

# Power Supply Coal Consumption Prediction Model Based on Long and Short-term Memory Neural Network

Mengjie Liu\*, Chenggang Zhen

School of Control and Computer Engineering, North China Electric Power University, Baoding 071003, China

\* Corresponding author: Mengjie Liu

**Abstract:** As a measure of the overall economic performance of thermal power plants and an important component of the variable cost of power generation, the prediction of coal consumption of power supply is of great significance for the bidding decision of power generation companies. In this paper, we propose an improved long and short-term memory (LSTM) model for calculating the coal consumption of thermal power units, using a large amount of data stored in the plant monitoring information system (SIS). Firstly, the data are pre-processed using the threshold determination method to filter out the data of stable operating conditions. Secondly, the maximum correlation minimum redundancy (mRMR) algorithm is used to determine the optimal set of special features. The results show that the prediction effect of the improved LSTM-based power supply coal consumption calculation model proposed in this paper is better than other models, and the calculation accuracy is higher, which is suitable for power supply coal consumption calculation.

**Keywords:** Power supply coal consumption; MRMR; LSTM.

## 1. Introduction

At present, the real-time calculation of power supply coal consumption mainly adopts data mining method. Based on the support vector machine technology, the literature [1] analyzed the sensitivity coefficients of the operating parameters on the coal consumption under different load conditions and provided guidance for the optimal operation adjustment of the unit. In the literature [2], the BP algorithm was used to establish the prediction model of coal consumption with strong dynamic response performance, but the coal consumption characteristics under different load conditions were not further studied. In the literature [3], the energy efficiency indexes that have a major impact on the coal consumption of power supply were selected by using the gray correlation analysis method and combined with the generalized regression neural network to predict the coal consumption of power supply. In the literature [4], a discrete model between load and coal consumption rate was established by using fuzzy clustering, rough set theory and other data mining methods to predict the current coal consumption characteristics of the unit and achieve the optimal load allocation of the unit in real time. In the literature [5], the mRMR algorithm was used to filter the optimal subset of features by ranking the importance of the initial input variables, which improved the prediction accuracy of the model. In the literature [6], a black box model was established based on the long short-term memory (LSTM) neural network, and the comparative validation showed that the LSTM model outperformed the SVR and BP models in all evaluation indexes.

In this paper, a 600MW domestic supercritical coal-fired unit is used as the research object, the stable working condition data is screened by the threshold determination method, the optimal set of special features is determined by the mRMR algorithm, and finally the power supply coal consumption prediction model is established based on LSTM. The application of the algorithm model was carried out with an ultra-supercritical million wet-cooled unit as the research

object.

## 2. Methodology

### 2.1. Screening of steady state data

When the unit is in variable operating conditions, fluctuations in operating parameters occur. Therefore, the collected data need to be pre-processed, i.e., stability judgment, and the data of non-stable operating conditions need to be excluded[7]. The determination of the stable operating conditions of the unit needs to be based on the important operating parameters that characterize the operating status of the unit. If the unit load, feedwater flow, reheat steam temperature, exhaust smoke temperature, and other parameters are all tested for steady state, sufficient stable data will not be available for model training. Therefore, in this paper, unit load, main steam pressure and reheat steam temperature are selected for steady-state condition determination.

The data for the steady-state condition were screened by the threshold method. Within 60 min, the difference between the maximum and minimum values of the unit load, main steam pressure and reheat steam temperature is less than a certain threshold, which is considered to be in steady-state conditions. Otherwise, the sampling slide window is pushed back clockwise for 5 min, and the steady state is re-judged until the sample meets the requirements.

### 2.2. MRMR-based key feature extraction

In probability theory and information theory, the concept of mutual information is introduced to measure the degree of interdependence between two random variables, reflecting both linear and nonlinear correlations between the two sets of variables. For 2 sets of discrete random variables  $X$  and  $Y$ , the mutual information between them is  $I(X;Y)$ .

The formula is as follows :

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)} \quad (1)$$

The mRMR calculates the correlation between each input variable and the output variable based on the mutual information theory, while taking into account the correlation between each input variable. The mRMR algorithm can find the subset of features from the initial input variables that have the highest correlation with the output variables and the lowest correlation between the input variables.

$$\max \left( \frac{1}{n} \sum_{X_i \in S_n} I(X_i; Z) - \frac{1}{n^2} \sum_{X_i, X_j \in S_n} I(X_i; X_j) \right) \quad (2)$$

An incremental search is used to sequentially select the optimal feature subset, which satisfies the following equation.

$$\max \left( \left| I(X_i; Z) - \frac{1}{m-1} \sum_{X_j \in S_{m-1}} I(X_i; X_j) \right| \right) \quad (3)$$

The mRMR algorithm was used to rank the degree of influence of all features and added to the feature subset in order. To eliminate the difference in magnitude between variables, the sample data need to be normalized by the formula:

$$M_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (4)$$

The final feature subset is best predicted by the prediction model when unit load, exhaust oxygen, condenser vacuum, total coal feed, ambient temperature, and low-level heat of incoming coal are taken.

### 2.3. LSTM-based coal consumption prediction for power supply

Long-term memory in learning is difficult to be preserved due to the gradient explosion or gradient disappearance problem of Recurrent Neural Network (RNN). In 1997, Hochreiter et al. first proposed Long-Term and Short-Term Memory Network, LSTM [8].

LSTM is developed on the basis of RNN recurrent neural network, which is a gated recurrent neural network. The internal structure of LSTM is shown in Figure 1. The key of LSTM is the memory cells, which have states throughout the network with only a small amount of information interaction, making the information stable during the long-term propagation. LSTM is implemented through input gates, forgetting gates and output gates to achieve protection and control of gradient information. Where the forgetting gate determines which information is discarded or retained from the cell state; the input gate updates the cell state; and the output gate determines the final output result.

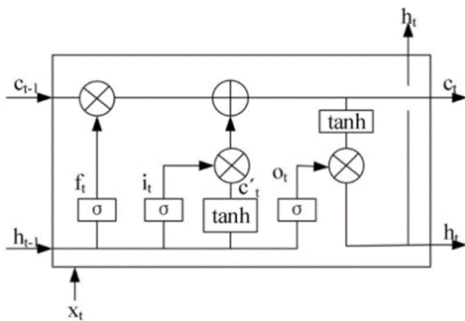


Figure 1. The structure of LSTM network

The equation for calculating each part of the neuron at time t is shown below.

$$f_t = \sigma(W_{sf}X_t + W_{hf}h_{t-1} + b_f) \quad (5)$$

$$i_t = \sigma(W_{si}X_t + W_{hi}h_{t-1} + b_i) \quad (6)$$

$$\begin{cases} c'_t = \tanh(W_{xc}X_t + W_{hc}h_{t-1} + b_c) \\ c_t = f_t * c_{t-1} + i_t * c'_t \end{cases} \quad (7)$$

$$o_t = \sigma(W_{xo}X_t + W_{ho}h_{t-1} + b_o) \quad (8)$$

$$h_t = o_t * \tanh(c_t) \quad (9)$$

## 3. Results and discussion

Figure 2 shows the comparison of the prediction effect of LSTM with error back propagation (BP) and support vector machines (SVM). The structure of LSTM in this paper is: 3-layer structure, optimizer is Adam, loss function is MSE, number of nodes in hidden layer is 20, learning rate is 0.001, 100 iterations. Compared with BP and SVM, LSTM has better prediction results.

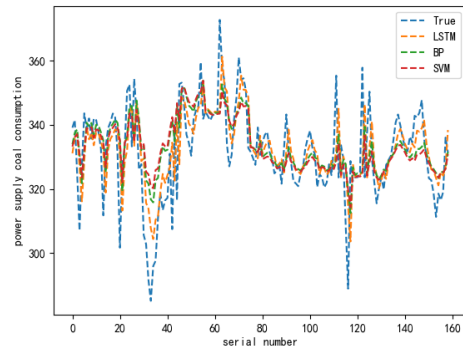


Figure 2. Prediction effect of three models

Table 1 shows the comparison of the prediction accuracy of the two networks.

Table 1. Comparison of prediction accuracy

models	MAE	RMSE	MAPE%
LSTM	7.619	10.552	2.321
BP	8.274	11.520	2.539
SVM	8.505	11.919	2.610

The calculated value of LSTM model of coal supply consumption of the unit is in good agreement with the actual value. The low load operation of the unit is unstable and the parameters fluctuate greatly, resulting in a relatively high error in the calculation of power supply coal consumption. The experimental results show that the prediction model used in this paper can achieve the calculation of power supply coal consumption in thermal power plants better.

## 4. Conclusion

LSTM is an efficient deep learning model that can handle large amounts of data efficiently. It is proven to be more accurate and applicable by comparing with other prediction methods. For the selection of the input parameters of LSTM, a maximum small-off minimum redundancy method is used to select the influences with high relevance to form the optimal set of special features. The prediction error is less than 3%, and the accuracy of the predicted coal consumption of thermal power units is high.

## References

- [1] Liu B.H., Fu Z.G., Wang P.K., et al (2018). Research on the application of big data mining data technology in the analysis of energy consumption of coal-fired power plant units. Proceedings of the CSEE, vol.12, no.38, p.3578-3587.

- [2] Du G., Liao L., Zhang D.L., Yang K (2011). Online simulation of coal consumption in thermal power plants based on BP neural network. *Computer Simulation*, vol.4, no.28, p.313-344.
- [3] Xu J.,Gu Y.J.,Wang Z.,Li Q.Q.,Yang N (2017). Research on energy efficiency characteristic indexes of coal power units and their benchmark values based on data mining. *Proceedings of the CSEE*, vol.7, no.37, p. 2009-2016.
- [4] Zeng D.L., Yang T.T., Cheng X, et al (2010). Application of data mining methods in real-time plant-level load optimization allocation. *Proceedings of the CSEE*, vol.11, no.30, p. 109-114.
- [5] Jin X.Z.,Liu Y.,Zhao W.J.,Yu J (2022). Prediction of SO<sub>2</sub> mass concentration at the outlet of thermal power plants based on mRMR and MA-RELM. *Journal of Chinese Society of Power Engineering*, vol.7, no.42, p.664-676.
- [6] Zheng T.X., He J., Zhang L.B (2020). LSTM neural network-based combustion timing estimation for hybrid fuel HCCI engines under complex operating conditions. *Chinese Journal of Scientific Instrument*, vol.10, no.41, p.100-110.
- [7] Opri I, Victor-Eduard C , Norior M , et al (2020). Parametric optimization of the thermodynamic cycle design for supercritical steam power plants. *Energy Conversion and Management*, no.208, p. 112587.
- [8] Liu Y, Li Q, Duan X, Zhang Y, Yang Z, Che D (2018). Thermodynamic analysis of a modified system for a 1000 MW single reheat ultra-supercritical thermal power plant. *Energy*, no.145, p. 25–37.r