Multiple Sequence Long and Short Memory Network Model for Corner Gas Concentration Prediction on Coal Mine Workings

Dengke Wang, Lizhen Zhao,* Tianxuan Hao, Yang Du, Jianting Shen, Yiju Tang, Jiupeng Gong, Fan Li, Xiao Yan, Zehua Wang, and Yu Fang

ABSTRACT: To further improve the accuracy of recurrent neural network in predicting the gas concentration in the upper corner of the mine tunnel, this paper proposes a method to construct a gas concentration prediction model based on multiple sequence long and short memory network, considering the spatial correlation between the gas concentration in the return airway and upper corner. The reliability of the model construction is improved by using the white noise test and smoothness test to verify the interpretability of the data in this paper and constructing supervised learning type data for gas concentration prediction model training and testing by means of data set division and data windowing. Through experimental comparison, grid search, and time series decomposition, the model algorithm, training parameters, and experimental results were combined to make an in-depth analysis of the influence of each parameter on the model training and the prediction. A training model of the spatially fused gas concentration prediction model with a network layer of 1 and a number of neurons of 32 as the model structure, Adam as the optimization algorithm, and a learning rate of 0.001 and a batch size of 32 as the training parameters was finally determined. The gas concentration prediction model trained in this paper performed well in the test set with a mean square error (MSE) of 0.0013, and its superiority was verified by comparing it with other models to provide some experience and basis for subsequent studies on gas concentration prediction in the upper corner.

1. INTRODUCTION

China’s coal-based energy resource endowment and stage of economic and social development determine that economic and social development will remain inseparable from coal for quite some time in the future. With the increase of mining depth, the gas content and pressure of coal seam keep increasing, and the gas accumulation in the roadway occurs very easily and causes gas explosion accidents in the production process. It is well known that in the process of coal mining, the surrounding coal rock body and neighboring layer gas fully unloads and gushes into the tunnel, and through the wind flow from the working face to the return tunnel, the gas concentration in the return tunnel is inextricably related to the gas concentration in the upper corner; if we can predict the future gas concentration timely, accurately, and efficiently from the massive data of different spatial and temporal distribution of the tunnel gas monitoring system, it is very important for the prediction and prevention of mine gas disasters.

At present, many experts and scholars at home and abroad have made a lot of meaningful work in the direction of gas concentration prediction, such as the establishment of gas concentration prediction methods based on multivariate distribution lag, support vector machines, BP (back propagation) neural networks, and other models, all of which have achieved good prediction results. However, with the popularization of coal mine monitoring and surveillance systems and the rapid development of big data mining and artificial intelligence in recent years, coal mining enterprises have accumulated huge amounts of gas time series data; recurrent neural networks in deep learning are also very good at solving such serial data problems; and a series of related studies have emerged. For example, Li et al.14 Song et al.,15 Zhang et al.,16 Ma et al.,17 Pan,18 and Fu et al.19 used recurrent neural networks to consider multiple factors and use multiparameter inputs to make predictions of gas concentration. Xun et al.20 combined convolutional neural networks with recurrent neural networks to provide new ideas for obtaining gas concentration time series data features. Zhang et al.21 carried out a multipart prediction model for gas concentration based on LSTM (long—short memory network), which improved the timeliness of prediction. Zhang et al.22

Received: August 12, 2022
Accepted: September 30, 2022
Published: October 13, 2022

http://pubs.acs.org/journal/acsodf
© 2022 The Authors. Published by American Chemical Society
37980
https://doi.org/10.1021/acsomega.2c05188
ACS Omega 2022, 7, 37980−37987
combined wavelet threshold noise reduction with LSTM to increase the performance of the model.

In summary, LSTM is popular among researchers as a recurrent neural network created for solving long series data. However, in the study of prediction models with multifactor inputs, few studies have considered the correlation between gas concentrations in adjacent regions, and the total number of samples used to construct the model is small, which affects the prediction accuracy of the model to some extent. Therefore, in order to further improve the prediction accuracy of the model, this paper uses nearly 60,000 gas concentration monitoring data from the corner of the working face and the return air tunnel of 3105 working face of Hebi no. 8 coal mine as samples to construct a multiseriess LSTM model for the prediction of gas concentration in the corner of the working face of coal mines in order to provide a reliable prediction tool for the prediction of gas concentration in coal mines.

2. THEORY AND MODELING

In order to reveal the law of nonlinear change of mine gas concentration with time, this paper intends to further improve the prediction accuracy of recurrent neural network prediction of gas extraction by enhancing the spatial correlation among the parameters and fusing the gas concentration monitoring data of two regions in the same time period, so as to establish a gas concentration prediction model based on multisequence long–short memory network, as shown in Figure 1.

The prediction model mainly includes the input layer, hidden layer, and output layer; the input data in the input layer includes two time series of corner gas concentration on the working face of the coal mine and gas concentration in the return air tunnel, and the data set needs to be normalized and windowed before the data is passed into the input layer; the hidden layer includes several LSTM layers, and the output of the input layer is passed into the second LSTM layer after the data is passed into the first LSTM layer. The output layer includes a layer of fully connected neural network, which receives the sequence of hidden layer and outputs the predicted value of gas concentration at the corner of the working surface of the coal mine through the fully connected neural network layer.

The work of building prediction models mainly includes data preprocessing, model training, and data prediction. Data preprocessing includes data set partitioning, data windowing, etc. Data set partitioning is to divide the collected data set into a training set, testing set, and validation set in a certain proportion, and the purpose of normalization is to make each parameter get equal treatment; data windowing is to transform the original ultralong time series data into a structure similar to supervised learning; model training is to select the most suitable hyperparameters to minimize the loss function through Adam’s optimization; data prediction is done by inputting data with the same structure as the training to derive predicted values after the model training is completed.

2.1. MLSTM Model Structure. LSTM controls the flow and loss of information by introducing forgetting gate \( f^{(t)} \), input gate \( i^{(t)} \), output gate \( o^{(t)} \), and cellular states \( c^{(t)} \), characterizing long-term memory and candidate states \( \tilde{c}^{(t)} \) to be deposited into long-term memory for control, and these gates avoid the problem of gradient disappearance and gradient explosion after multistage backpropagation.

The forgetting gate (which determines what proportion of information in the cellular state is selectively forgotten) is as follows:

\[
f^{(t)} = \sigma(\tilde{c}^{(t-1)}W_{\text{hf}} + x^{(t)}W_{\text{hf}} + b_f)
\]

(1)

Input gate (determines what percentage of the information will be stored in the current cell state)

\[
i^{(t)} = \sigma(h^{(t-1)}W_{\text{hi}} + x^{(t)}W_{\text{hi}} + b_i)
\]

(2)

Output gate (determines what percentage of information in the cellular state will be stored in short-term memory output)

\[
o^{(t)} = \sigma(h^{(t-1)}W_{\text{ho}} + x^{(t)}W_{\text{ho}} + b_o)
\]

(3)

where \( \sigma \) is the sigmoid activation function, \( x^{(t)} \) is the input at the current moment, \( h^{(t-1)} \) is the short-term memory at the previous moment, \( W_{\text{hf}}, W_{\text{hi}}, W_{\text{ho}} \) and \( W_{\text{ho}} \) are the cycle weight matrices to be trained, representing the proportion of information obtained from the short-term memory at the previous moment for \( f^{(t)}, i^{(t)}, \) and \( o^{(t)} \), respectively, \( W_{\text{hf}}, W_{\text{hi}}, W_{\text{ho}} \) and \( W_{\text{ho}} \) are the input weight matrices to be trained for the proportion of information obtained from the input at the
current moment, and \( b_p, b_q \), and \( b_s \) are the bias terms to be trained. The short-term memory can be expressed as

\[
h^{(t)} = \sigma^{(t)} \odot \tanh(c^{(t)})
\]

where \( \sigma \) is the activation function and \( c^{(t)} \) is the cell state at the current moment, which can be expressed as

\[
c^{(t)} = c^{(t-1)} + i^{(t)} \odot \tilde{c}^{(t)}
\]

where \( c^{(t-1)} \) is the cellular state at the previous moment and \( \tilde{c}^{(t)} \) is the candidate cellular state, which can be expressed as

\[
\tilde{c}^{(t)} = \tanh(h^{(t-1)}W_{hc} + x^{(t)}W_{xc} + b_c)
\]

where \( W_{hc} \) is the loop weight matrix to be trained, \( W_{xc} \) is the input weight matrix of the fully connected neural network, and \( b_c \) is the bias term to be trained.

LSTM has been successfully used in numerous fields such as unconstrained handwriting recognition, speech recognition, handwriting generation, machine translation, generating captions for images, and parsing.\(^{24}\)

### 2.2. Calculation of the Loss Function.

In this paper, the short-term memory \( h^{(t)} \) derived from each moment of the long-short memory network before the last layer of the mine gas concentration prediction model is directly output to the next layer of the long-short memory network, while the last layer of the long-short memory network only passes the output \( H^{(T)} \) of the last moment into a layer of the fully connected neural network.

The output \( H^{(T)} \) of the last layer of the long-short memory network can be expressed as

\[
H^{(T)} = \sigma^{(T)} \odot \tanh(c^{(T)})
\]

Its use as input gives the hidden layer of the fully connected neural network layer:

\[
F = H^{(T)}W_{Fy} + b_f
\]

where \( F \) is the hidden layer matrix of the fully connected neural network and \( W_{Fy} \) is the input weight matrix of the fully connected neural network. The output of the fully connected neural network can be expressed as

\[
\hat{y} = FW_{fy} + b_y
\]

where \( \hat{y} \) is the predicted value of the output of the fully connected neural network, \( W_{fy} \) is the weight matrix of the output \( \hat{y} \) to be trained of the fully connected neural network, and \( b_y \) is the bias term to be trained of the fully connected neural network.

MSE (mean square error) is often used as a loss function for regression problems to measure how well the predicted values match the true values. The loss function is calculated directly with the following equation:

\[
J = \frac{1}{n} \sum_{i=1}^{n} (\hat{y} - y)^2
\]

where \( J \) is the total loss function and \( y \) is the true value.

### 2.3. Gradient Calculation for Back Propagation over Time.

Then, the gradient of the output \( \hat{y} \) can be derived as

\[
\frac{\partial J}{\partial \hat{y}} = \hat{y} - y
\]
The gradients of the weights to be trained in the output gate $o^{(i)}$ are
\[ \frac{\partial J}{\partial W_{ho}} = \sum_{t=1}^{T} [\delta_h^{(t)} \tanh(c^{(t)}) o^{(t)} (1 - o^{(t)}) ]^T h^{(t-1)} \]  
(19)

The optimization problem in this paper can be simply formulated as follows: finding a set of parameters $W$ and $b$ on the neural network by the gradients $\partial J / \partial W$ and $\partial J / \partial b$ of each weight to be trained computed in the previous section, which significantly reduces the cost function $J(W, b)$. Adam is a learning rate adaptive optimization algorithm capable of updating neural network weights iteratively based on training data. The parameters suggested by its proposers in the original paper are chosen as follows: step size (learning rate) $L_r = 0.001$, exponential decay rate of moment estimation $\rho_1 = 0.9$, $\rho_2 = 0.999$, and a small constant $\sigma = 10^{-8}$ for numerical stabilization. The update of its weights to be trained requires initializing the first-order moment variable $s = 0$, the second-order moment variable $r = 0$, and the time step $t = 0$. The cycle starts after the initialization is completed. The gradient of the weights to be trained is calculated and the time step is added by 1 to update the biased first-order moment estimates and biased second-order moment estimates:
\[ s^{(t+1)} = \rho_s s^{(t)} + (1 - \rho_s) \frac{\partial J}{\partial \omega} \]  
(20)
\[ r^{(t+1)} = \rho_r r^{(t)} + (1 - \rho_r) \left( \frac{\partial J}{\partial \omega} \odot \frac{\partial J}{\partial \omega} \right) \]  
(21)

Correction of deviation of first-order moments from second-order moments
\[ \hat{s}^{(t+1)} = s^{(t+1)} / (1 - \rho_s) \]  
(22)
\[ \hat{r}^{(t+1)} = r^{(t+1)} / (1 - \rho_r) \]  
(23)

Update the parameters to be trained
\[ W^{new} = W - L_r \left( \frac{\hat{s}^{(t+1)}}{\sqrt{\hat{r}^{(t+1)} + \sigma}} \right) \]  
(24)

3. EXPERIMENT SECTION

3.1. Data Sample. In order to have enough samples for model training, the author concentrated on collecting 29,742 pieces of gas concentration monitoring data from May 5, 2021 to August 18, 2021 for the corner angle gas concentration monitoring data on 3105 working face and 29,742 pieces of gas concentration monitoring data on the return air lane of Hebi no. 8 coal mine, totaling 59,484 pieces (the extraction time interval of each two pieces of data was 5 min). The results of the smoothness test and white noise test of gas concentration monitoring data show that the $p$-value of both tests is 0 (indicating that the series is considered as a smooth series with 100% confidence level and is not a pure random series), which is less than the significance level of 0.05. Therefore, the time series data of gas concentration in the backwind lane and upper corner used in this paper are smooth and nonwhite noise and are interpretable.

Therefore, this paper constructs the original data into a supervised learning sample structure by adding windows to the data (principle as in Figure 2), so that 96 upper corner gas concentration monitoring data and 96 return airway gas concentration monitoring data are used as input features, and the 97th upper corner gas concentration monitoring data is used as the corresponding label, generating a total of 29,646 data samples, of which 20,000 sets are training data and 9646 sets are test data.

3.2. Tuning of Learning Rate. The learning rate is an important parameter; too large or too small will cause the model training difficult to converge. In this paper, we first find the maximum learning rate suitable for this model by the dynamic learning rate method to ensure the model training effect and accelerate the convergence of the model. The learning rate for training is set as follows:
\[ L_r = 10^{(epoch/20) - 8} \]  
(25)

where epoch is the number of iterations of model training; the learning rate will increase with the increase of training epoch and stop training using the early stop method (stop training when the model validation loss no longer improves because of the increase in learning rate); the learning rate here is the appropriate learning rate for this model. As shown in Figure 3, the loss value decreases during the model training iterations, and the loss stops decreasing when the learning rate is $10^{-3}$, which is the recommended learning rate for Adam’s algorithm, and stops training when the learning rate is $10^{-2}$ by the early stop method. In order to make the model have the best training effect, $10^{-3}$ is used as the learning rate for model construction in this paper.

3.3. Tuning of Learning Rate Batch Size and Neuron Number Tuning Based on Grid Search Method.

To experimentally derive the optimal combination of batch size and number of neurons parameters, this paper first set the learning rate to 0.01 as suggested by Goodfellow et al., 24 the proposers of Adam, in their original paper, and used the grid search method for the batch size ($16, 32, 48, 64, 80, 128$) and the number of neurons ($16, 32, 48, 64, 80, 128$) for cross-validation was performed to compare the minimum test set.

Figure 2. Data set windowing.
MSE (mean square error) that could be achieved after training using each set of parameters, and the results are shown in Figure 4. It can be seen from Figure 4 that when the batch size and number of neurons are combined as (16, 64), (16, 80), and (32, 32) during training, the test set has the smallest MSE with the true value and the best prediction. The number of neurons is larger when the batch size and the number of neurons are set to (16, 80) than (16, 64), which consumes more computational resources, so this combination of parameters cannot be considered.

To verify the differences between the two parameter combinations, the two sets of parameters were trained separately again and stopped by the early stop method; the variation pattern of the error with the number of iterations was obtained for the training; and the differences in the parameters and running times between the two trainings are shown in Table 1. In the case of achieving the same predicted MSE, the training time for the (32, 32) combination is 7 min 50 s and the training time for the (16, 64) combination is 14 min 15 s. The number of iterations for both is similar, 62 and 58. The reason for the difference in training time between the two is that the backward calculation of the error and updating the weights to be trained frequency of the model training is 1250 times when the batch size is equal to 16. In the long–short memory network, the number of neurons is the number of short-term memory, and there are four weights to be trained related to short-term memory in each layer of the long–short memory network, resulting in about four times more weights to be updated when the number of neurons is 64 than when the number of neurons is 32, which are 21,153 and 83,265, respectively. Therefore, with similar prediction results and considering the performance consumption of model training, the model parameter batch size and the number of neurons are set to 32 in this paper.

3.4. Layer Tuning of Neural Networks. However, the deeper the network is, the more computational resources will be used for training the model, resulting in higher hardware configuration and longer computation time required for model construction; the deeper the network layer is, the more parameters to be trained will be included in the neurons, which will also easily lead to overfitting. To find the number of network layers suitable for the gas concentration prediction model, four network layers (1, 2, 3, 4) were trained one by one to examine the network performance.

The specific performance of the four network layers is shown in Table 2. The results show that although the errors of the four network layers are decreasing during the training process, the validation errors have rebounded to a certain extent after decreasing to a certain degree, which indicates that all four network layers have a certain degree of overfitting phenomenon in the training, and it is not difficult to find that the situation is significantly better when the number of network layers is 1 than when the numbers of network layers are 2, 3, and 4.

In terms of the training MSE (mean square error), the network reaches 0.0011 for either three or four layers, which is better than 0.0012 for two layers and 0.0013 for one layer. However, the MSE of the test set does not decrease as much as the training MSE for three or four layers but increases from 0.0014 to 0.0015 for one or two layers, and the number of parameters to be updated gradually increases with the number of layers, which leads to a longer training time. It can be seen that when the number of layers of the neural network is 1, the model consumes less time for training, has the smallest validation MSE, is the least likely to be overfitted, and has the most stable MSE during training because of the few optimization parameters.

![Figure 3. Variation of training error under different learning rates.](image)

![Figure 4. The grid search method was used to cross-verify batch size and number of neurons.](image)

| batch size | number of neurons | number of weights | training time | iterations | number of parameter updates | MSE of test set |
|------------|------------------|------------------|--------------|------------|---------------------------|----------------|
| 32         | 32               | 21,153           | 470 s        | 62         | 625                       | 0.0015         |
| 16         | 64               | 83,265           | 855 s        | 58         | 1250                      | 0.0015         |

| number of network layers | number of weights | MSE of training set | MSE of test set | training time |
|--------------------------|-------------------|---------------------|-----------------|--------------|
| 1                        | 4513              | 0.0013              | 0.0014          | 390 s        |
| 2                        | 12,833            | 0.0012              | 0.0014          | 794 s        |
| 3                        | 21,153            | 0.0011              | 0.0015          | 1077 s       |
| 4                        | 29,473            | 0.0011              | 0.0015          | 1324 s       |
Through the above analysis, a mine gas concentration prediction model with a 1-layer long—short memory network and a number of neurons of 32 was finally selected, and the model was trained with a learning rate of 0.001 and a batch size of 32 using Adam as the optimization algorithm. The MSE obtained for the final obtained mine gas concentration spatial fusion prediction model in the test was 0.0013.

4. RESULTS AND DISCUSSION

4.1. Validation of Prediction Effect. To verify the advancedness of the prediction effect of the model constructed in this paper, the gas concentration prediction effect of the model in two recent related papers was compared with the prediction effect of this paper, as shown in Table 3, where the number of samples denotes the number of data samples used to construct the model, and the number of spatial features denotes the feature parameters of several regions considered in the input data.

From the data in the table, it can be seen that literature 14 constructs the model considering the relationship between the gas concentration in three areas, namely, the return air tunnel, the working face, and the upper corner, and the number of spatial features is 3, which is more than this paper. However, its sample size is only half of that of this paper, resulting in a higher MSE of prediction results than this paper. The model in literature 25 does not consider the influence of gas concentration in adjacent areas; the number of spatial features is 1; and the number of samples is only half of that in this paper, so the predicted MSE is higher than that in both literature 25 and this paper. In summary, the MLSTM gas concentration prediction model constructed in this paper has a certain degree of improvement over the other two papers, and the MSE of the prediction results is 90 and 23.5% lower than that of literature 25 and 14, respectively.

| number of network layers | number of weights | MSE of training set | MSE of test set | training time |
|--------------------------|------------------|---------------------|----------------|--------------|
| 1                        | 45,13            | 0.0013              | 0.0014         | 390 s        |
| 2                        | 12,833           | 0.0012              | 0.0014         | 794 s        |
| 3                        | 21,153           | 0.0011              | 0.0015         | 1077 s       |
| 4                        | 29,473           | 0.0011              | 0.0015         | 1324 s       |

4.2. Predictive Effect Analysis. The prediction effect of the test set is shown in Figure 5, from which it can be seen that the model has a good prediction effect for the gas concentration in the upper corner. To analyze the prediction effect of the mine gas concentration prediction model in depth, the latter 1742 data of the test set data were decomposed in a time series with 1 day (288 data) as shown in Figure 6. The time series usually contains three parts: the trending part, (c) periodic variations, and (d) irregular variations.

Among them, the long-term trend refers to a continuous trend that manifests itself over a considerable period of time, as shown in Figure 6b; cyclical variation refers to a short-term cyclical fluctuation of fixed length and magnitude formed by the influence of cyclical changes in the time series data, as shown in Figure 6c; irregular variation is formed by the influence of chance factors, as shown in Figure 6d. These three factors can affect the development and change of time series through different combinations.

The decomposed long-term trend section shows that the gas concentration monitoring data used in this paper has a trend of decreasing, then increasing, then decreasing gas concentration between 0.3 and 0.38%. For the periodic variation, the data segments in Figure 6 have regular short-term periodic fluctuations with a 1 day cycle and a range of −0.1 to 0.3%. For the random section, the high points of gas concentration in the upper corner all belong to this. For the random part, the high points of gas concentration in the upper corner are all part of this part, which has fluctuations in the range of −0.5 to 1.25%.

It can be seen that the prediction of the model is more accurate at low gas concentration and can fit the trend of its concentration completely because this part of the data has good trend and periodicity, while the prediction at the point of sudden change of gas concentration is not as good as the former because this part of the data has randomness, but the model also makes a corresponding degree of fitting to this part and has some effect on the prediction of gas concentration exceeding the limit.

5. CONCLUSIONS

(1) In this paper, considering the spatial distribution relationship between the upper corner on the working face of coal mines and the gas concentration in the return airway, a multisequence input long and short memory model for predicting the gas concentration in the corner angle on the working face of coal mines is...
constructed, and the back propagation update principle of this model is derived to provide reference for future related research work.

(2) The white noise test and smoothness test are used to verify the interpretability of the data in this paper, and supervised learning-like data for training and testing of gas concentration prediction models are constructed by means of data set partitioning and data windowing to ensure the reliability of the models.

(3) By combining the model algorithm, training parameters, and experimental results through experimental comparison, grid search, and time series decomposition, the influence of each parameter on the training and prediction of the model was analyzed in depth, and finally, the gas concentration prediction model with Adam as the optimization algorithm, a learning rate of 0.001, and a batch size of 32 as the training parameters performed well in the test set. The final prediction MSE is 0.0013, which is expected to provide reference for future research on gas concentration prediction in coal mines based on recurrent neural networks.

■ AUTHOR INFORMATION

Corresponding Author
Lizhen Zhao — College of Safety Science and Engineering, Henan Polytechnic University, Jiaozuo 454000, China; orcid.org/0000-0001-8883-9799; Email: safezlz@foxmail.com

Authors
Dengke Wang — State Key Laboratory of Mining Response and Disaster Prevention and Control in Deep Coal Mines, Anhui University of Science and Technology, Huainan 232001, China; College of Safety Science and Engineering, Henan Polytechnic University, Jiaozuo 454000, China; State Collaborative Innovation Center of Coal Work Safety and Clean-efficiency Utilization, Jiaozuo, Henan 454000, China; State Key Laboratory Cultivation Base for Gas Geology and Gas Control, Henan Polytechnic University, Jiaozuo, Henan 454000, China
Tianxuan Hao — College of Safety Science and Engineering, Henan Polytechnic University, Jiaozuo 454000, China; State Collaborative Innovation Center of Coal Work Safety and Clean-efficiency Utilization, Jiaozuo, Henan 454000, China; State Key Laboratory Cultivation Base for Gas Geology and Gas Control, Henan Polytechnic University, Jiaozuo, Henan 454000, China
Yang Du — College of Safety Science and Engineering, Henan Polytechnic University, Jiaozuo 454000, China
Jianting Shen — Dafang Lvtang Coal Mine Co., Ltd., Bijie, Guizhou 551604, China
Yiju Tang — College of Safety Science and Engineering, Henan Polytechnic University, Jiaozuo 454000, China
Jiupeng Gong — Dafang Lvtang Coal Mine Co., Ltd., Bijie, Guizhou 551604, China
Fan Li — College of Safety Science and Engineering, Henan Polytechnic University, Jiaozuo 454000, China; orcid.org/0000-0003-4632-5663
Xiao Yan — Dafang Lvtang Coal Mine Co., Ltd., Bijie, Guizhou 551604, China
Zehua Wang — College of Safety Science and Engineering, Henan Polytechnic University, Jiaozuo 454000, China
Yu Fang — College of Safety Science and Engineering, Henan Polytechnic University, Jiaozuo 454000, China

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.2c05188

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was funded by the National Natural Science Foundation of China (nos. 52174174 and 52274193), the Special Program for Basic Research of Key Scientific Research
Projects of Colleges and Universities in Henan Province of China (no. 21zx004), the Science and Key R & D promotion projects of Henan Province (nos. 222102320172 and 222102320413), and the Innovative Scientific Research Team of Henan Polytechnic University in China (T2022-1).

REFERENCES

(1) Xie, H. P.; Ren, S. H.; Xie, Y. C. Development opportunities of the coal industry towards the goal of carbon neutrality [J/OL]. J. China Coal Soc. 2021, 46, 07−15.
(2) Ding, Y.; Zhu, B.; Li, S. G.; Lin, H. F.; Wei, Z. M.; Li, L. M.; Long, H. Accurate identification and efficient drainage of relieved methane in goaf of high outburst mine. J. China Coal Soc. 2021, 46, 3565−3577.
(3) Jufeng, Z.; Shiliang, S.; Yi, L.; Bo, Y.; Fanghua, W.; Kuan, W. Intelligent early warning system of gas and coal spontaneous combustion disaster based on big data: data characteristics, application structure and key technologies. China Saf. Sci. J. 2021, 31, 60−66.
(4) Zhang, J.; Xu, K.; Wang, B.; Wang, Y. Extraordinarily serious gas explosion accidents in coal mines: analysis of causes and research on management mode. China Saf. Sci. J. 2016, 26, 73−78.
(5) Yang, L.; Liu, H.; Mao, S. J.; Shi, C. Dynamic prediction of gas concentration based on multivariate distribution lag model. J. China Univ. Min. Technol. 2016, 45, 455−461.
(6) Yan, H. Research on Mine Gas Accident Risk Prediction Technology Based on Support Vector Machine. Taiyuan University of Technology. 2018.
(7) Wang, K.; Wang, J. Prediction of Gas Concentration based on SVM Optimized by Memetic Algorithm. Coal. 2020, 29, 4−6.
(8) Li, D.; Sun, Z.; Li, M.; Hou, Y.; Mao, S.; Niu, Y. AWLSSVM Gas Prediction Research Based on Chaotic Particle Swarm Optimization. Saf. Coal Mines. 2020, 51, 193−198.
(9) Wei, L.; Bai, T.; Fu, H.; Yi, Y. New gas concentration dynamic prediction model based on the EMD-LSSVM. J. Saf. Environ. 2016, 16, 119−123.
(10) Qian, J.; Qiu, C.; Li, Z.; Wu, X. Gas Emission Quantity Prediction Based on Particle Swarm Optimization of SVM and Deep Learning Network. Saf. Coal Mines 2016, 47, 173−176.
(11) Zhang, W. Research on Gas Concentration Prediction Based on Support Vector Machine and Immune Genetic BP. Xi’an University of Science and Technology. 2017.
(12) Yao, Q.; Qiu, B. Mine Gas Concentration Prediction Algorithm Based on Improved BP Neural Network. Coal Technol. 2017, 36, 182−184.
(13) Liu, Y.; Zhao, Q.; Hao, W. Study of Gas Concentration Prediction Based on Genetic Algorithm and Optimizing BP Neural Network. Min. Saf. Environ. Prot. 2015, 42, 56−60.
(14) Li, S. G.; Ma, L.; Pan, S. B.; Shi, X. Research on prediction model of gas concentration based on RNN in coal mining face. Coal Sci. Technol. 2020, 48, 33−37.
(15) Song, S.; Li, S.; Zhang, T.; Ma, L.; Pan, S.; Gao, L. Research on a Multi-Parameter Fusion Prediction Model of Pressure Relief Gas Concentration Based on RNN. Energies 2021, 14, 1384−1402.
(16) Zhang, T.; Song, S.; Li, S.; Ma, L.; Pan, S.; Han, L. Research on Gas Concentration Prediction Models Based on LSTM Multidimensional Time Series. Energies 2019, 12, 161−176.
(17) Ma, L.; Pan, S.; Dai, X.; Song, S.; Shi, X. Gas concentration prediction model of working face based on PSO-Adam-GRU. J. Xi’an Univ. Sci. Technol. 2020, 40, 363−368.
(18) Pan, S. Research on Multi Parameter Fusion Prediction of Gas Concentration Based on RNN. M.S. Xi’an University of Science and Technology. 2020.
(19) Fu, H.; Liu, Y.; Xu, N.; Zhang, J. Research on Gas Concentration Prediction Based on Multi-Sensor-Deep Long Short-Term Memory Network Fusion. Chin. J. Sens. Actuators 2021, 34, 784−790.
(20) Xun, X.; Su, C.; Li, W.; Pang, M. Coal Mine Gas Concentration Prediction Based on CNN-LSTM. Mod. Inf. Technol. 2020, 4, 149−152.
(21) Zhang, Z.; Zhu, Q.; Li, Q.; Zhang, E.; Liu, H. Construction and application of ARIMA prediction model of gas concentration based on Python. J. North China Inst. Sci. Technol. 2020, 17, 23−28.
(22) Zhang, X.; Liu, F.; Li, X. Coal Mine Gas Concentration Prediction Based on Wavelet Denoising and Recurrent Neural Network. Coal Technol. 2020, 39, 145−148.
(23) Hochreiter, S.; Schmidhuber, J. Long Short-term Memory. Neural Comput. 1997, 9, 1735−1780.
(24) Goodfellow, I., Bengio, Y., Courville, A. Deep Learning. Posts & Telecom Press: Beijing, 2017; pp. 248−250.
(25) Jia, P.; Zhang, Z.; Liang, R.; Liu, H.; Miao, Y. Gas Concentration Prediction Method Based on PSO-CNN-aBiGRU. Min. Res. Dev. 2021, 41, 76−81.
(26) Fang, X. Python Data Mining Practice. Publishing House of Electronics Industry, Beijing, 2021; pp. 123−135.