of P divided by the sum of the relative reachable distances from all points within the data set to P, which is expressed as $\rho_k(P)$. The expression is as follows:

$$\rho_k(P) = \frac{|N_k(P)|}{\sum_{o \in N_k(P)} d_k(P, O)} \tag{3}$$

The LOF is obtained by calculating the ratio of the average local reachable density of k points around the sample point to the local reachable density of this point, which is expressed by equation (4).

$$lof_k(P) = \frac{\sum_{o \in N_k(P)} \frac{\rho_k(O)}{\rho_k(P)}}{|N_k(P)|} \tag{4}$$

2.2. Fuzzy support vector regression. To improve the anti-interference ability of SVR, FSVR is obtained by combining the fuzzy membership value with SVR. Several sample points are required to train the model; however, the contribution of different sample points to the final result is different. Therefore, we assign a fuzzy membership value, λ,to each sample based on the LOF to obtain the fuzzy training set $D' = \{(x_1, y_1, \lambda_1), (x_2, y_2, \lambda_2), \ldots, (x_m, y_m, \lambda_m)\}, y_i \in \mathbb{R}, 0 \leq \lambda \leq 1$, where the specific setting method for λ_m is described in the subsection "Calculating the value of fuzzy membership" in Section 3 of this paper. The mathematical expression of the model is provided by equation (5). The fuzzy membership degree λ is the contribution of the sample point to the final regression results, that is, a larger λ is assigned to a normal sample or a sample with a larger contribution to the model training, and a smaller λ is assigned to an abnormal outlier data point or a sample with a smaller contribution to the final result. This can reduce the interference of some of the sample's outlier points or meaningless samples on the model to ensure that the final prediction results are closer to the true value, resulting in increased robustness of the model.

$$min\{\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} \lambda_m(\xi_i + \xi_i^*)\}$$

$$s.t. \begin{cases} y_i - f(x_i) \leq \varepsilon + \xi_i \\ f(x_i) - y_i \leq \varepsilon + \xi^* \\ \xi_i, \xi_i^* \geq 0, i = 1, 2, \ldots, n. \end{cases} \tag{5}$$

The above formula is an optimization problem with constraints, where ε is an insensitive loss function; ξ_i and ξ_i^* are slack variables, which describe the loss caused by the sample away from the regression interval; and C is the penalty factor, which refers to the penalty for the sample whose regression function error is greater than ε. The smaller the value of C, the smaller the loss of the sample and, consequently, the smaller the loss of the objective function. Points far from the interval have larger errors; therefore, the penalty factor C will make the regression function more sensitive to such points to further fit these points. To calculate the corresponding loss, the absolute value should be greater than ε. In particular, considering $f(x)$ as the center, we set an interval band width of 2ε. When the training sample falls into the interval band, the prediction result is correct.

The Lagrange multiplier $a_i, a_i^*, \mu_i, \mu_i^*$ is introduced to obtain the Lagrange equation, and the partial derivative of the equation with respect to w, b, ξ_i, ξ_i^* is obtained, such that the partial derivative is 0. The results of partial derivatives are returned to the Lagrange equation, and the equation of the dual problem is obtained as follows:

$$max \sum_{i=1}^{n} y_i(a_i^* - a_i) - \varepsilon(a_i^* + a_i) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i^* - a_i)(a_j^* - a_j)\kappa(x_i, x_j)$$

$$s.t. \begin{cases} \sum_{i=1}^{n}(a_i^* - a_i) = 0 \\ 0 \leq a_i, a_i^* \leq C\lambda_m \end{cases} \tag{6}$$

When the above process satisfies the Karush–Kuhn–Tucker conditions, the value of b is obtained as follows:

$$b = y_i + \varepsilon - \sum\nolimits_{i=1}^{n}(a_j^* - a_j)\kappa(x_i, x_j) \tag{7}$$

After feature mapping, w is expressed as follows:

$$w = \sum\nolimits_{i=1}^{n} \phi(x_i)(a_i^* - a_i) \tag{8}$$

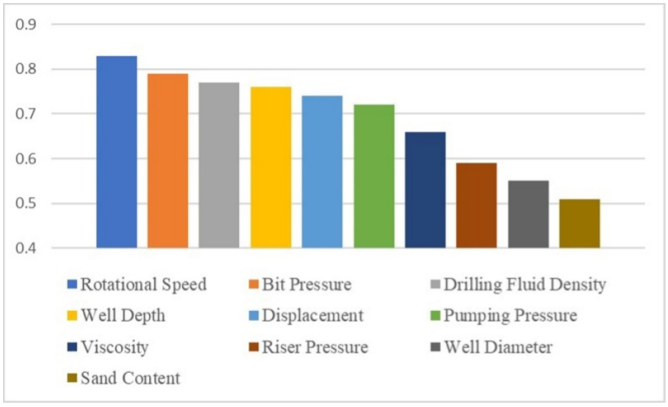


Fig. 3.1: Ranking of grey correlation degree

Table 3.1: Partial training sample - Part 1

rotational speed $/(r \cdot min^{-1})$	bit pressure /KN	density$/(g \cdot cm^{-3})$	well depth$/m$
120	20	1.10	214.10
160	30	1.14	365.03
176	50	1.19	1733.02
176	60	1.20	1876.11

By substituting w into $f(x) = w^T\phi(x) + b$, the final expression in equation (9) is obtained, where κ is the kernel function. Through this, the sample points can be mapped from low-dimensional space to high-dimensional space, which reduces the calculation difficulty. This model uses the Gaussian radial basis kernel function, which can better deal with high-dimensional samples and involves fewer parameters, consequently gaining good stability.

$$f(x) = \sum_{i=1}^{n}(a_i^* - a_i)\kappa(x, x_i) + b \tag{9}$$

3. Experiments and results.

3.1. Analysis of influencing factors on ROP. The parameters related to mechanical drilling speed can be divided into controllable and uncontrollable. Uncontrollable parameters are objective parameters related to formation conditions, such as pore pressure and fracture pressure. Controllable parameters are parameters that can be adjusted by technology and equipment and can be further divided into drilling fluid parameters, hydraulic parameters, and mechanical parameters. Drilling fluid parameters include density, rheological parameters, etc. Hydraulic parameters include bit pressure drop, etc. Mechanical parameters include drilling pressure, rotational speed, etc.

The data used in the model are from a block in the Shunbei Oilfield.There are a total of 11173 samples in the dataset.The mechanical drilling speed is assumed as the parent sequence, and its influencing parameters are assumed as the subsequence. The correlation degree ranking of the influencing parameters of ROP is obtained by grey correlation analysis and calculation, as shown in Figure 3.1.

Figure 3.1 shows that the correlation degree between each parameter and mechanical drilling rate is significantly different. The first seven parameters exhibiting a high correlation degree, namely, rotational speed, bit pressure, drilling fluid density, well depth, displacement, pumping pressure, and viscosity, are assumed to be the input of the model. Some training samples are shown in Table 3.1.

3.2. Preprocessing of data.

Table 3.2: Partial training sample - Part 2

displacement/$(L \cdot s^{-1})$	pumping pressure/Mpa	viscosity/m	ROP/$(m \cdot h^{-1})$
35	6	52	2.16
38	11	57	2.01
35	13	60	1.30
35	13	56	0.72

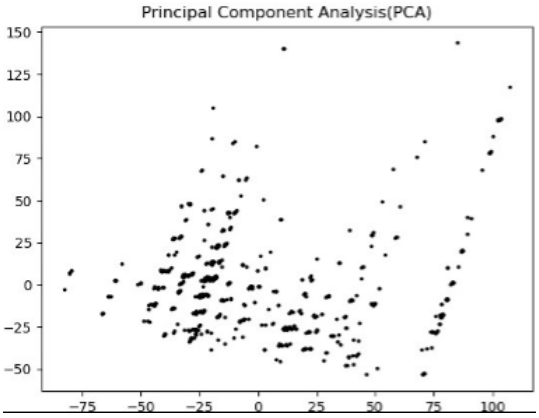


Fig. 3.2: Visualization results of 2D data

3.2.1. Detecting outliers. Owing to the introduction of the concept of nearest neighbours, the method based on LOF can better identify outliers located at the edge of the sample class and has better recognition effect. To display the results of LOF outlier detection conveniently, PCA dimension reduction was used to reduce the multidimensional data to 2D for visualization. The results are shown in Figure3.2.

The model calculates the LOF for each data point and determines whether the point is an outlier based on the size of the LOF value. According to its definition, if the LOF value is close to 1, it indicates that the sample density is similar to its neighbourhood density, and the sample is classified as a normal data point. It is believed that a LOF value between 1 and 1.5 corresponds to normal data. If the LOF value is significantly greater than 1, it indicates that the sample density is less than its neighbourhood density, and the sample may be classified as an outlier. The 2D data obtained after dimensionality reduction are used for outlier detection, and the results are shown in Figure 3.3, where the red circles mark each data point. The size of the red circle represents the size of the LOF value, that is, the radius of the red circle determines the outlier degree of the data point. The larger the radius of the red circle, the lower the sample density; hence, the sample may be outliers. A smaller radius of the red circle indicates that the point is within the same cluster as the surrounding neighbourhood points and the degree of outliers is low, which indicates a normal data point. The outliers were labelled -1 and the normal points were labelled 1.

3.2.2. Normalizing the data. To eliminate the influence of different parameter ranges on the model and improve the generalization ability of the model data processing, the data should be normalized. The data normalization interval is [-1,1], and the expression is as follows:

$$x_{norm} = \frac{(y_{max} - y_{min})(x - x_{min})}{x_{max} - x_{min}} + y_{min} \tag{10}$$

where x_{norm} and x represent the values after and before data normalization, respectively; x_{max} and x_{min} represent the maximum and minimum values before data normalization; and y_{max} and y_{min} represent the maximum and minimum values after normalization.

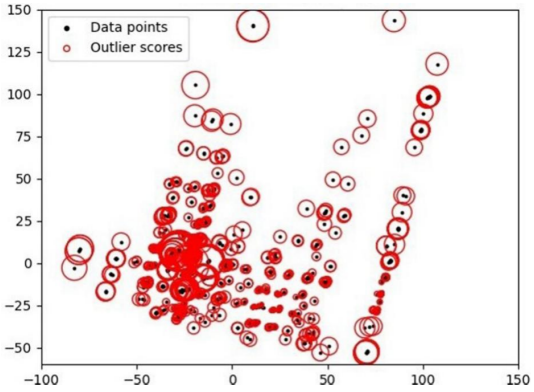


Fig. 3.3: LOF value of sample point

Table 3.3: Partial prediction results of SVR.

prediction of ROP for SVR	actual value of ROP	error	Is it a outlier (SVR)	LOF	membership value	Is it a outlier (LOF)
3.09	3.20	0.11	1	1.0251	0.8996	1
5.26	8.37	3.11	-1	1.7096	0.6203	-1
7.04	10.12	3.08	-1	1.8139	0.6059	-1
6.21	9.41	3.20	-1	1.7325	0.6147	-1

3.3. Training the fuzzy support vector regression model.

3.3.1. Training the standard SVR model. Normal training classifies data sets into training and test sets, which renders data sets not fully applicable to training. This can be avoided by cross-validation, wherein it is possible to use all data for training and testing. This model uses ten-fold cross-validation, wherein the data set is divided into ten sets, of which nine are training sets and one is a test set, which is used for an experiment. Subsequently, a new test set is chosen from the training set, and the original test set is used for training. Thus, ten experiments were conducted. Finally, the model with the smallest average absolute error in the ten experiments was assumed as the final model, and the average value of the ten average absolute errors was used to evaluate the model.

To verify the accuracy of the labelled outliers in the subsection "Detecting outliers," the standard SVR model was used for ROP prediction. The data set whose absolute difference between the predicted ROP value and real value was greater than 3, which was temporarily recorded as an abnormal outlier and marked as -1, while other normal data were marked as 1. Some prediction results are listed in Table 3.3.

There are 81 groups of abnormal outlier data marked by the standard SVR model. Compared with the 72 groups of outlier data detected using the LOF, there are 61 groups of data consistent between both models, thereby indicating that the LOF outlier detection recognition rate reaches 84.72%.

3.3.2. Calculating the value of fuzzy membership. The membership function value λ_m in this study is determined by the value of the LOF, expressed as

$$\lambda_m = \begin{cases} (1-\theta)^{\mu} + \sigma & \overline{lof} < lof_k(p) \le lof_{max} \\ 1-\theta & lof_{min} \le lof_k(p) \le \overline{lof}' \end{cases}, \theta = \frac{lof_k(p) - lof_{min}}{lof_{max} - lof_{min}} \tag{11}$$

Here, $\mu \ge 2$ (in this study, n is taken to be 10); σ is a sufficiently small positive real number less than 1; and lof_{min},lof_{max}, and $\overline{lof}$ are the minimum, maximum, and average values in the sample LOF, respectively.

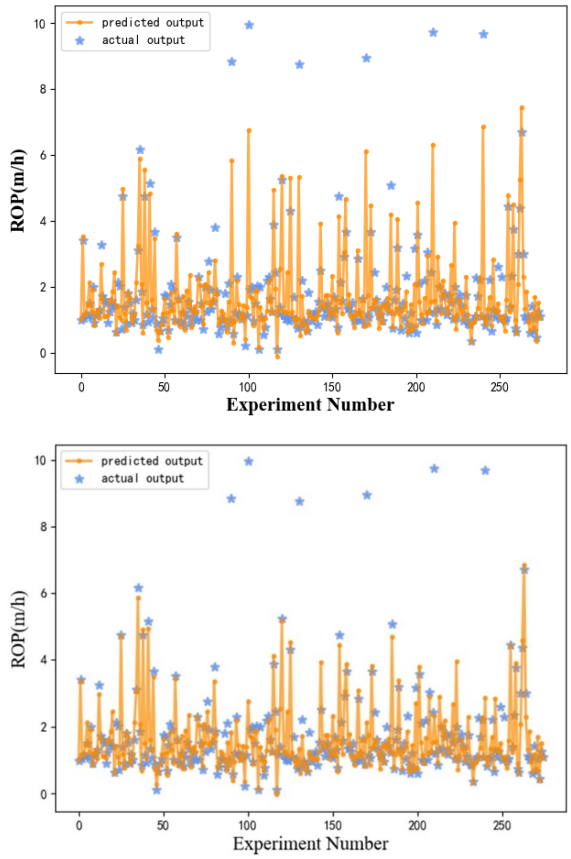


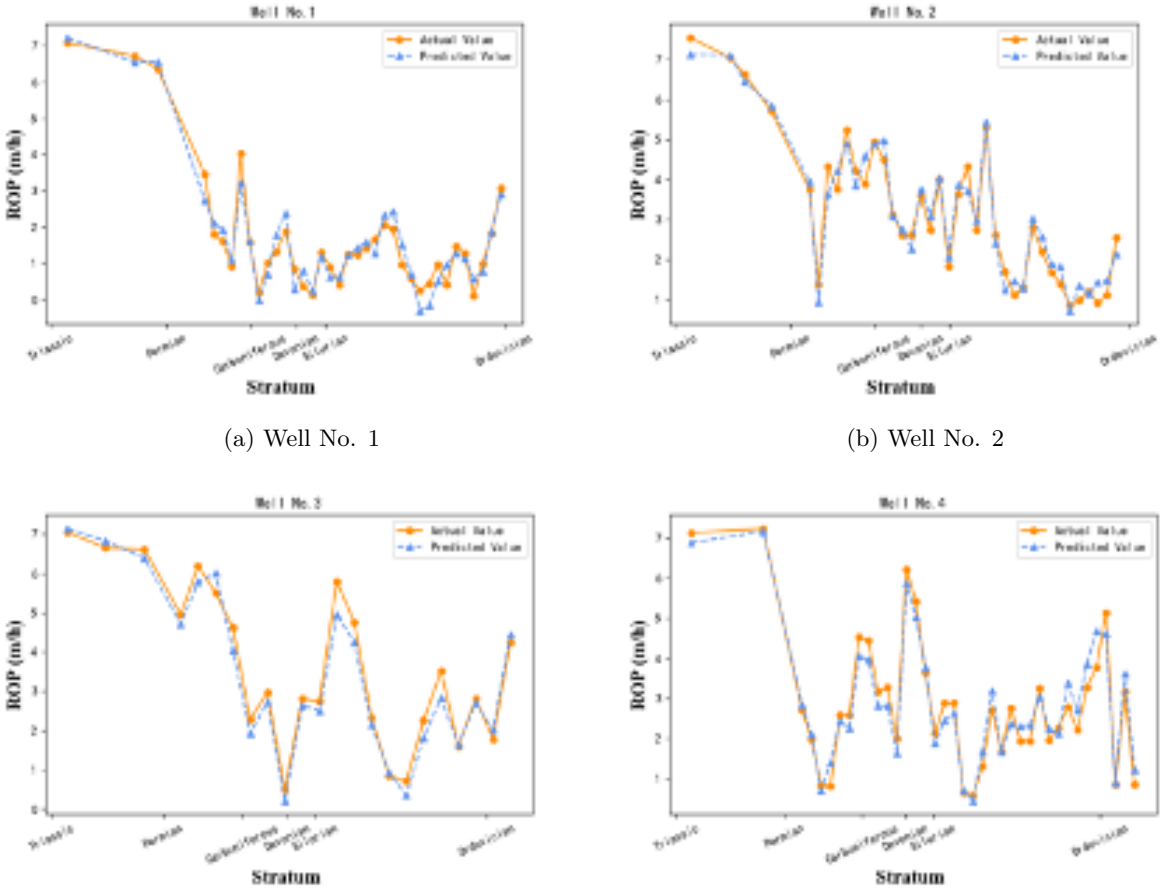
Fig. 3.4: (a) Results before model improvement (b) Results after model improvement

3.3.3. Training the fuzzy support vector regression model. The calculated fuzzy membership value is attached to each data point, and the fuzzy data set $D' = \{(x_1, y_1, \lambda_1), (x_2, y_2, \lambda_2), \ldots, (x_m, y_m, \lambda_m)\}$ is sent to the FSVR model for training. The purpose is to control the samples with a large LOF and reduce their membership value to ensure that they have a smaller influence on the model. For the samples with a small LOF, the membership value is large to ensure that they have a significant contribution to the model. This training still uses ten-fold cross-validation. The results before and after improvement are shown in Figures 4 and 5, respectively. It can be observed that for the outliers above the general trend of the data, the standard SVR model (Figure 3.4) shifts the outliers significantly; thus, it is prone to overfitting and reduces generalization ability. However, the FSVR model with the fuzzy membership value added (Figure 5) is less affected by outliers and does not shift significantly toward them, thus demonstrating stronger anti-interference ability.

The four subplots within Figure 3.5 are the prediction results of ROP in different formations of four wells within the Shunbei Oilfield using the improved FSVR model. Figures 3.5(a)–(d) utilize well Nos. 1–4, respectively. The abscissa shows different strata, and the ordinate is ROP. It can be observed from the figures that the error between the predicted and real values is controlled within 1, and the prediction effect is good.

3.4. Contrast experiment.

To evaluate the prediction accuracy of the model for normal data, the abnormal outliers of the dataset were removed, and three models mentioned in the Introduction section— ordinary SVR, BP neural network, and PSO-MLP—were used, in addition to the proposed FSVR model, to predict ROP. The errors of these four models were compared, and the results are listed in Table 3.4. It can be observed from Table 3.4 that the



(a) Well No. 1 (b) Well No. 2

(c) Well No. 3 (d) Well No. 4

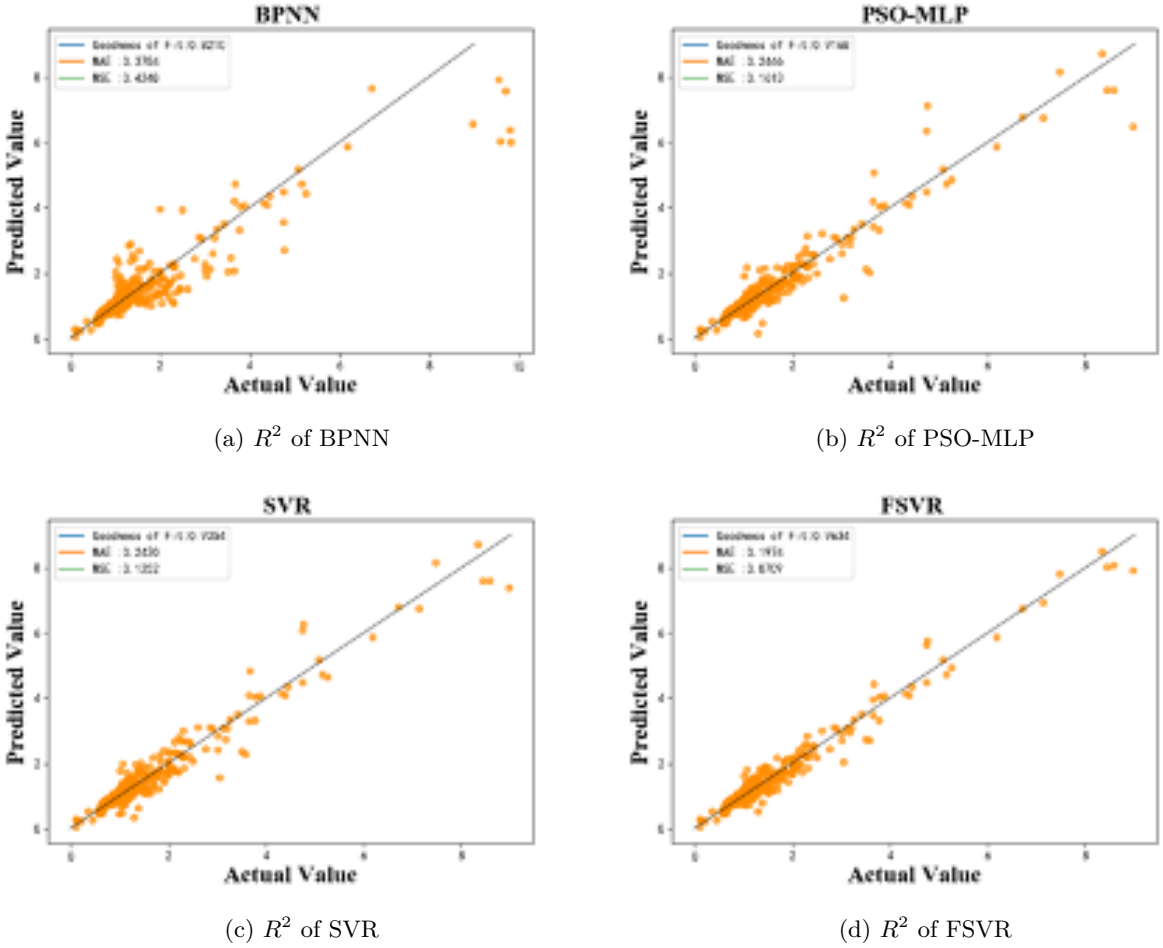
Fig. 3.5: Prediction results of FSVR

Table 3.4: Comparison of errors.

model	R^2	MAE	MSE
FSVR	0.9634	0.1974	0.0709
SVR	0.9354	0.2430	0.1252
PSO-MLP	0.9168	0.2446	0.1613
BPNN	0.8210	0.3784	0.4248

mean absolute error (MAE) and mean squared error (MSE) of the FSVR model are smaller than those of the other models. The four graphs in Figure 3.6 show the goodness of fit (R^2) of the four models; R^2 of the FSVR model is closer to 1, thereby indicating that the predicted value of the FSVR model is closer to the real value; therefore, the regression fitting effect is enhanced.

According to the comparative experiments, the improved FSVR model is not sensitive to abnormal outlier data, which can reduce the interference of outliers within the model. Additionally, the model has higher prediction accuracy for normal data points.



(a) R^2 of BPNN

(b) R^2 of PSO-MLP

(c) R^2 of SVR

(d) R^2 of FSVR

Fig. 3.6: Graphs showing goodness of fit

4. Conclusions. In this study, the FSVR model was established to address the problem of interference of outlier samples in a model, which are not considered by ROP prediction models based on machine learning algorithms. Parameters such as rotational speed and bit pressure of a block in the Shunbei Oilfield were selected as the input through grey correlation analysis. According to the LOF, a fuzzy membership value was assigned to each group of samples and was used to weight the penalty factor. The model was subsequently trained by the fuzzy training set. Attaching a smaller membership value to the outliers reduces the interference induced in the model. The normal data points were assigned a larger membership value to ensure that they have a greater influence. The results demonstrated the following.

1. The improved FSVR model has better anti-interference ability than other models, such as standard SVR, and is not prone to overfitting.
2. R2 of the improved FSVR model is 0.9634, and the MAE is 0.1974. Therefore, the improved FSVR model has a smaller prediction error, higher accuracy, and better regression fitting effect for ROP of normal sample points.
3. The model has a good ability to predict ROP, which can provide support for the optimization of drilling parameters in the next step, thereby improving drilling efficiency and saving drilling costs.

Acknowledgments. The authors are grateful to High-level and scarce talent research Initiation Foundation of STIEI under Grant No. GCC2023007 and No. GCC2023034, and the National Natural Science Foundation of China under Grant No. 51974090.

REFERENCES

[1] L.F.F.M Barbosa, A. Nascimento, M.H. Mathias et al. (2019) "Machine learning methods applied to drilling rate of penetration prediction and optimization-A review", *Journal of Petroleum Science and Engineering*, 183, 106332.
[2] C. Soares and K. Gray (2019) "Real-time predictive capabilities of analytical and machine learning rate of penetration (ROP) models", *Journal of Petroleum Science and Engineering*, 172, 934-959.
[3] C.S. Li (2013) "Research on drilling speed prediction method based on multiple regression analysis", *Science Technology and Engineering*, 13(7), 1740-1744.
[4] C.G. Wu, M. Zhao and Z.W. Guo (2015) "Study on drilling rate prediction based on principal component analysis", *Mining Research and Development*, 35(10), 84-86.
[5] X.H. Su, J.M. Sun, X. Gao et al. (2019) "Prediction method of drilling rate of penetration based on GBDT algorithm", *Computer Applications and Software*, 36(12), 87-92.
[6] S. Kahraman (2016) "Estimating the penetration rate in diamond drilling in laboratory works using the regression and artificial neural network analysis", *Neural Processing Letters*, 43(2), 523–535.
[7] L. Yang, T.Y. Liu, W.J. Ren et al. (2021) "Fuzzy Neural Network for Studying Coupling between Drilling Parameters", *ACS Omega*, 6(38), 24351-24361.
[8] Y. Zhao, T. Sun, J. Yang et al. (2019) "Extreme learning machine-based offshore drilling ROP monitoring and real-time optimization", *China Offshore Oil and Gas*, 31(6), 138-142.
[9] Q. Li, F.T. Qu, J.B. He et al. (2021) "Prediction model of mechanical ROP during drilling based on BAS-BP", *Journal of Xi'an Shiyou University (Natural Science Edition)*, 36(6), 89-95.
[10] M.Z. Xu, M.H. Wei, S. Deng et al. (2021) "Application of multi-model ensemble learning in prediction of mechanical drilling rate", *Computer Science*, 48, 619-622.
[11] M. Sabah, M. Talebkeikhah, D.A. Wood et al. (2019) "A machine learning approach to predict drilling rate using petrophysical and mud logging data", *Earth Science Informatics*, 12(3), 319–339.
[12] Y. Zhou, X. Chen, H. Zhao et al. (2021) "A novel rate of penetration prediction model with identified condition for the complex geological drilling process", *Journal of Process Control*, 100, 30–40.
[13] O.S. Ahmed, A.A. Adeniran, and A. Samsuri (2019) "Computational intelligence based prediction of drilling rate of penetration: A comparative study", *Journal of Petroleum Science and Engineering*, 172, 1-12.
[14] R.A. Hamid, J.S.H. Mohammad, and A. Masoud (2017) "Drilling rate of penetration prediction through committee support vector regression based on imperialist competitive algorithm", *Carbonates and Evaporites*, 32(2), 205-213.
[15] S. Moslemnejad and J. Hamidzadeh (2021) "Weighted support vector machine using fuzzy rough set theory", *Soft Computing*, 25(13), 8461-8481.
[16] D. Gupta and B. Richhariya (2018) "Entropy based fuzzy least squares twin support vector machine for class imbalance learning", *Applied Intelligence*, 48(11), 4212-4231.
[17] S. Chen, J. Cao, and F. Chen (2020) "Entropy-based fuzzy least squares twin support vector machine for pattern classification", *Neural Processing Letters*, Vol. 51(11), 41-66.
[18] M. Asadolahi, M.G. Akbari, and G. Hesamian (2021) "A robust support vector regression with exact predictors and fuzzy responses", *International Journal of Approximate Reasoning*, 132, 206-225.
[19] Z. Xue, R. Zhang, and C. Qin (2020) "An adaptive twin support vector regression machine based on rough and fuzzy set theories", *Neural Computing and Applications*, Vol. 32(9), 4709-4732.
[20] S. Chakravarty, H. Demirhan, and F. Baser (2020) "Fuzzy regression functions with a noise cluster and the impact of outliers on mainstream machine learning methods in the regression setting", *Applied Soft Computing*, 96, 106535.
[21] J. Liu (2021) "Fuzzy support vector machine for imbalanced data with borderline noise", *Fuzzy Sets and Systems*, 413, 64-73.
[22] A.T. Nguyen, T. Taniguchi, and L. Eciolaza (2019) "Fuzzy control systems: Past, present and future", *IEEE Computational Intelligence Magazine*, 14(1), 56-68.
[23] T. Jeon and I. Paek (2021) "Design and verification of the LQR controller based on fuzzy logic for large wind turbine", *Energies*, 14(1), 230.
[24] Z. Fan, R. Chiong, and Z. Hu (2020) "A fuzzy weighted relative error support vector machine for reverse prediction of concrete components", *Computers & Structures*, 230, 106171.
[25] W. Bao, H. Wang, J. Chen et al. (2020) "Life prediction of slewing bearing based on isometric mapping and fuzzy support vector regression", *Transactions of the Institute of Measurement and Control*, 42(1), 94-103.

Edited by: Bradha Madhavan
Special issue on: High-performance Computing Algorithms for Material Sciences
Received: Dec 25, 2023
Accepted: Mar 21, 2024