WILEY | Hindawi

Research Article

Long Short-Term Memory Networks and Bayesian Optimization for Predicting the Time-Weighted Average Pressure of Shield Supporting Cycles

Wanzi Yan,[1] **Junhui Wang** ,[1] **Jingyi Cheng** ,[1] **Zhijun Wan** ,[1] **Keke Xing,**[1]
and Kuidong Gao[2]

[1]*Key Laboratory of Deep Coal Resource Mining (CUMT), Ministry of Education of China, School of Mines, China University of Mining and Technology, Xuzhou 221116, China*
[2]*Shandong Province Key Laboratory of Mine Mechanical Engineering, Shandong University of Science and Technology, Qingdao 266590, China*

Correspondence should be addressed to Junhui Wang; wangjunhui@cumt.edu.cn and Jingyi Cheng; 5759@cumt.edu.cn

Characteristic parameters of shield supporting in fully mechanized mining, especially time-weighted average pressure (TWAP), are crucial for the analysis and prediction of roof weightings in longwall panels. Despite the leap-forward development of underground data collection and transmission, mining and regional correlation analysis of massive shield data remains challenging. In this study, a hybrid machine learning model integrating the long short-term memory (LSTM) networks and the Bayesian optimization (BO) algorithm was developed to predict TWAP based on the setting pressure (SP), revised setting pressure (RSP), final pressure (FP), number of yielding (NY), TWAP in the last supporting cycle (TWAP (last)), and loading rate in each period. Statistical measures including the mean square error and mean absolute error were used to validate and compare the prediction performances of the BP model, the LSTM model, and the BO-LSTM model. Furthermore, sensitivity studies were carried out to evaluate the importance of input parameters. The results show that the BO-LSTM model is robust in predicting TWAP. FP and TWAP (last) are the most important input parameters in TWAP prediction, followed by RSP and NY. Moreover, the total importance scores of loading rates reach 0.229, indicating the necessity of including these parameters into the dataset. The proposed BO-LSTM model is capable of predicting TWAP which serves for shield-roof status intelligent perception.

1. Introduction

With the surge in energy demands and mining intensity, shallow resources are decreasing while deep resources are gradually exploited [1]. High ground stress induced by deep mining complicates mine pressure behaviors, including partial roof fall, rib spalling [2], shield crushing, and even rock burst [3]. Mine pressure monitoring is an important means to ensure safe and efficient production in coal mines. Mine pressure data in the roof-control area are comprised of basic data (roof-to-floor convergence, shield load, and leg closure) and statistical data (fissure development, the number of gob steps, etc.) [4]. In recent years, microseismic signals and elec-

tromagnetic radiation signals have been included in the dataset [5]. In fact, time-weighted average pressure (TWAP) in the dataset, which belongs to shield supporting feedback and one of the judgment bases for roof weighting, is significant for weighting prediction [6]. As electronic technology and sensor theories advance, the information transmission speed and accuracy have achieved a leap-forward development thanks to the shield electrohydraulic monitoring technology [7]. However, deep analysis on nonlinear and nonstationary shield pressure signals remains challenging, and there is a lack of correlation analysis on information monitored by different stations. Machine learning (ML) is a branch of artificial intelligence that has experienced tremendous progress in the

last two decades. Therefore, research on ML-based TWAP prediction of shields in the fully mechanized working face is meaningful for realizing the intelligent perception and coupling adaptation of shield-roof status.

ML has been extensively applied to mining engineering at an astounding rate primarily for the purposes of mine disaster monitoring (e.g., gushing water, gas, and mine pressure [8–11]) and roadway excavation technology (e.g., coal-rock interface recognition and digital drilling [12–17]). With respect to mine pressure monitoring, Li et al. [18] optimized the back-propagation (BP) neural network by using the particle swarm optimization (PSO) algorithm and established a PSO-BP model for assessing rock burst risk. Tan et al. [19] optimized the BP neural network by using the genetic algorithm (GA) and established a GA-BP mine pressure prediction model for predicting the weighting interval and intensity. Sun et al. [20] processed the monitored mine pressure data by means of chaotic analysis and established a shield pressure prediction model based on a support vector machine. The above researches all strived to predict roof weighting based on the traditional neural network. Although some researchers optimized the neural network by using the PSO algorithm, they failed to fully consider the time sequence of mine pressure data. Problems remain in the adaptability between intelligent optimization algorithms and neural networks. In addition, researchers rarely built a weighting prediction model considering the load-bearing information of shields (e.g., TWAP), and the characteristic parameters of shield supporting cycle lacks in-depth exploration. In summary, the weighting prediction model can still be optimized to a large extent.

Despite the application of various ML algorithms to roof weighting prediction, the applicability and mobility of these algorithms are questioned for two reasons. First, mine pressure behaviors possess strong periodicity which is reflected by the time sequence change in shield pressure data, so it is essential to choose a model suitable for dealing with time sequence problems. Second, due to the improper choice of input variables, a model matching the current engineering background can hardly be adjusted adaptively to the best performance in another engineering background. Long short-term memory (LSTM), first proposed by Hochreiter and Schmidhuber [21], is an excellent deep learning network widely adopted in fields such as natural language processing and time sequence prediction [22–24]. Since its establishment, the LSTM model demonstrates extraordinary performance for mine pressure prediction [25]. It is absolutely essential to select a suitable intelligent optimization algorithm to adjust its parameters. The commonly used optimization algorithms include the GA, the simulated annealing (SA) algorithm, the PSO algorithm, and the Bayesian optimization (BO) algorithm. Among these algorithms [26–32], the BO algorithm whose iterations are few is able to quickly find the optimal value without wasting resources. Such a feature enables this algorithm to be applicable to the coal mine field.

Several contributions of the study are summarized as follows: (1) Loading rates in each period are applied to TWAP prediction for the first time. (2) The BO-LSTM hybrid model is proposed to better predict the TWAP value in the next supporting cycle by setting appropriate input parameters, and this value is regarded as the criterion for weighting prediction. The study is mainly aimed at efficiently and accurately predicting the TWAP value in the next shield supporting cycle through the BO-LSTM model and serving for the intelligent perception of shield-roof status. The input parameters include setting pressure (SP), final pressure (FP), loading rates in each period, and number of yielding (NY) [33–36]. Considering that model development and validation require sufficient available shield data, the 22307 fully mechanized working face of the Bulianta Coal Mine was taken as the study area. The Pearson correlation coefficient was adopted for analyzing correlation between the input parameters and the output parameter. Meanwhile, the sensitivity between the input parameters and the output parameter was analyzed in depth.

2. Dataset Preparation

2.1. Description of the Study Area. The Bulianta Coal Mine, located in Inner Mongolia (Figure 1(a)), is selected as the study area. The buried depth of the 22307 working face is 88-245 m, and the average thickness, face length, and strike length are 7.25 m, 301 m, and 4954.05 m, respectively. The geological column is exhibited in Figure 2. The lithological formations above the coal seam (thickness 85-230 m) consist of mudstone and sandstone, while the thickness of loose layer is 8-23 m.

The 22307 working face is equipped with 150 ZY/18000/32/70D shields whose numbers are given in Figure 1(b). A PM32 electrohydraulic controller and a pressure sensor are arranged on each shield (Figure 1(c)). The data acquisition and transmission method is presented in Figure 1(d) [37]. Specifically, the pressure data collected by the sensors are conveyed through the CAN bus network to the ground server and the host computer, and then the automatic control of the shields is realized through the decision of the host computer. The self-developed Status of Shield and Roof IntelliSense (SSRI) system, integrated with the data mining and cycle analysis technology, is capable of efficiently and accurately analyzing massive monitoring data and cooperates with the host computer to make decisions intelligently. Detailed architecture and working principle of the SSRI system can be found in the literature [38].

With the aid of the SSRI system, a total of over 56,000 pieces of data are extracted from the massive pressure data of 80 shields in the 22307 working face (Figure 3(a)). In this study, three shields, i.e., 1#, 40#, and 80#, from the end to the middle of the working face are selected as the research object. The processed pressure data of the three shields are displayed in Figure 3(b). The pressure data of each shield contains n supporting cycles reflecting the relationship between the shield and the surrounding rock.

2.2. Data collection

2.2.1. Output. The output variable is TWAP. The shield supporting cycle refers to the pressure variation process of a shield, including its setting, loading/unloading, and forward moving. The cycles can be obtained from the pressure data in Figure 3(b). TWAP [39], a characteristic quantity reflecting the pressure variation in shield cycles, is the basis for
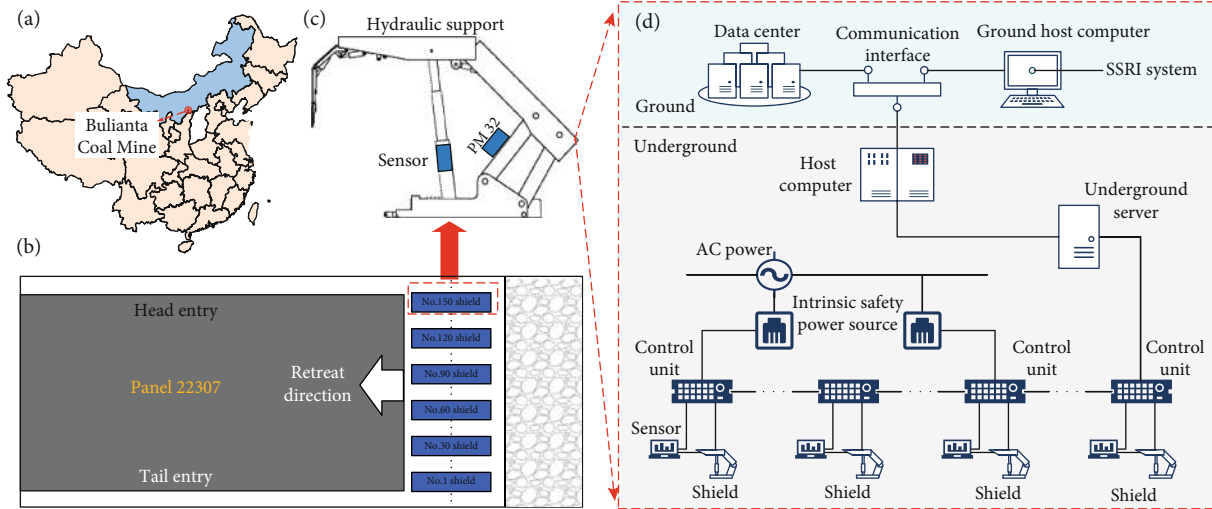
FIGURE 1: Description of the study area: (a) location of the Bulianta Coal Mine; (b) positions and numbers of shields; (c) positions of pressure sensors and PM32 controllers; (d) data acquisition and transmission method and SSRI system.

Lithology	Geological column	Thickness (m)	Depth (m)
Coal 1-2#		4.80	52.68
Sandy mudstone		4.68	57.36
Fine sandstone		2.90	60.26
Medium sandstone		29.88	90.14
Sandy mudstone		1.76	91.90
Coal 2-2#		7.47	99.37

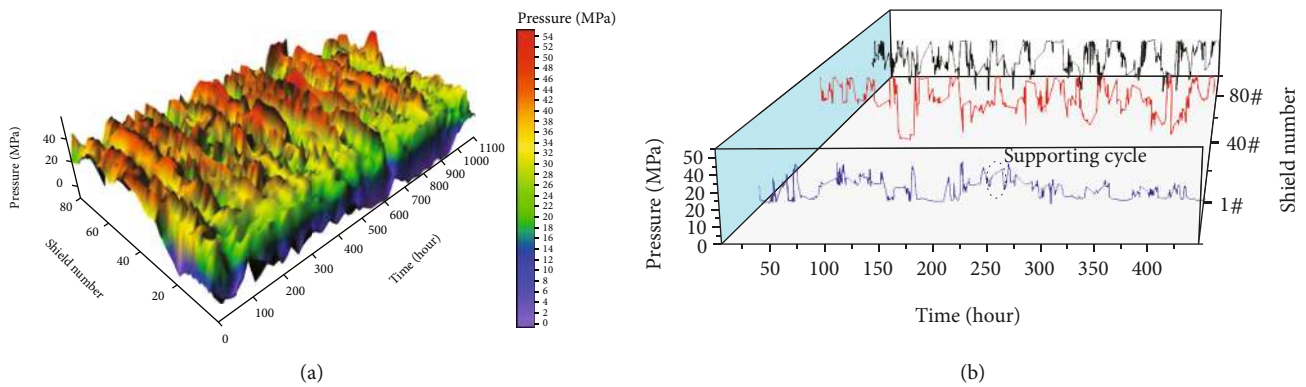Coal seam was being mined

FIGURE 2: Geological column.

FIGURE 3: Processed pressure data: (a) pressure data of shields in the 22307 working face; (b) partial processed pressure data of 1#, 40#, and 80# shields.
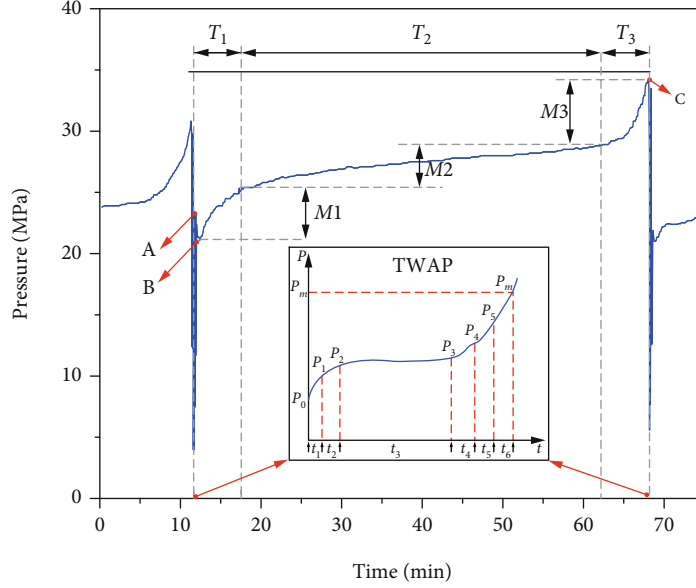
FIGURE 4: Shield supporting cycle curve.

judging periodic weighting. Its value can be calculated by equation (1) according to the time sequence variation curve of shield pressure in each supporting cycle (Figure 4).

$$P_t = \frac{(1/2)\sum_{n=1}^{m}(P_{n-1} + P_n)t_n}{\sum_{n=1}^{m}t_n}, \qquad (1)$$

where P_t is TWAP (MPa); P_n is the shield pressure (MPa); t_n is the total time (min).

In this study, the SSRI system can identify the number of cycles in each pressure data set and calculate the TWAP value in each cycle. The processed results of the pressure data (Figure 3(a)) indicate that each shield contains 600 supporting cycles. The overall TWAP data within 1100 h are shown in Figure 5.

2.2.2. Input Variables. The input variables are multidimensional feature parameters of supporting cycles. As can be observed in Figure 4, the typical shield supporting cycle can be simplified into three periods in accordance with the loading/unloading characteristics: initial period (T_1), relatively stable period (T_2), and cutting influenced and neighboring shield movement period (T_3). Based on the loading characteristics of the supporting cycle in different stages, the SSRI system can extract 27 characteristic parameters by employing multiple algorithms such as Naive Bayes, clustering, classification, and correlation. In addition to SP and FP, loading rates and NY are also included in these parameters as input variables for TWAP prediction.

The 27 characteristic parameters of TWAP factors are subjected to gray correlation analysis. Among them, 9 parameters with high correlation degrees, namely, SP, revised setting pressure (RSP), FP, NY, loading rates in the initial period (LRIP 5 min and LRIP 10 min), loading

rate in the relatively stable period (LRRSP), loading rate in the cutting influenced and neighboring shield movement period (LRCMP), and TWAP in the last supporting cycle (TWAP last), are regarded as input variables in this study.

(1) SP and RSP. SP, an active supporting force provided by the shield to the roof, plays a crucial role in ground controlling in the working face (point A in Figure 4). After shield pressure reaches SP, it will first experience a short-term (1-3 min) drop and then rise slowly or stabilize. The shield leg pressure at this moment is the actual SP provided by the shield to the roof, and it is defined as the RSP (point B in Figure 4).

(2) FP. The FP of a shield (point C in Figure 4), which is normally considered as the maximum supporting load in the whole supporting cycle, is attained when its neighboring shields are lowered and advanced.

(3) Loading Rates. Loading rates refer to the rates of load changes in the T_1, T_2, and T_3 stages in Figure 4. They reflect the state of roof activity in different time periods. Among the three, LRIP = M_1/T_1, LRRSP = M_2/T_2, and LRCMP = M_3/T_3. It is noteworthy that LRIP (5 min) and LRIP (10 min) refer to the ratio of resistance change to time within the first 5 min and 10 min in the T_1 stage, respectively.

(4) NY. When the shield pressure reaches or exceeds the rated shield capacity, high-pressure emulsion overflows for protecting the shield leg under the condition that working face production is suspended for a long time or that the roof activity is intense. The process of yield valve opening and closure is called a yielding cycle. The number of yielding cycles is an important indicator reflecting the intensity of roof activity.
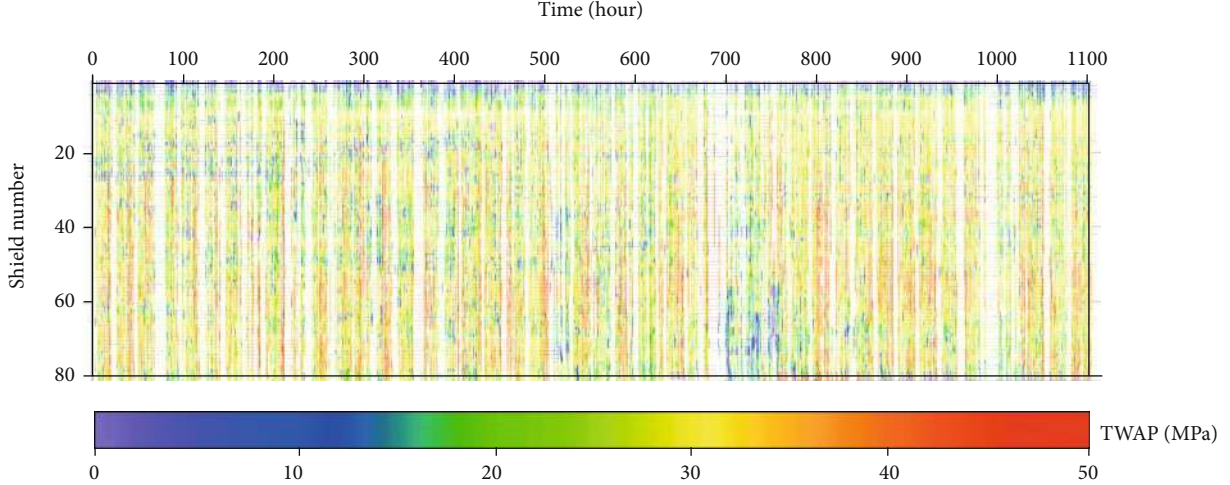
Time (hour)

FIGURE 5: TWAP data in 600 supporting cycles of the shields.

The data of 9 feature parameters in this study are preliminarily analyzed in Table 1.

3. Methods Used

3.1. Fundamental Theory of LSTM and BOA

3.1.1. Long Short-Term Memory Networks (LSTM). LSTM neural networks are especially applicable to processing highly time-dependent problems. Roof weighting has notable periodic characteristics, as the pressure waveform is a strong time series whose front and back inputs are considerably correlated. Thus, LSTM neural networks are applied to predicting TWAP of shields.

The LSTM neural unit is optimized on the basis of the recurrent neural network (RNN) by adding control gates inside the units. The addition of a new cell state input enables the LSTM to remember and retain historical information. The entire neural unit is controlled by the input gate, the forget gate, and the output gate.

The internal structure of the LSTM neural unit at three consecutive times is illustrated in Figure 6 where x is the input vector of the neural unit at the current time; h is the output vector of the neural unit; C is the cell state; σ and $\tan h$ are both activation functions; f, i, and O are the forget gate, the input gate, and the output gate, respectively; and W and b are the weight and the deviation matrix, respectively.

Equations (2)–(7) show the internal operations followed by neural units:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right), \qquad (2)$$

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right), \qquad (3)$$

$$\tilde{C}_t = \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right), \qquad (4)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t, \qquad (5)$$

$$O_t = \sigma\left(W_O \cdot [h_{t-1}, x_t] + b_O\right), \qquad (6)$$

$$h_t = O_t \cdot \tanh\left(C_t\right). \qquad (7)$$

TABLE 1: Descriptive statistics of inputs and output.

Parameter	Min value	Max value	Mean value	Standard deviation
TWAP	2.00	46.19	27.55	8.93
SP	1.8	46.2	23.38	6.79
FP	1.8	47.7	31.26	11.00
RSP	1.5	46.2	22.13	7.24
LRIP (5 min)	-4.39	5.30	0.38	0.93
LRIP (10 min)	-2.45	4.83	0.31	0.68
LRRSP	-0.38	0.46	0.02	0.05
LRCMP	-8.60	5.36	0.35	0.84
NY	0	19	0.41	1.62

3.1.2. Bayesian Optimization (BO) Algorithm. The BO algorithm is a kind of probability distribution optimization algorithm used for the automatic tuning of ML hyperparameters. The algorithm, whose core idea is to give an optimized objective function and update the posterior distribution of the objective function by continuously adding sample points, is mainly oriented to solving complex black box problems with multimodality, nonconvexity, high dimensionality, and high evaluation costs. Established on the basis of the Gaussian process, it considers the previous parameter information and constantly updates the prior. It not only greatly saves the optimization cost due to the small number and fast speed of iterations, but it also boasts good performance against both convex and nonconvex optimization problems. The basic flow of the BO algorithm is listed in Figure 7.

3.2. Modelling Methodology. Shield data, a kind of strong time series data, feature instability, nonlinearity, and periodic uncertainty. In this study, a prediction model regarding shield pressure data is built based on LSTM networks. Since data of various shields are not synchronized with significant differences, the BO algorithm is combined with the LSTM model so that the hyperparameters match the data characteristics of each shield.
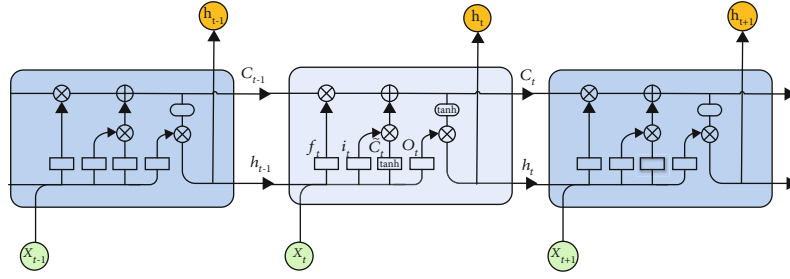
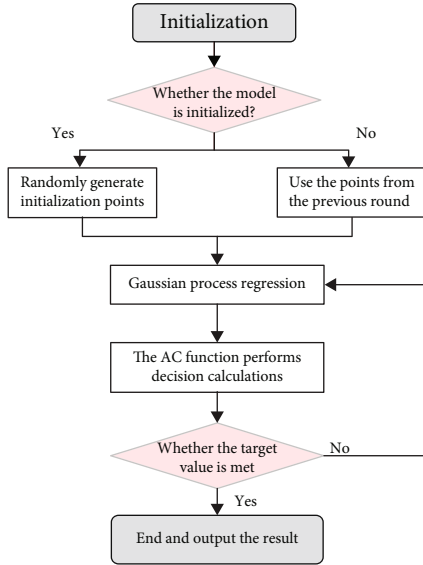FIGURE 6: Structure of the LSTM neural unit.

FIGURE 7: Basic flow of the BO algorithm.

The framework of the BO-LSTM model is listed in Figure 8. It is known that both the structure and model parameters of an artificial neural network exert an influence on the performance of the model, so the number of units in the LSTM layers, the parameter of the dropout layer, the batch size, and the optimizer learning rate are selected as the optimization objects of the Bayesian optimizer. The model is composed of an input layer, two LSTM layers, a dropout layer, and an output layer. The loss function adopts the mean square error (MSE), and the model training process is optimized through the Adam algorithm. The optimization value ranges of hyperparameters are set as follows: number of units in the LSTM layers (1, 80), batch size (1, 80), optimizer learning rate (0.0001, 0.01), and parameter of the dropout layer (0.1, 0.6).

The BP model and the LSTM model are taken as contrast models in this study. The BP model is comprised of an input layer, two hidden layers (20 units), a dropout layer (0.2 parameter), and an output layer. The LSTM model consists of an input layer, a LSTM layer (20 units), a dropout layer (0.2 parameter), and an output layer. The learning rate of the Adam optimizer is 0.01.

The flow of the BO-LSTM model is as follows:

Step 1. To divide the experimental data into the training set and the test set.

Step 2. To initialize the BO algorithm by taking batch size, optimizer learning rate, and dropout layer parameter in the LSTM network model as optimization objects.

Step 3. To randomly calculate the current function distribution according to the formula.

Step 4. To adjust the current function distribution according to the strategy selected by the selection function.

Step 5. To judge whether the termination condition is met. If it is met, the optimal hyperparameter values are returned; otherwise, Step 4 is returned.

Step 6. To construct the LSTM network model with the optimal hyperparameters.

Step 7. To conduct training and testing.

3.3. Statistical Assessment. The assessment of this study was carried out using MSE and the mean absolute error (MAE) [40], the most common indicators for ML model validation and comparison. The equations of the two indexes are as follows:

$$ \mathrm{MAE} = \frac{1}{N} \sum_{n=1}^{N} |\widehat{y_n} - y_n|, \tag{8} $$

$$ \mathrm{MSE} = \frac{1}{N} \sum_{n=1}^{N} (y_n \wedge - y_n)^2. \tag{9} $$

where N is the number of instances; $\widehat{y}_n$ is the value predicted by the model; y_n is the true value. The two indexes can both represent the error between $\widehat{y}_n$ and y_n. A smaller value of the index is indicative of a smaller deviation of the predicted value from the true value and better performance of the model.

4. Results and Discussion

4.1. Correlations among Inputs and Output. The relationships between the 9 input variables and the output variable have been investigated based on the Pearson correlation coefficient [41]. Figure 9 is a distribution diagram of correlation coefficients among variables. The results show that SP, LRIP, FP, NY, and RSP are all positively correlated with TWAP, and the loading rates in the three stages (T_1, T_2, and T_3) will
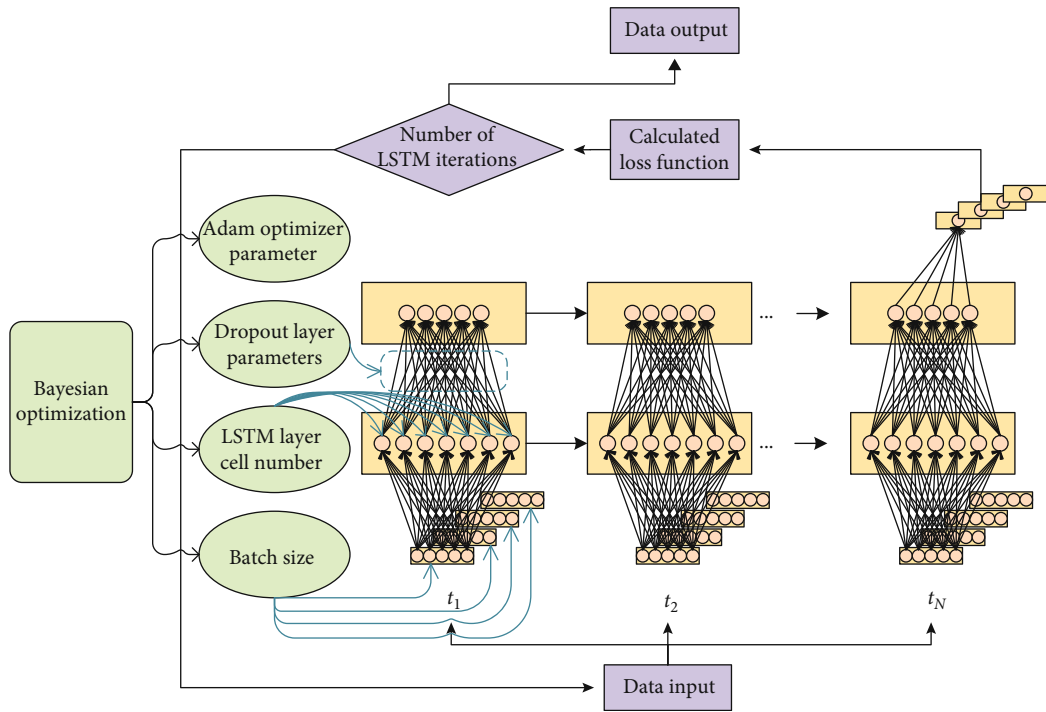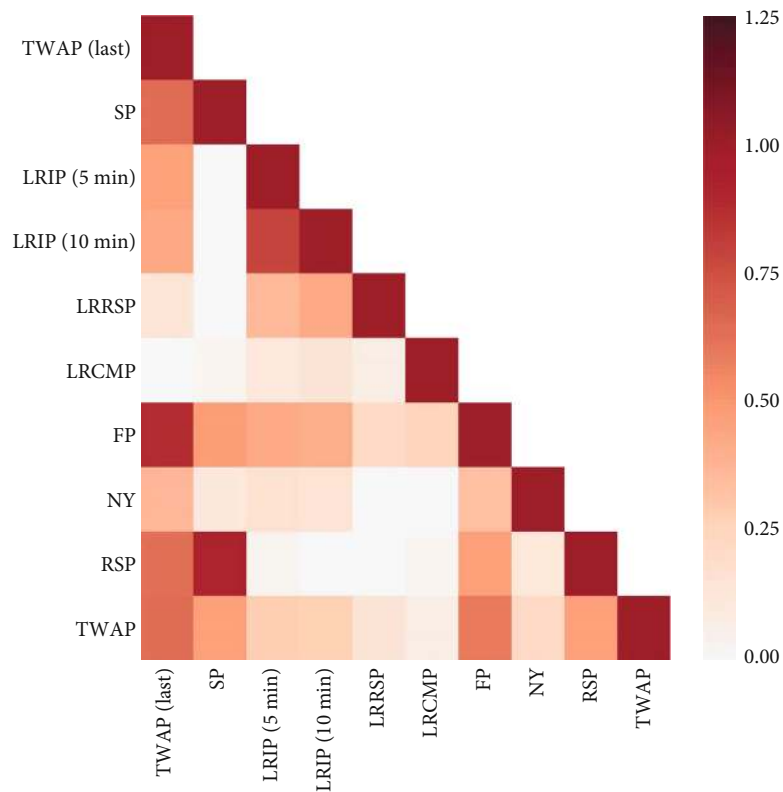
FIGURE 8: Framework of the BO-LSTM model.

FIGURE 9: The Pearson correlation coefficients.

(a)

(b)

(c)

Number of supporting cycles

- ■ Original value
- ● BP predictive value
- ● LSTM predictive value
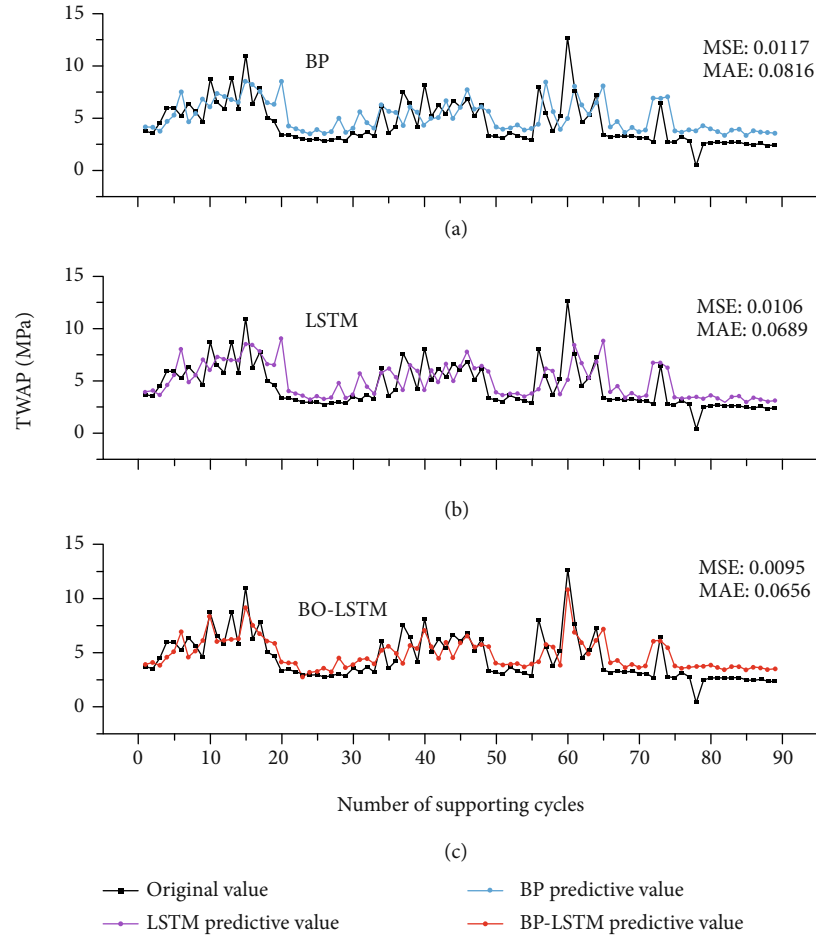- ● BP-LSTM predictive value

FIGURE 10: Comparison among the TWAP prediction performances of three models for 1# shield.

affect the variation of the TWAP value. NY will lead to constant fluctuations of shield supporting pressure. When the yield valve opens too frequently, the TWAP value will fluctuate accordingly.

4.2. Predictive Capability of the Models

4.2.1. TWAP Prediction for 1# Shield. The TWAP values predicted by the three models on the test sets and the actual TWAP values for 1# shield are compared in Figure 10. Among the 297 pieces of 1# shield data (Sept. 27, 2014-Oct. 17, 2014), 70% are taken as the training set and 30% are used as the test set.

Figure 10(a) displays the results predicted by the BP model, the MSE value and MAE value being 0.0117 and 0.0816, respectively. Figure 10(b) exhibits the results predicted by the LSTM model, the MSE value and MAE value being 0.0106 and 0.0689, respectively. Figure 10(c) gives the results predicted by the BO-LSTM model. After being tuned by the BO algorithm, the hyperparameters of the BO-LSTM model are as follows: the number of units in the LSTM layer is 42, the batch size is 52, the optimizer learning rate is 0.0094, and the dropout layer parameter is 0.58. In this case, the MSE value and MAE value are 0.0095 and 0.0656, respectively.

4.2.2. TWAP Prediction for 40# Shield. The TWAP values predicted by the three models on the test sets and the actual TWAP values for 40# shield are compared in Figure 11. Among the 662 pieces of 40# shield data (Sept. 27, 2014-Nov. 11, 2014), 70% are taken as the training set and 30% are used as the test set.

Figure 11(a) displays the results predicted by the BP model, the MSE value and MAE value being 0.0194 and 0.1058, respectively. Figure 11(b) exhibits the results predicted by the LSTM model, the MSE value and MAE value being 0.0165 and 0.0924, respectively. Figure 11(c) gives the results predicted by the BO-LSTM model. After being tuned by the BO algorithm, the hyperparameters of the BO-LSTM model are as follows: the number of units in the LSTM layer is 77, the batch size is 65, the optimizer learning rate is 0.0061, and the dropout layer parameter is 0.24. In this case, the MSE value and MAE value are 0.0155 and 0.0886, respectively.

4.2.3. TWAP Prediction for 80# Shield. The TWAP values predicted by the three models on the test sets and the actual TWAP values for 80# shield are compared in Figure 12. Among the 645 pieces of 80# shield data (Sept. 27, 2014-Nov. 11, 2014), 70% are taken as the training set and 30% are used as the test set.

(a)

(b)

(c)

Number of supporting cycles

— Original value   — BP predictive value
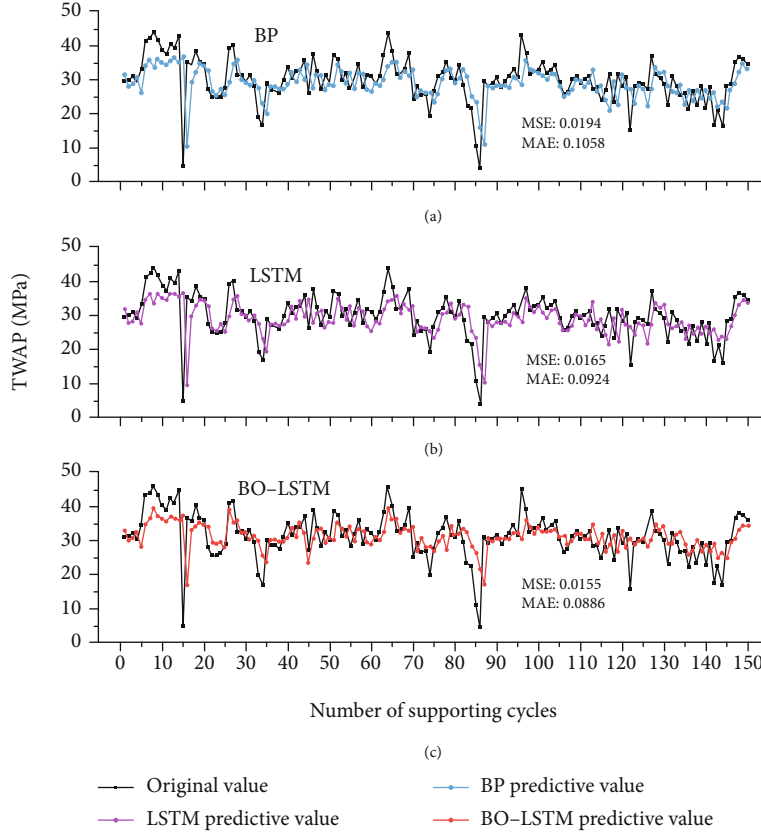— LSTM predictive value   — BO–LSTM predictive value

FIGURE 11: Comparison among the TWAP prediction performances of three models for 40# shield.

Figure 12(a) displays the results predicted by the BP model, the MSE value and MAE value being 0.0334 and 0.1442, respectively. Figure 12(b) exhibits the results predicted by the LSTM model, the MSE value and MAE value being 0.0331 and 0.1436, respectively. Figure 12(c) gives the results predicted by the BO-LSTM model. After being tuned by the BO algorithm, the hyperparameters of the BO-LSTM model are as follows: the number of units in the LSTM layer is 69, the batch size is 28, the optimizer learning rate is 0.0041, and the dropout layer parameter is 0.42. In this case, the MSE value and MAE value are 0.0315 and 0.1362, respectively.

4.2.4. Analysis of Prediction Results. The errors of the prediction results by the BP, LSTM, and BO-LSTM models are shown in Figure 13, according to which the BO-LSTM model achieves the best prediction performance. Compared with the BP model, the LSTM model succeeds in reducing the errors when it is used to predict the TWAP data of shields. This result suggests that the LSTM model is more suitable for processing TWAP data with a notable time sequence than the BP model. Moreover, the BO-LSTM model whose hyperparameters have been automatically tuned by the BO algorithm boasts an improved performance compared with the LSTM model which has not received BO optimization, with the errors of the three shields by the BO-LSTM model all being smaller than those by the LSTM model and the BP

model. The result confirmed the feasibility of the BO algorithm in improving the performance of the LSTM model.

The BP model presents the largest errors because its TWAP prediction is based on the data of the previous moment. It lacks the memory of data characteristics in the past. The LSTM model demonstrates a better prediction performance because it memorizes data characteristics in the past. However, it fails to be well compatible with the difference in data characteristics of each shield because the shield weightings in the 22307 working face are not synchronous while the LSTM model adopts the same set of hyperparameters to train the data of shields. In contrast, the BO-LSTM model boasts the smallest prediction errors because it not only possesses the advantages of the LSTM model but also finds out the hyperparameters that are the most suitable for the current data in accordance with the distinct characteristics of shields.

With respect to the spatial evolution, the MSE value and MAE value increase from 1# shield to 80# shield, demonstrating that the prediction error of the shield (80#) in the middle is larger than that of the shield (1#) in the end. The reason is that compared with 1# shield, 80# shield corresponds to a smaller weighting interval and experiences more fluctuant supporting pressure [42].

In short, the BO-LSTM model proposed in this study can quickly and effectively predict the TWAP values of shields at different positions. Its performance may be improved due to
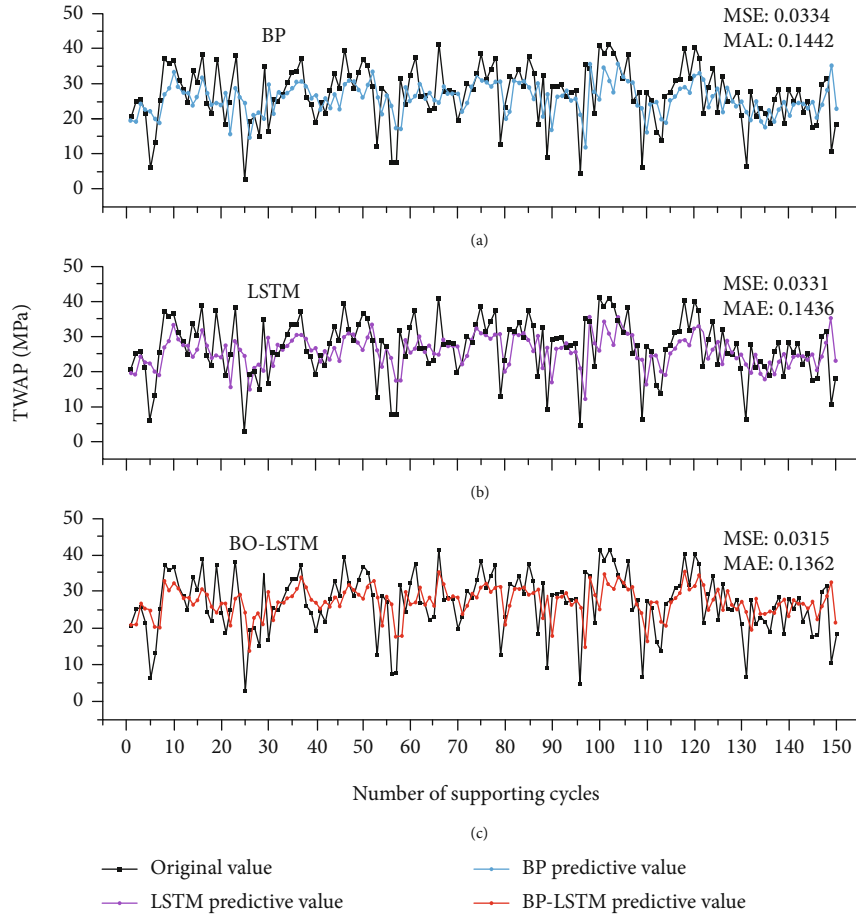
FIGURE 12: Comparison among the TWAP prediction performances of three models for 80# shield.
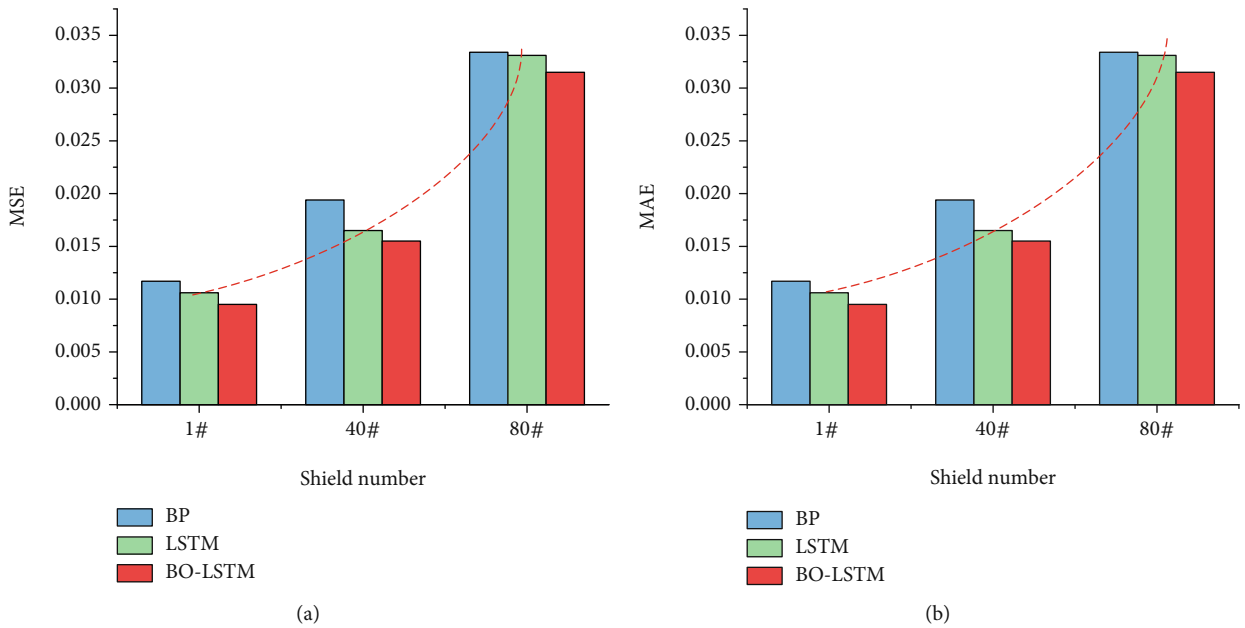

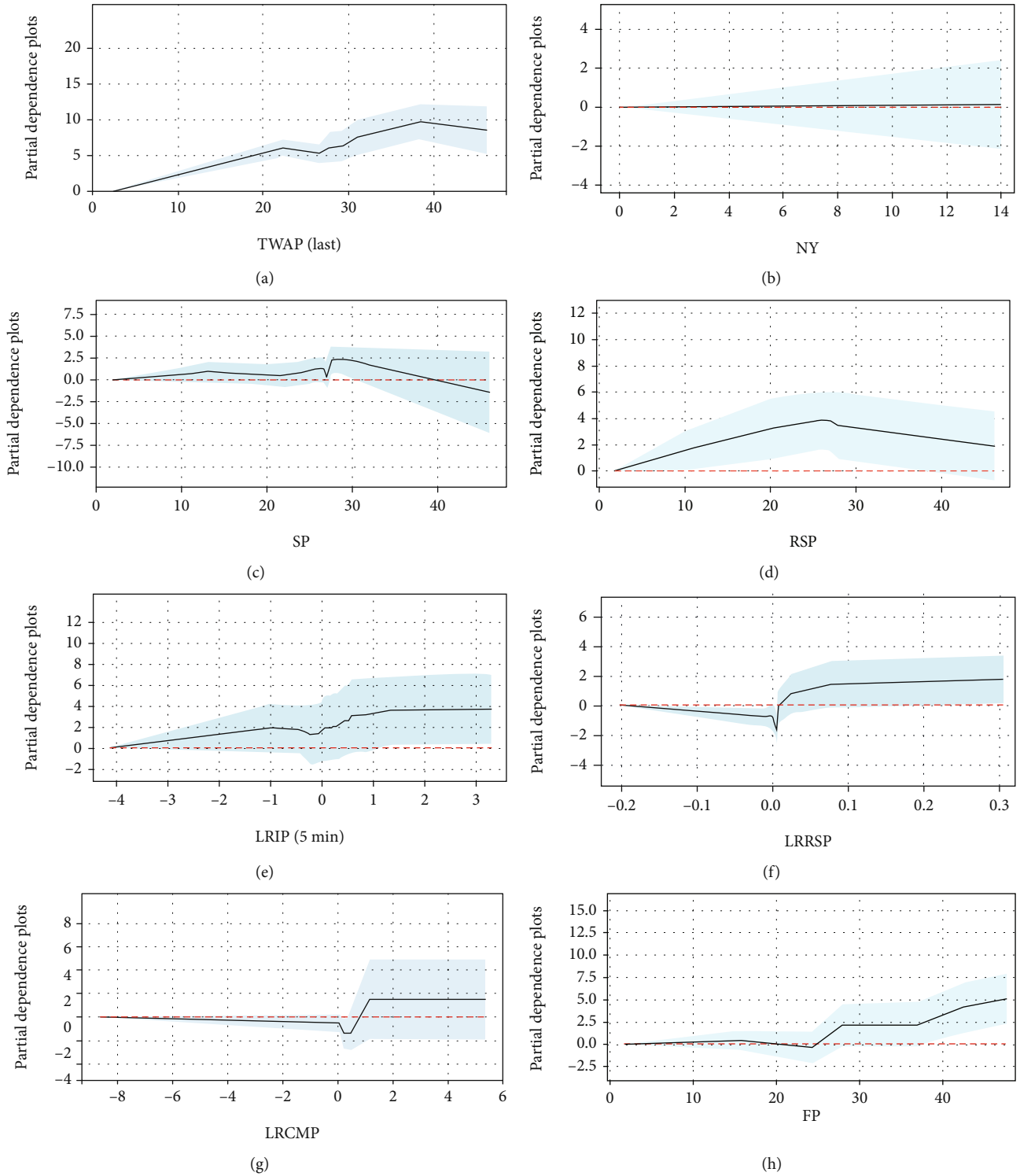FIGURE 13: Comparison among errors by the BP, LSTM, and BO-LSTM models.

FIGURE 14: Partial dependence of the influencing variables on TWAP prediction.

the different conditions of coal seam occurrence. One advantage of hybrid ML models like the BO-LSTM model is that they can process large and complex data. Hence, such models are suitable for TWAP prediction of shields in the fully mechanized mining face.

4.3. Sensitivity Study of Input Variables. To better grasp the influence of each input variable on prediction of TWAP values, the sensitivity of input parameters was analyzed by partial dependency plots (PDP), and the importance of input parameters was analyzed by the algorithm XGBoost [43, 44].
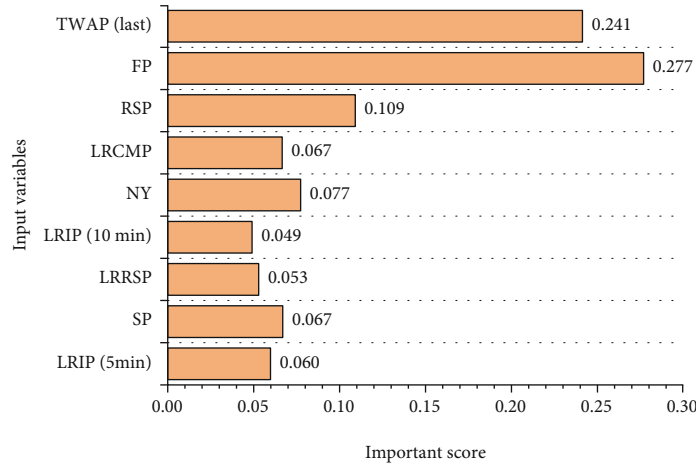
FIGURE 15: Importance scores of input variables on TWAP prediction.

PDP is a popular tool to determine the sensitivity of the output variable on the individual input variables [45]. The horizontal axis represents all possible values of the input variable, and the vertical axis represents the mean value of the predicted output value. Figures 14(a)–14(h) demonstrate the partial dependence of input variables (excluding LRIP (10 min)) and the output variable. The results reveal that loading rates in different stages, FP, and TWAP (last) are positively correlated with the TWAP value. SP and RSP are positively correlated with the TWAP value before 27 MPa, whereas they are negatively correlated with it after this pressure, indicating that 27 MPa is the best SP. The TWAP value varies more intensely with the changes in FP and TWAP (last), compared with NY, SP, RSP, and loading rates in different stages.

To illustrate the importance of input variables more clearly, the importance scores (IS) of all input variables were deduced through normalizing the summation (Figure 15). The results manifest that FP (IS: 0.277) exerts the greatest influences on TWAP prediction, followed by TWAP (last) (IS: 0.241), RSP (IS: 0.109), NY (IS: 0.077), SP (IS: 0.067), and loading rates in different stages (IS: mostly around 0.06) in turn. From IS, it can be seen that FP and TWAP (last) are the primary variables influencing the TWAP prediction performance. The two parameters are also the basis for judging roof weighting in traditional mining, and this explains why the TWAP value varies more intensely with the changes in FP and TWAP (last) in Figure 14.

Indeed, traditional feature parameters (e.g., FP and TWAP (last)) remarkably influence the TWAP value in the next supporting cycle, which has been proved by IS. Nevertheless, the total IS of LRIP, LRRSP, and LRCMP reaches 0.229, so it can be concluded that loading rates in different stages are also indispensable for TWAP prediction.

In a supporting cycle, the feature parameters which are later in time play a more important role in TWAP prediction. For instance, LRCMP is more important than LRIP, and FP is more important than SP and RSP. Meanwhile, the shield supporting characteristic with a higher loading rate plays a greater role in TWAP prediction.

In addition, NY also exerts a certain influence on TWAP prediction. The importance of RSP on TWAP prediction is larger than that of SP, probably because when the shield is initially set, it has not fully contacted the roof and SP cannot obviously reflect the roof condition. This phenomenon exists commonly in shield supporting cycles. For example, in the 22304 working face of the Buertai Coal Mine in China and the LW23 working face of the Emerald Coal Mine in the United States, 48.5% and 71.9% of shield supporting cycles underwent a load decrement of over 1 MPa after the shield was initially set, and the shield supporting cycles within which the load decrement lasted shorter than 1 min accounted for 70.3% and 96%, respectively.

In general, the above research results verify the importance of the thoroughly explored input variables to a certain extent. It is believed that further research based on this viewpoint will have far-reaching significance for the further development of TWAP prediction.

5. Conclusions

In this study, the BO-LSTM model, a hybrid ML model integrating the advantages of the LSTM model and the BO algorithm, was proposed to predict the shield TWAP values for the purpose of shield-roof status intelligent perception. A total of over 56000 pieces of regionally correlated shield pressure data were collected from the Bulianta Coal Mine, and the SSRI system assisted in identifying the supporting cycles and extracting characteristic parameters from them. Nine parameters sharing large gray correlation degrees with the output variable (TWAP), namely, SP, RSP, FP, NY, loading rates, and TWAP (last), were selected as the input variables. Statistical measures including MSE and MAE were used to validate and compare the prediction performance of the BP model, the LSTM model, and the BO-LSTM model. Moreover, PDP and IS were employed to evaluate the importance of input parameters in the model study. The main conclusions are as follows.

(1) Among the three ML models, the BO-LSTM model achieves the best performance in terms of MSE and MAE, followed by the LSTM model and the BP model in turn. Compared with the BP model, the LSTM model is more applicable to processing data with a notable time sequence. Furthermore, the BO algorithm can find the hyperparameters most suitable for the current data in accordance with the distinct characteristics of shields, thus further improving the prediction accuracy

(2) The assessment criteria increase from the shield (1#) in the end to the shield (80#) in the middle, because the former experiences less fluctuant TWAP than the latter

(3) The input variables extracted by the SSRI system are effective for TWAP prediction. The PDP and IS results reveal that FP and TWAP (last) are the most important parameters in TWAP prediction, followed by RSP and NY. Besides, it is necessary to revise SP. The total importance scores of loading rates in different stages reach 0.229, indicating that loading rates are also indispensable for TWAP prediction. Meanwhile, the shield supporting characteristic with a higher loading rate plays a greater role in TWAP prediction. In a supporting cycle, the characteristic parameters which are later in time play a more important role in TWAP prediction than those which are earlier in time. For instance, FP is more important than RSP, and LRCMP is more important than LRIP

Abbreviations

TWAP: Time-weighted average pressure
SP: Setting pressure
RSP: Revised setting pressure
FP: Final pressure
NY: Number of yielding
LRIP: Loading rates in the initial period
LRRSP: Loading rate in the relatively stable period
LRCMP: Loading rate in the cutting influenced and neighboring shield movement period
ML: Machine learning
LSTM: Long short-term memory networks
BO: Bayesian optimization
BP: Back-propagation neural network
PSO: Particle swarm optimization
SA: Simulated annealing
GA: Genetic algorithm
SSRI: Status of Shield and Roof IntelliSense system
MSE: The mean square error
MAE: Mean absolute error.

Data Availability

The data used to support the findings of this study are available from the corresponding authors upon request.

References

[1] M. Cai and E. T. Brown, "Challenges in the mining and utilization of deep mineral resources," *Engineering*, vol. 3, no. 4, pp. 432-433, 2017.

[2] S. Liu, Z. Wan, Y. Zhang, S. Lu, X. Ta, and Z. Wu, "Research on evaluation and control technology of coal pillar stability based on the fracture digitization method," *Measurement*, vol. 158, p. 107713, 2020.

[3] S. S. Peng, *Longwall Mining*, CRC Press, Boca Raton, 2019.

[4] M. Qian, P. Shi, and J. Xu, *Mine Pressure and Rock Formation Control*, China University of Mining and Technology Press, Xuzhou, China, 2010.

[5] L. M. Dou, X. Q. He, and E. Y. Wang, "Electromagnetic emission technique of monitoring rock burst and its application," *Journal of China Coal Society*, vol. 29, no. 4, pp. 396–399, 2004.

[6] S. S. Peng, *Coal Mine Ground Control*, Woodhead Publishing, Cambridge, 3rd ed edition, 2008.

[7] S. B. Li, W. S. Wei, and J. F. Niu, "Statement on electric and hydraulic control of hydraulic powered support and mining face automation technology," *Coal Science & Technology*, vol. 35, no. 11, pp. 1–5, 2007.

[8] J. Y. Xiao, M. M. Tong, Q. Fan, and C. J. Zhu, "Fuzzy neural network model applied in the mine water inrush prediction," in *International Conference on Image Processing & Pattern Recognition in Industrial Engineering*, Xi'an, China, 2010.

[9] Q. Meng, M. A. Xiaoping, and Y. Zhou, "Prediction of mine gas emission with improved particle swarm support vector machine," *Mining Safety & Environmental Protection*, vol. 42, no. 2, pp. 1–5, 2015.

[10] X. Xu, X. Rui, Y. Fan, T. Yu, and Y. Ju, "Forecasting of coalbed methane daily production based on T-LSTM neural networks," *Symmetry*, vol. 12, no. 5, p. 861, 2020.

[11] J. Lin, T. Ren, Y. Cheng, J. Nemcik, and G. Wang, "Cyclic N2 injection for enhanced coal seam gas recovery: a laboratory study," *Energy*, vol. 188, p. 116115, 2019.

[12] W. Hou, "Identification of coal and gangue by feed-forward neural network based on data analysis," *International Journal of Coal Preparation & Utilization*, vol. 39, no. 1, pp. 33–43, 2017.

[13] H. Chiranth and G. Ken, "Evaluation of coupled machine learning models for drilling optimization," *Journal of Natural Gas Science & Engineering*, vol. 56, pp. 397–407, 2018.

[14] Y. Yang, X. Li, H. Li, D. Li, and R. Yuan, "Deep Q-Network for optimal decision for top-coal caving," *Energies*, vol. 13, no. 7, p. 1618, 2020.

[15] C. Qi, Q. Chen, A. Fourie, and Q. Zhang, "An intelligent modelling framework for mechanical properties of cemented paste backfill," *Minerals Engineering*, vol. 123, pp. 16–27, 2018.

[16] C. Qi, A. Fourie, Q. Chen, and Q. Zhang, "A strength prediction model using artificial intelligence for recycling waste tailings as cemented paste backfill," *Journal of Cleaner Production*, vol. 183, pp. 566–578, 2018.

[17] C. Qi, X. Tang, X. Dong, Q. Chen, A. Fourie, and E. Liu, "Towards intelligent mining for backfill: a genetic programming-based method for strength forecasting of cemented paste backfill," *Minerals Engineering*, vol. 133, pp. 69–79, 2019.

[18] H. Li, Z. Li, R. He, and Y. Yan, "Rock burst risk evaluation based on particle swarm optimization and BP neural network," *Journal of Mining & Safety Engineering*, vol. 31, no. 2, pp. 203–231, 2014.

[19] T. Tan, Z. Yang, F. Chang, and K. Zhao, "Prediction of the first weighting from the working face roof in a coal mine based on a GA-BP neural network," *Applied Sciences*, vol. 9, no. 19, p. 4159, 2019.

[20] G. Sun, H. P. Zhou, and K. L. Sun, "Prediction of mine pressure monitoring data based on support vector machine," *Journal of Fuyang Teachers College*, vol. 30, no. 3, pp. 46–49, 2013.

[21] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[22] C. Xiong and M. Su, "IARNN-based semantic-containing double-level embedding Bi-LSTM for question-and-answer matching," *Computational Intelligence and Neuroscience*, vol. 2019, 10 pages, 2019.

[23] M. Ibrahim, A. Alsheikh, Q. Al-Hindawi, S. Al-Dahidi, and H. ElMoaqet, "Short-time wind speed forecast using artificial learning-based algorithms," *Computational Intelligence and Neuroscience*, vol. 2020, 15 pages, 2020.

[24] J. Y. Choi and B. Lee, "Combining LSTM network ensemble via adaptive weighting for improved time series forecasting," *Mathematical Problems in Engineering*, vol. 2018, 8 pages, 2018.

[25] Y. X. Zhao, Z. L. Yang, B. J. Ma, H. H. Song, and D. H. Yang, "Deep learning prediction and model generalization of ground pressure for deep longwall face with large mining height," *Journal of China Coal Society*, vol. 45, no. 1, pp. 54–65, 2012.

[26] Y. ÖZgür, "Penalty function methods for constrained optimization with genetic algorithms," *Mathematical & Computational Applications*, vol. 10, no. 1, pp. 45–56, 2005.

[27] F. Stochino and F. L. Gayarre, "Reinforced concrete slab optimization with simulated annealing," *Applied Sciences*, vol. 9, no. 15, 2019.

[28] T. Wu, X. Shi, L. Liao, C. Zhou, H. Zhou, and Y. Su, "A capacity configuration control strategy to alleviate power fluctuation of hybrid energy storage system based on improved particle swarm optimization," *Energies*, vol. 12, no. 4, 2019.

[29] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," *Advances in Neural Information Processing Systems*, vol. 4, 2012.

[30] J. Zhou, Y. Qiu, S. Zhu, D. J. Armaghani, M. Khandelwal, and E. T. Mohamad, *Estimating of the TBM Advance Rate under Hard Rock Conditions Using XGBoost and Bayesian Optimization*, Underground Space, 2020.

[31] J. Zhou, X. Li, and X. Shi, "Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines," *Safety Science*, vol. 50, no. 4, pp. 629–644, 2012.

[32] J. Zhou, Y. G. Qiu, S. L. Zhu et al., "Optimization of support vector machine through the use of metaheuristic algorithms in forecasting TBM advance rate," *Engineering Applications of Artificial Intelligence*, vol. 97, p. 104015, 2021.

[33] J. Cheng, Z. Wan, S. S. Peng, S. Liu, and Y. Ji, "What can the changes in shield resistance tell us during the period of shearer's cutting and neighboring shields' advance?," *International Journal of Mining Science & Technology*, vol. 25, no. 3, pp. 361–367, 2015.

[34] C. Jing-Yi, Z. Yi-Dong, C. Liang, J. Ming, G. Wei, and G. Lin-Sheng, "Study of loading and running characteristic of hydraulic support in underhand mining face," *Archives of Mining Sciences*, vol. 62, no. 1, 2017.

[35] J. Cheng, Z. Wan, and Y. Ji, "Shield-roof interaction in longwall panels: insights from field data and their application to ground control," *Advances in Civil Engineering*, vol. 2018, 18 pages, 2018.

[36] S. S. Peng, "What can a shield leg pressure tell us?," *Coal Age*, vol. 103, pp. 54–57, 1998.

[37] G. Wang, G. Zhao, and H. Ren, "Analysis on key technologies of intelligent coal mine and intelligent mining," *Journal of China Coal Society*, vol. 44, no. 1, pp. 34–41, 2019.

[38] J. Y. Cheng, Z. J. Wan, S. S. Peng et al., "Technology of intelligent sensing of longwall shield supports status and roof strata based on massive shield pressure monitoring data," *Journal of China Coal Society*, vol. 45, no. 6, pp. 2090–2103, 2020.

[39] D. Park and D. Deb, "Longwall strata control and maintenance system—a stethoscope for longwall mining," *Mining engineering-new york then littleton colorado-*, vol. 51, no. 10, pp. 49–53, 1999.

[40] F. M. Tseng, H. C. Yu, and G. H. Tzeng, "Applied hybrid grey model to forecast seasonal time series," *Technological Forecasting and Social Change*, vol. 67, no. 2-3, pp. 291–302, 2001.

[41] S. Prion and K. A. Haerling, "Making sense of methods and measurement: Pearson product-moment correlation coefficient," *Clinical Simulation in Nursing*, vol. 10, no. 11, pp. 587-588, 2014.

[42] F. Du, R. Yuan, J. Zheng, and G. Song, "Mechanism of abnormal strata pressure of mining under coal pillar in close distance shallow coal seams," *Meitan Xuebao/journal of the China Coal Society*, vol. 42, pp. 24–29, 2017.

[43] H. Zheng, J. Yuan, and L. Chen, "Short-term load forecasting using EMD-LSTM neural networks with a Xgboost algorithm for feature importance evaluation," *Energies*, vol. 10, no. 8, p. 1168, 2017.

[44] J. Zhou, Y. G. Qiu, D. J. Armaghani et al., "Predicting TBM penetration rate in hard rock condition: a comparative study among six XGB-based metaheuristic techniques," *Geoscience Frontiers*, vol. 12, no. 3, p. 101091, 2021.

[45] C. Qi, A. Fourie, and Q. Chen, "Neural network and particle swarm optimization for predicting the unconfined compressive strength of cemented paste backfill," *Construction & Building Materials*, vol. 159, pp. 473–478, 2018.