Research on Main Steam Temperature Prediction Model Based on Improved LSTM Algorithm

Ling Zheng, Xin Ye and Fei Chen
Department of Computer Science and Technology, North China Electric Power University, Faculty of Control and Computer Engineering, room E1112, No. 2 Beinong Road, HuiLongguan Town, Changping District, Beijing, China
Email: 450377299@qq.com

Abstract. In order to improve the reliability and accuracy of the main steam temperature trend prediction, a main steam temperature prediction model based on improved LSTM is proposed. Firstly, uses the grey correlation analysis method to select the important influencing factors. Then, a linear structure is introduced into the LSTM structure to construct a main steam temperature prediction model. Finally, based on the historical operating data of the thermal power unit, a simulation experiment is performed to compare the prediction error of the output of the RNN model, the LSTM model, and the improved LSTM model. The results show that the method has higher prediction accuracy for the main steam temperature. At the same time, compared with other traditional methods, this method has better fitting effect, which can be well applied in practical engineering.

Keywords. Main steam temperature; grey correlation analysis; deep learning; improved LSTM algorithm; regression prediction.

1. Introduction
The main steam temperature is a typical large delay and large inertia object which has many influencing factors, so the design of an advanced and reasonable thermal power unit coordinated control system needs to master the dynamic characteristics of the object in advance. PID is one of the earliest developed control strategies in the thermal control system. Various types of PID controllers still have a dominant position in the thermal process control due to their clear physical meaning, easy adjustment and robustness. However, because the controlled object has characteristics such as ignorance, time-varying and delay, the traditional PID controller is not very effective in the control process of practical application. With the increasing enrichment and optimization of computer algorithms, scholars have introduced advanced control technologies such as intelligent control and predictive control into the field of process control, which have achieved fruitful results. Wu [1] combined fuzzy control and PID control to design a PID controller to control the main steam temperature; Jin [2] built a Support Vector Machine (SVM) model to predict the main steam temperature; Fang [3] used improved genetic algorithm to optimize PID controller parameters online; Wang [4] combined the BP network with ant colony algorithm to adjust PID parameters.

Recurrent Neural Network (RNN) is a kind of neural network specially used to process time series data. The input of the hidden layer includes the output of the input layer and the same hidden layer at the previous moment, which is its most obvious feature. This series network structure can effectively maintain the dependency relationship between data, which is very suitable for time series data prediction. Due to problems of “gradient explosion” and “gradient disappearance” of traditional RNN,
Long Short-Term Memory (LSTM) network is a variant structure of RNN, which is designed based on the idea of gating to avoid long-term dependence problems. It is characterized by being able to remember long-term information and realize the optimization of processing long sequence data. However, neural networks are fitted to the linear part of the model by superimposing and nesting nonlinear functions, which leads to a reduction in the overall convergence speed and prediction accuracy. In order to solve this problem, Pao [5] proposed a random vector functional link net, which is a neural network that combines nonlinear and linear and can be used in pattern recognition, system identification and real-time control.

Given that the controlled objects all have obvious timing correlation, a main steam temperature prediction model based on improved LSTM is constructed for the thermal power Modulation Control System (MCS). After the model is verified, it is combined with Model Predictive Control (MPC) technology to coordinate control the main steam temperature. This model adds a linear fitting part to the LSTM network, which makes the improved network have faster convergence speed and approximation ability in addition to long-term sequence processing ability, thereby improving the adaptability of the network. This paper first analyzes the main influencing factors by grey correlation analysis algorithm, then analyzes advantages and limitations of RNN and LSTM, and constructs a model based on improved LSTM. The historical data provided by the operating unit is taken as an example to verify the superiority and feasibility. The results show that prediction accuracy of the improved LSTM model is significantly better than the traditional RNN model and slightly better than the LSTM model.

2. Analysis of Influencing Factors
The main steam temperature predicted by this model is the steam temperature before the second desuperheating of the superheater. The factors affecting the steam temperature are complex, but the main process variables are the first desuperheating water flow, main steam flow, fuel quantity, air volume, etc. This paper selects 5100 historical data of each of the 11 variables and main steam temperature from operating units of a power plant and calculates the correlation degree between each factor and the main steam temperature by grey correlation analysis method, then analyzes influencing factors by comparing the correlation degree.

2.1. Method
Grey correlation analysis method [6] can quantitatively describe the change trend and comparison of the system. The degree of correlation between curves is determined by the geometric similarity between the reference data column and several comparison data columns. This method can analyze the correlation degree between many factors during the change process, so it is widely used in dynamic system analysis. The calculation steps are as follows:

1. Determine the reference data sequence and the comparison data sequence. Here the reference data sequence is \( X_0 \), and the comparison sequence is \( X_i \).
2. Dimensionless processing of data. Before performing correlation analysis, the data needs to be dimensionlessly processed. There are two dimensionless processing methods: initial value processing and average processing. This article uses the average processing method, which uses the following formula to calculate:

\[
Y_i(k) = \frac{X_i(k)}{X_i}, \quad k = 1,2, \ldots, n; \ i = 0,1, \ldots, m
\]  

(1)

3. Calculate correlation coefficient. Use the following formula to calculate:

\[
\xi_i(k) = \frac{\min \min |Y_i(k) - Y_j(k)| + \rho \max \max |Y_i(k) - Y_j(k)|}{|Y_i(k) - Y_j(k)| + \rho \max \max |Y_i(k) - Y_j(k)|}
\]

(2)

Among them, \( \rho \) is called a resolution coefficient. The smaller the \( \rho \), the greater the resolution. This paper takes \( \rho \) as 0.5 to calculate correlation coefficient.
(4) Calculate correlation degree. The correlation coefficients at each moment are integrated to form a correlation degree which shows the overall comparison of the correlation. Use the following formula to calculate:

\[ r_i = \frac{1}{N} \sum_{k=1}^{N} \xi_i^r(k) \]  

(3)

Among them, \( r \in (0,1] \). The greater the value of \( r \), the greater the correlation between the influencing factors and the system; otherwise, the smaller the correlation between the two.

(5) Correlation degree ranking. Sort the influencing factors according to the correlation degree, which explains the correlation between factors and system.

2.2. Correlation Analysis of Factors Affecting Main Steam Temperature

This paper uses MATLAB software to calculate correlation degree of various factors affecting main steam temperature according to the above steps, the results are shown in Table 1.

| #  | Factor                                         | Correlation degree |
|----|-----------------------------------------------|--------------------|
| 1  | First desuperheating water flow               | 0.8929             |
| 2  | Fuel quantity                                 | 0.8409             |
| 3  | Steam temperature after first desuperheating  | 0.8198             |
| 4  | Main steam flow                               | 0.8031             |
| 5  | Boiler main control variable                  | 0.7401             |
| 6  | Water temperature in the economizer inlet     | 0.7131             |
| 7  | Air volume                                    | 0.7084             |
| 8  | Main steam pressure                           | 0.7051             |
| 9  | Steam turbine valve position                  | 0.6992             |
| 10 | Generator active power                        | 0.6972             |
| 11 | Boiler main control set value                 | 0.6591             |

As can be seen from Table 1, the specified threshold is set to 0.8, and the correlation is considered to be good when the correlation degree exceeds 0.8. When the correlation degree is between 0.6 and 0.8, the correlation is not bad. The results show that the four process variables of first desuperheating water flow, fuel quantity, steam temperature after first desuperheating and main steam flow are the main factors affecting main steam temperature. Therefore, these four variables are used to predict the main steam temperature.

3. Model

LSTM effectively overcomes the difficulties of “gradient disappearance” and “gradient explosion” in RNN, and has been widely used in tasks related to time series learning, such as speech recognition, language model, machine translation etc. However, the linear parts of the network are approximated by non-linear functions. Non-linear functions of multiple hidden layers and multiple nodes need to be built and superimposed, which requires high computing power of the equipment used, prolongs the calculation time and approximates the accuracy reduce. Therefore, based on the study of LSTM, this paper constructs an improved LSTM network which combines linear and nonlinear. Figure 1 is a comparison diagram of the network structure of LSTM and improved LSTM. By comparing figures 1a and 1b, it can be seen that the traditional LSTM directly obtains the output value through the full connection calculation of the last hidden layer nodes, while the improved LSTM adds a direct connection matrix A of input vector and network output, so that the linear features in the input vectors are directly mapped to the output.
Figure 1. Comparison of network structure between LSTM and improved LSTM.

The main steam temperature prediction model based on the improved LSTM includes not only the non-linear composition of the traditional LSTM model but also the linear composition. The linear part of the network no longer needs to be approximated by non-linear functions, which can effectively improve the overall convergence speed and approximation ability, and thus the network Adaptability. Its output layer is expressed as:

$$o_t = \sum_{j=1}^{k} V_j h_{t,j} + \sum_{j=1}^{l} A_j x_{t,j} + b$$

Among them, $l$ represents the number of dimensions of the input vector, and $k$ represents the number of nodes in the last hidden layer.

The training process of the network is to determine the connection strength of each input, namely the weights, scaling parameters and translation parameters between nodes. This model uses Adam optimization algorithm to modify parameters. Adam optimization algorithm is an optimization algorithm of stochastic gradient descent algorithm, which was proposed by Kingma and Ba in 2015 [7]. Adam algorithm can be expressed as:

$$
\begin{align*}
    m_t &= \beta_1 m_{t-1} + (1 - \beta_1) g_t \\
    v_t &= \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \\
    m'_t &= \frac{m_t}{1 - \beta_1^t}, \quad v'_t = \frac{v_t}{1 - \beta_2^t} \\
    W_{t+1} &= W_t - \frac{\alpha}{\sqrt{v'_t} + \epsilon} m'_t
\end{align*}
$$

Among them, $m_t$ and $v_t$ represent the first- and second-order momentum terms; $m'_t$ and $v'_t$ represent the correction values of the first- and second-order momentum terms; $W_{t}$ represents the parameters of the model at t-th time; $g_t = \Delta J (W_t)$ represents t times the gradient size of the iterative cost function with respect to $W$; $\alpha$ represents the initial global learning rate; $\beta_1$ and $\beta_2$ represent dynamic values; $\epsilon$ is a parameter with a small value, in order to avoid the denominator being 0.

Parameter adjustment of Adam algorithm are relatively simple and the default parameters can handle most of the problems, so the model in this paper uses the default settings of TensorFlow to set parameters, namely $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 1e-08$.

4. Experiment

The main steam temperature is affected by multiple factors such as the fuel quantity and the desuperheating water flow, and has the characteristics of large inertia, large delay, etc. Accurately predicting the trend can effectively help control the thermal power unit coordinately. This paper designs a main steam temperature prediction model based on improved LSTM, and validates this model using historical operating data of 800 MW generating units in a domestic plant collected from
2019-1-1 to 2019-2-18 (seven consecutive weeks) collected through Distributed Control System (DCS).

4.1. Data Preprocessing
The historical data collected by DCS is 1 group per 1 second, and the data is sampled at a time interval of 5 seconds to obtain 850,000 groups of data. The main influencing factors were selected as the independent variables of the prediction model, and the main steam temperature is set as the dependent variable.

All data are normalized, that is, each feature component is normalized to a specified range to ensure that the input attribute of a larger value does not affect the input attribute of a smaller value. Finally, the normalized data is divided into 8: 1: 1 and the training set, validation set and test set are obtained for experiments.

4.2. Evaluation Indicators
The experiment uses Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) as evaluation indicators to measure whether the predicted value is close to the actual value. The formula is as follows:

\[
\epsilon_{\text{MAPE}} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% 
\]

(6)

\[
\epsilon_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} 
\]

(7)

Among them, n is the number of predicted samples; \(y_i\) represents actual main steam temperature at the i-th time; \(\hat{y}_i\) represents predicted main steam temperature at the i-th time. When \(\epsilon_{\text{MAPE}}\) and \(\epsilon_{\text{RMSE}}\) are small, it indicates that the network prediction value is close to the actual value and has high prediction accuracy.

4.3. Experimental Results
The main steam temperature prediction model based on the improved LSTM is compared with the model based on RNN and LSTM algorithms. RNN, LSTM and the improved LSTM all use three hidden layers, and the number of nodes in each layer is 32. The number of training iterations is 1000. The activation function of RNN model is a tanh function. The step length of each model is 720 seconds, and each time the temperature is predicted in the next 360 seconds.

The historical data of the operating unit was used as the data source, and three prediction models were applied to predict the temperature. Table 2 lists the average \(\epsilon_{\text{MAPE}}\) and \(\epsilon_{\text{RMSE}}\) values for each model.

| Model       | \(\epsilon_{\text{MAPE}}\) | \(\epsilon_{\text{RMSE}}\) |
|-------------|-----------------------------|-----------------------------|
| RNN         | 11.280%                     | 3.053                       |
| LSTM        | 1.734%                      | 1.752                       |
| Improved LSTM | 1.226%                    | 1.163                       |

It can be seen from table 2 that the average \(\epsilon_{\text{MAPE}}\) of the improved LSTM model is 1.226%, and the result is significantly better than the RNN model and slightly better than the LSTM model. The average \(\epsilon_{\text{RMSE}}\) of the improved LSTM model is 1.163, which is 1.301 and 0.589 lower than the RNN model and LSTM model, respectively. The improved LSTM model has higher approximation accuracy and further improves the prediction effect.
To further analyze the model performance, table 3 shows the experimental results of five test samples. It can be seen that the maximum $e_{\text{MAPE}}$ of the improved LSTM model is 1.457%, the minimum is 1.041%, the maximum $e_{\text{RMSE}}$ is 1.380, and the minimum is 0.863. All the sample prediction errors of the improved LSTM model are smaller than errors of the RNN model and LSTM model, and the error fluctuation is small, which means that it has better robustness and stability.

Table 3. Comparison of prediction results of test samples.

| Model         | 1   | 2    | 3    | 4    | 5    |
|---------------|-----|------|------|------|------|
| RNN           | 8.475% | 10.867% | 14.691% | 9.901% | 12.496% |
| LSTM          | 1.829% | 1.906% | 1.738% | 1.417% | 1.779% |
| Improved LSTM | 1.041% | 1.457% | 1.283% | 1.051% | 1.300% |
| $e_{\text{MAPE}}$ | 1   | 2    | 3    | 4    | 5    |
| RNN           | 4.115 | 3.325 | 2.791 | 2.324 | 2.709 |
| LSTM          | 1.605 | 2.259 | 1.846 | 1.394 | 1.880 |
| Improved LSTM | 1.380 | 1.312 | 0.929 | 1.108 | 0.863 |

Figure 2 shows the main steam temperature prediction results of the three models for the same half-hour. It can be intuitively found that the RNN model has the poor prediction effect and can’t grasp the basic trend; the prediction results of LSTM model and the improved LSTM model are closer to the trend of the actual main steam temperature curve; compared with LSTM model, the improved LSTM model fits actual curve better, and the deviation in the peak and trough areas of the curve is smaller. Compared with the traditional RNN model and the LSTM model, the curve fitting effect of the improved LSTM model is more accurate, and law capturing ability is more advantageous.

5. Conclusion
Aiming at the problems of various factors affecting main steam temperature of thermal power units and their characteristics such as large inertia and large delay, a main steam temperature prediction model based on improved LSTM was proposed, which is verified by practical examples. The main conclusions of this paper are:

(1) The grey correlation analysis is used to determine the main influencing factors of the main steam temperature, which are used to predict;

(2) Using the time series characteristics, a main steam temperature prediction model based on the improved LSTM was established. This model adds a linear fitting part to the LSTM network, so that the improved network can have faster convergence speed and approximation ability in addition to long-term sequence processing ability;
(3) Based on the historical data of the operating unit, the RNN model, the LSTM model and the method of this paper are used to predict the same set of data. The results show that the improved LSTM model is superior to the traditional RNN model and LSTM model in terms of prediction accuracy and fitting effect. Therefore, the improved LSTM model has great significance in main steam temperature prediction.

References
[1] Wu C, Wang C and Li B 2018 Application of fuzzy self-tuning PID cascade control in main steam temperature control *International Electronic Elements* **26** 122-5.
[2] Li Q, Zhang H, Peng D, Guo Y, Wang N and Sun Y 2017 Multi-variable modeling research for main-steam temperature of power station boiler based on improved difference evolution algorithm *Journal of System Simulation* **29** 1712-18.
[3] Wang Q, Ma C, Xiao L and Cui R 2013 Application of BP neural network optimized by ant colony algorithm to main steam temperature control *Control and Instruments in Chemical Industry* **40** 834-7.
[4] Fang Y, Yi F and Hu W 2013 Genetic algorithm-based generalized predictive PID control and its application to main steam temperature system *Engineering Journal of Wuhan University* **46** 386-92.
[5] Pao Y H, Park G H and Sobajic D J 1994 Learning and generalization characteristics of the random vector functional-link net *Neuro Computing* **6** 163-80.
[6] Deng J 1982 Grey control system *Journal of Huazhong University of Science and Technology (Natural Science Edition)* **10** 9-18.
[7] Kingma D and Ba J 2014 Adam: A method for stochastic optimization *Computer Science* **1**.