

Received June 15, 2020, accepted July 2, 2020, date of publication July 16, 2020, date of current version September 9, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3009679

# Time Series Risk Prediction Based on LSTM and a Variant DTW Algorithm: Application of Bed Inventory Overturn Prevention in a Pant-Leg CFB Boiler

FENG HONG<sup>1</sup>, JIYU CHEN<sup>1</sup>, ZHIYU ZHANG<sup>2</sup>, RUI WANG<sup>2</sup>, AND MINGMING GAO<sup>2</sup>

<sup>1</sup>State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China

<sup>2</sup>School of International, Beijing University of Posts and Telecommunications, Beijing 100876, China

Corresponding author: Feng Hong (hongf@ncepu.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2016YFB0600205; and in part by the Fundamental Research Funds for the Central Universities under Grant 2020MS013.

**ABSTRACT** The pant-leg design is typical for higher capacity circulating fluidized bed (CFB) boilers because it allows for better secondary air penetration, maintaining good air-coal mixing and efficient combustion. However, the special risk, nominated as bed inventory overturn, remains a big challenge and it hinders the application of pant-leg CFB boilers. For a time series risk, it is critical to do the bed inventory overturn prevention to leave enough time for the adjustment. This paper proposed a new framework combing long short-term memory (LSTM) and dynamic time warping (DTW) methods to do the risk prediction. Pattern matching of data difference discrimination is employed for DTW algorithm, instead of the traditional Euclidean metric. The pattern matching has the merits in reduction of calculation and improvement of the adaptability to variables with different dimensions. After variable processing of the time series data by the variant DTW algorithm, the bed pressure drop prediction model is established based on the LSTM structure in this framework. Compared with some traditional prediction method, the framework in this paper has achieved superior results in the application of bed inventory overturn prevention.

**INDEX TERMS** Risk prediction, time series, LSTM, CFB, bed inventory overturn.

## I. INTRODUCTION

The circulating fluidized bed (CFB) boiler is believed to be an improvement over the conventional pulverized coal furnace in some respects [1]. Operation of industrial CFB boilers has confirmed some advantages like fuel flexibility, low NO<sub>x</sub> emissions, high sulphur capture efficiency and so on [2]. As for the CFB furnace, the heat import is proportional to bed cross-sectional area while the heat absorption is proportional to the perimeter of the furnace [3]. With the growing required capacity, both the furnace volume and heat transfer surface increase accompanied by the contradiction that the former increases faster, naturally in need of higher furnace height to control the temperature [4]. However, the furnace height of a 300MW CFB boiler is mostly limited to about 50m due to commercial consideration.

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Liu.

To have extra heat transfer surface, the external heat exchanger (EHE) could be a good alternative option to a scaled-up CFB boiler [5], and the corresponding structure of the furnace is called the pant-leg [6]. A pant-leg bottom CFB boiler with EHE has several advantages [7]:

(1) The reheat steam temperature can be adjusted through regulating the control valves of EHE rather than the spray water, improving the efficiency of power unit.

(2) Bed temperature can be controlled flexibly and reliably.

(3) EHE increases the heat storage of the unit to a certain extent, which could enhance the fuel flexibility.

The industrial operations prove that the pant-leg structure can improve the mixing of air and solid in the furnace and reduce the carbon content in fly ash greatly. However, the two independent distributors at the bottom of the pant-leg boiler will always cause bed inventory imbalance. When the imbalance aggravates, one special phenomenon, nominated as bed inventory overturn, will occur and cause the shutdown of

boiler without timely adjustment [8]. Because of the considerable negative influence on CFB boiler operation security, more and more attention is paid to the risk prediction and prevention of bed inventory overturn. To understand the mechanisms underlying the bed inventory overturn, Wang [9] looked into the bed material imbalance on both sides related to the bed inventory overturn. Liu *et al.* [10] proposed the reasons for bed inventory overturn based on mechanism operation analysis, and the imbalances of bed pressure are specially analyzed. Basu *et al.* [11] had showed the details about the pressure balance behavior affected by the loop-seal operation. The effects of particle diameter, particle density and gas distributor design on the hydrodynamics of CFB were studied by Qi *et al.* [12]. Li experimentally studied the lateral transfer of solid particles in a small-scale, cold CFB riser with pant-leg structure [13]. A compounded mathematical model of pressure drop was established, and it was concluded that the main reason for lateral migration of solid particles is the lateral pressure gradient of the gas phase in the CFB and once the pressure balance is broken, it is difficult to restore the balance without timely adjustment of the primary air fan. Therefore, monitoring the risk trend of bed inventory overturn is vital, which can be considered as a typical accident with time series characteristics. A prediction method of bed pressure fluctuation in CFB riser was proposed by Zhao *et al.* [14], and this method reconstructed the phase space trajectory evolution by establishing discrete dynamic mapping equation. Afsin Gungor [15] established a CFB axial pressure distribution prediction model based on particle method (PBA).

Some of mechanism model methods above have been well applied in laboratory situations, providing clear analyzation of both reasons and modes for the bed inventory overturn. However, due to the high complexity and low adaptability of the mechanism modeling, the existed methods are difficult to be applied in industrial situation to do the bed inventory overturn prevention. Therefore, mechanic-based prediction model, rather than a simple risk predictor, is employed for condition design application in most investigation. With the developments of machine learning and data management system equipped in modern industrial process, more and more data driven methods are applied in performance monitoring and prediction.

Using the least squares support vector machine (LSSVM) to construct the model, a dynamic model was developed to predict the bed temperature of a CFB boiler [16]. Li proposed a method based on wavelet decomposition (WD) and a second-order gray neural network combined with an augmented Dickey-Fuller (ADF) test is proposed to improve the accuracy of load forecasting [17]. A novel hybrid ensemble deep learning (HEDL) approach was presented for deterministic and probabilistic low-voltage load forecasting. And the deep belief network (DBN) is applied to low-voltage load point prediction with the strong ability of approximating nonlinear mapping [18]. A forecasting study of hydro-electricity consumption in Pakistan was presented based

on Auto-Regressive Integrated Moving-Average (ARIMA) modeling, and the research was useful in better planning and management for future [18]. Que proposed a data-driven integrated framework for health prognostics for steam turbines, which is based on extreme gradient boosting (XGBoost) and dynamic time warping (DTW). And the proposed framework has achieved good results in practical application [19].

Since deep learning specialized in abstracting complex relationship among variables with multiple layers, the risk prediction model, with time series characteristics, could be established by analysing temporal relevance of operating industrial data. Recurrent neural network (RNN) is one of the effective algorithms that can accommodate relevance between consecutive time steps [20]. Long short-term memory (LSTM) units are suggested as a possible solution to the vanishing gradient problem noticed in the simple RNN [21]. The sequence to sequence approach based on LSTM has been previously employed in speech recognition, speech emotion classification machine translation applications [22], short-term weather forecasting [23], medium-to-long term electricity consumption for commercial and residential buildings [24]. For the bed temperature monitoring in a 300 MW CFB unit, Li [25] presented a 2D-interval prediction model based on LSTM. The results revealed that the model structure could effectively described the characteristic of bed temperature of CFB unit.

Based on the characteristics of bed inventory overturn, this paper, combining LSTM and DTW algorithms, proposed a new monitoring framework to predict the risk and prevent the overturn. DTW would be employed to extract the temporal dynamic characteristics, while LSTM has unique superiority in time series analysis, thus leading to an effective framework to prevent the overturn.

The rest of paper is organized as follow, Chapter 2 describes of investigated object, and Chapter 3 introduces the method used in risk prediction. The prediction framework will be discussed in Chapter 4. After that, the verification results are shown in Chapter 5, and Chapter 6 draws the final conclusion.

## II. OBJECT DESCRIPTION

### A. THE GENERAL LAYOUT OF THE INVESTIGATED BOILER

This paper mainly investigates a 300MW coal-fired CFB boiler, which belongs to the 1# unit of JoinLion power plant in China. It is a subcritical reheat boiler that respectively represent the typical pant-leg CFB furnace currently.

The material balance in the main loop of the pant-leg CFB furnace can be shown in Fig. 1. The coal and limestone are recycled many times to increase the fuel combustion efficiency as well as improve the utilization rate of limestone. To separate the heavier particles from the flue gas and return to the furnace for recirculation, two cyclones are arranged at each side of the furnace that with a pant-leg type in lower part of boiler. Circulating solids captured by cyclones enter loop seal and external heat exchanger (EHE) installed at the end of each cyclone standpipe respectively. A cone valve is

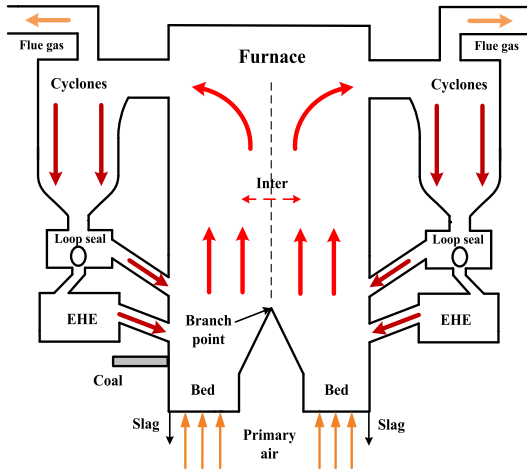


FIGURE 1. The material balance in the main loop of a pant-leg CFB furnace.

set at the inlet of each EHE, then the portion of low and high temperature solids return to the furnace are controlled by adjusting cone valves opening [28].

Once the bed inventory occurs, the operator has to adjust the air valve to improve the air flow rate in the leg with defluidization but turn down the flow rate in the leg with little bed inventory. If adjustments are effective and timely, the imbalance would be reduced and the bed material would restore to the origin. Then the risk of bed inventory overturn could be avoided. Due to the hysteresis and complicated dynamic of the bed inventory process, it is quite difficult to do the risk prediction. The untimely or inaccurate adjustments usually cause the shuttle of bed inventory between the two legs. The risk of bed inventory overturn is a typical accident with time series characteristics. It can make a great contribution to the effective and timely adjustment strategy with the precise risk prediction of the bed inventory imbalance between the two legs in a pant-leg CFB boiler.

**B. THE BED INVENTORY OVERTURN IN CFB**

The pant-leg structure can strengthen the penetration of secondary air in large CFB boilers to improve the fluidization and combustion efficiency. However, the unique structure of two independent distributors at the bottom will always cause bed inventory imbalance between the two legs. If the imbalance aggravates further, the bed inventory overturn would happen. As shown in Fig. 2, the bed pressure drop in the right leg increases while the air flow rate decreases, until the bed materials in the left leg is blown out and transferred into the right leg. The material quantity in the dense phase area of the legs on both sides is different due to the deviation between two air distribution plates. As a result, the bed pressure drop in the right leg gets further increase until the bed inventory is too vast to be fluidized by the primary air.

**III. METHODS OF THE ALGORITHM**

**A. LONG SHORT-TERM MEMORY**

Based on the recursive structure, RNN algorithm has the ability to memorize results of different layers, thus solving the

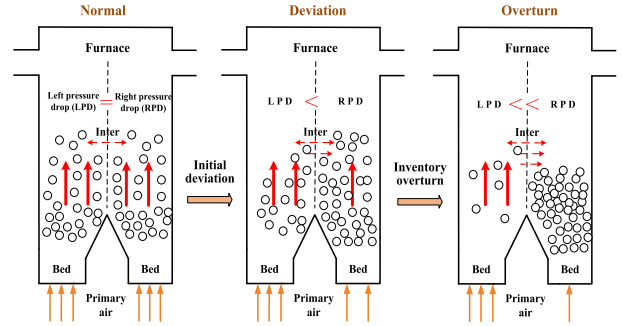


FIGURE 2. The process of inventory overturn; The inventory overturn accident comes from the positive feedback of transverse material flow in the furnace. This kind of positive feedback makes materials accumulate on one side of the furnace.

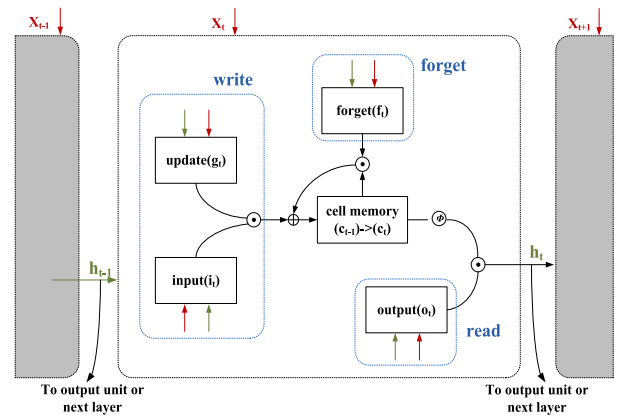


FIGURE 3. Layers in a LSTM cell.

time series problems successfully and remarkably. However, there is a demerit for RNN, known as long-term dependence, which means if dealing with the key points in message needs to employ former information with long distance, RNN tends to have errors and even break down. The LSTM algorithm has avoided this disadvantage via improving the structure, and it is displayed as Fig. 3.

Compared with the single layer in RNN structure, the process for LSTM is more complicated. As shown in Fig. 3, there are four neural networks interacting with each other in a particular way within one single LSTM cell. Represented as the horizontal line through the top of the diagram, the state of the cell in LSTM is transferred as a conveyor belt, running through the whole chain structure with only a few minor linear interactions. The output of the specific layer is to convey the state of the cell into  $C_t$  from  $C_{t-1}$ .

Three control gates are employed to command the state of the cell, including forget gate, input gate and output gate. In the forget gate processing, input data  $x_t$  and state data  $h_{t-1}$  would be concentrated and calculated with the equation as follows.

$$f_t = \sigma(W_f \circ x_t + U_f \circ h_{t-1} + b_f) \tag{1}$$

where  $\sigma$  is the sigmoid function, and  $\circ$  is the Hardamand product operation.

The result  $f_t$  would be put into the state  $C_{t-1}$ . If  $f_t$  is 0, the forget gate would delete all the information. On the contrary, the information would be saved thoroughly. And the expression of the sigmoid layer  $\sigma$  is

$$S(x) = \frac{1}{1 + e^x} \quad (2)$$

To add new information for the whole process, input gate needs deciding whether to save or delete information with the equation.

$$i_t = \sigma(W_i \circ [h_{t-1}, x_t] + b_i) \quad (3)$$

In the meanwhile, the updating factor of cell status, the vector  $\tilde{C}_t$ , would be generated with the layer operation  $\tanh$ .

$$\tilde{C}_t = \tanh(W_C \circ [h_{t-1}, x_t] + b_C) \quad (4)$$

The forget gate and the input gate provide essential elements to convert the cell status into  $\tilde{C}_t$ .

$$C_t = f_t \circ C_{t-1} + i_t \circ \tilde{C}_t \quad (5)$$

The output information would be assigned by the output gate, which would be based on the state of the cell as well as going through some filtration. Once the sigmoid layer has determined the output content of the cell, the  $\tanh$  layer would be employed to transfer the state of the cell between -1 and 1, following with calculating the output of the sigmoid layer.

$$o_t = \sigma(W_o \circ [h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = o_t \circ \tanh(C_t) \quad (7)$$

## B. DYNAMIC TIME WARPING

DTW is a similarity measurement method which is able to match and map the time series morphology by bending the time axis. It can measure the time series data with the same length, and has the ability to measure the similarity among time series with various length. The best merit is represented in its insensitivity to the abnormal and abrupt point in time series, for which the asynchronous similarity comparison is implemented well.

Take two time series  $Q$  and  $U$  for analysis, where  $Q = \{q_1, q_2, \dots, q_n\}$ , and  $U = \{u_1, u_2, \dots, u_n\}$ . By calculating the Euclidean distance among points within the two sequences, the distance matrix  $D_{n \times m}$  for these data points can be implemented as (8), and  $d(i, j)$  is the Euclidean distance between  $q_i$  and  $u_j$ . The matrix  $D_{n \times m}$  represents distances among data of various time points in the two sequences.

$$D_{n \times m} = \{d(i, j)\}_{n \times m}, \quad 1 \leq i \leq n, 1 \leq j \leq m \quad (8)$$

The specific similarity between two sequences can be calculated by DTW, and DTW algorithm finds the shortest distance path in the distance matrix. The similarity between two sequences is characterized by the sum of distances on the path. DTW is dedicated to find a continuous path  $H = \{h_1, h_2, \dots, h_s\}$  such that the sum of all elements in the path is minimal while three necessary requirements are satisfied, including boundary limits, continuity and monotonicity.

Optimal path ( $H$ ) is searched by implementing dynamic programming, constructing the cost matrix  $D_c$  (also known as the cumulative matrix) based on the matrix  $D_{n \times m}$  to record the shortest path from the beginning point to the end point, which would be obtained mainly relying on the following steps:

- 1) The element in the first row, the first column is the element in the first row, the first column of  $D_{n \times m}$ .
- 2) The values of elements sitting in other locations ( $D_c(i, j)$ ) would be calculated step by step, by the formula:

$$D_c(i, j) = \min(D_c(i-1, j-1), D_c(i-1, j), D_c(i, j-1)) + D_{n \times m}(i, j) \quad (9)$$

The element in the final row, the final column of  $D_c$  illustrates the distance between the two sequences  $Q$  and  $U$ , shown the output of DTW meanwhile.

## IV. DEVELOPMENT OF THE RISK PREDICTION FRAMEWORK

This paper proposed a new framework combing LSTM and DTW methods to do the risk prediction. Firstly, with analyzing the data correlation and filtering out irrelevant variables, the time series relationship among data has been adjusted by DTW method. In the processing part, the filtered data is normalized and non-linear conversion is performed by the Sigmoid function. Secondly, the data is employed to train the prediction model and obtain the prediction value, which could carry out early warnings. Instead of conventional Euclidean metric, DTW employs the data-difference-discrimination pattern matching, which could reduce the calculation and improve the adaptability of variables in different dimensions. Once the time series data are preprocessed by variant DTW algorithm, the bed pressure drop prediction model would be established based on LSTM algorithm. The whole process is illustrated in the following figure.

### A. VARIABLE PROCESSING OF THE TIME SERIES DATA BY THE VARIANT DTW ALGORITHM

DTW is used to analyze the correlation of variables and obtain the analysis results, according to which unnecessary variables would be filtered out and the variable data structure would be adjusted. DTW model is computationally intensive in practical application, especially for a large number of time series data. And Euclidean metric is used as distance formula in the traditional DTW model, which is disadvantageous for the variables with different dimensions.

In this paper, a pattern matching method is used in DTW distance formula. The continuous time series is converted into discrete representation features, which not only simplifies the computation, but also enhances the compatibility of data in different dimensions. Continuous data are divided into three categories, rising, maintaining and falling. The concept of pattern matching is used to replace the Euclidean metric as a representation of the differences between data. In the following (10), 1, 0.5, 0 represent rise, hold and

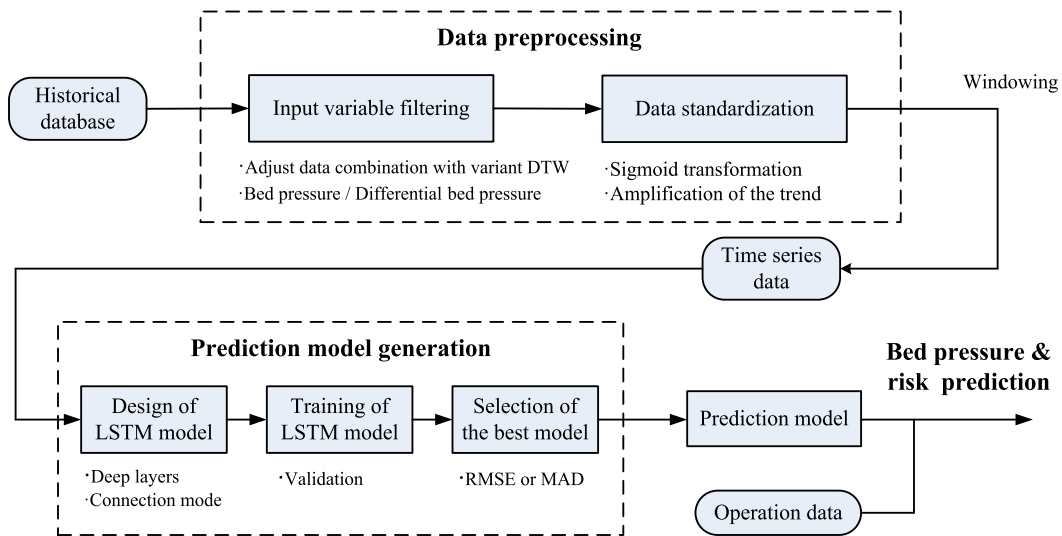


FIGURE 4. Flowchart of the proposed time series risk prediction framework.

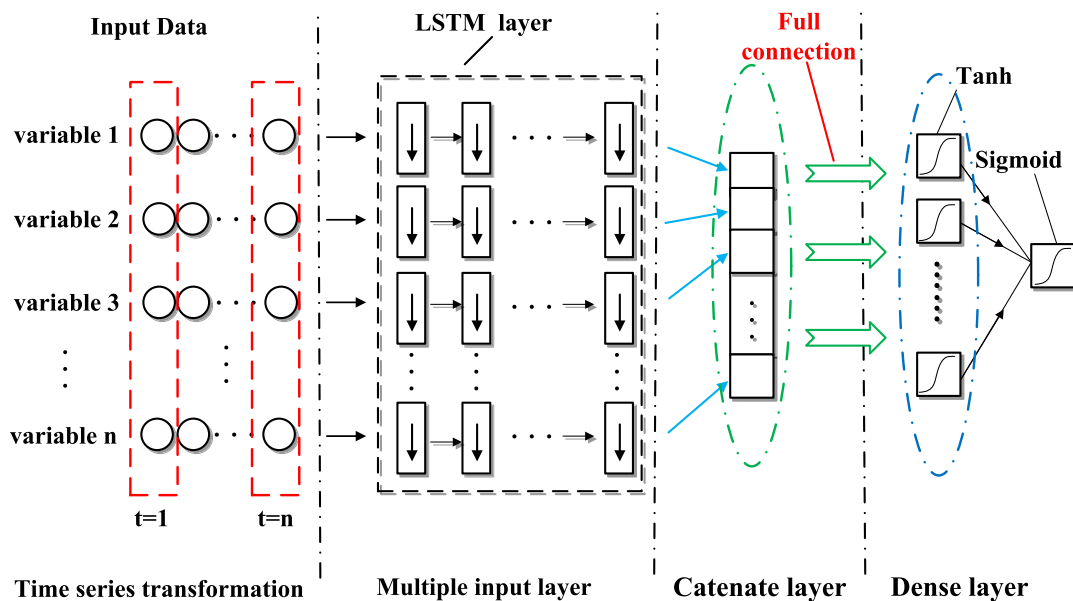


FIGURE 5. The structure of LSTM model for bed pressure drop prediction.

fall models respectively

$$d(i, j) = \begin{cases} 0, & o_i = o_j \\ 0.5, & o_i = -o_j \\ 1, & else \end{cases} \quad (10)$$

In the practical industrial process, as a consequence of the large inertia characteristics of objects, the influence among data, like the coal supply and the unit load, tends to be out of synchronization. When the coal supply changes, the unit load often lags for some time before it changes. Therefore, in data pre-processing, we should not only analyse the correlation between the data but also consider the inertia between the data. The optimization path of DTW indicates the direction of time series compression, which enables data to be inertia processed and be lagged to a certain extent.  $y = x + b$  is used to fit the optimization path. The number of lag samples

is obtained according to the lookback of the model and the absolute value of  $b$ .

### B. DATA STANDARDIZATION AND SELECTION OF MODEL LOSS FUNCTION

Data features are extracted from various dimensions to avoid inaccuracy of model precision and confusion of optimization trajectory when training model. The prediction target is bed pressure, so the bed pressure is normalized by min-max shown as (11).

$$y_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (11)$$

where  $x_{\min}$  and  $x_{\max}$  represent the minimum and maximum values in the data,  $y_i$  represents the data after min-max standardization, and  $x_i$  represents the data before processing. Linear change of data between 0 and 1 also benefits for training and initialization of network model weights and offsets.

TABLE 1. Correlation coefficient of left bed pressure.

Variable	DTW correlation coefficient
Load instruction	533.5
Total Fuel Quantity Directive	598.5
The corrected total fuel quantity	221
4# Weighing Coal Feeder Instantaneous Flow	440.5
3# Weighing Coal Feeder Instantaneous Flow	342
2# Weighing Coal Feeder Instantaneous Flow	307
1# Weighing Coal Feeder Instantaneous Flow	526.5
Speed of slag cooler A	203.5
Speed of slag cooler B	483
Speed of slag cooler C	438
Speed of slag cooler D	457.5
Speed Setting of 2# Chain Scraper Coal Feeder	367.5
Speed feedback of 4# chain scraper coal feeder	251
Valve Position of Conical Electric Regulating Valve of Left 1# Return Valve	672
Valve Position of Conical Electric Regulating Valve for Right Side 1# Return Valve	366.5
Lower Left secondary air	382.5
Lower right secondary air	371.5
Upper left secondary air	724
Upper right secondary air	429.5
Left primary wind	290
Right primary wind	159.5

In the actual industrial process, there are usually some wrong data points or mutation points. Min-Max normalization method cannot solve the influence of these error points or mutation points on the overall data. On the contrary, this kind of normalization is so sensitive to the abnormal data that the overall trend of the data would be not clear under the influence of a single bad value. However, linear change cannot solve the problem distinguishably. In this paper, Sigmoid function is used to convert data nonlinearly. Since the domain of Sigmoid function is  $(-\infty, +\infty)$  and the corresponding domain is  $(0, 1)$ , the information-loss problem would be avoided when converting data with Sigmoid function. If the computer precision allows, when a sudden change point occurs, its conversion value will only be infinitely close to 1 or 0, and the value of the sudden change point can be restored by inverse transformation. Sigmoid function is shown as follows:

$$S(x) = \frac{1}{1 + e^{-x}} \tag{12}$$

For the normalized data, we transform Sigmoid function to adapt to our data. The formula is as follows:

$$y = S\left\{\frac{10}{b-a}x - \left(\frac{10a}{b-a} + 5\right)\right\} \tag{13}$$

where y represents the converted data and x represents the data before conversion. a and b respectively represent the extent of expansion of data trends. The values of a and b are determined according to the results of data normalization.

**C. BED PRESSURE DROP PREDICTION MODEL BASED ON THE LSTM**

The conventional LSTM network, including LSTM layers and full connection layers, performs inadequately in data prediction with high complexity, while high-depth layers requires longer training time and more advanced machine configuration. Thus, the structure combining with deepening width, full connection layer and LSTM layer is applied shown as Fig. 5.

**V. RESULTS AND DISCUSSIONS OF A CASE STUDY**

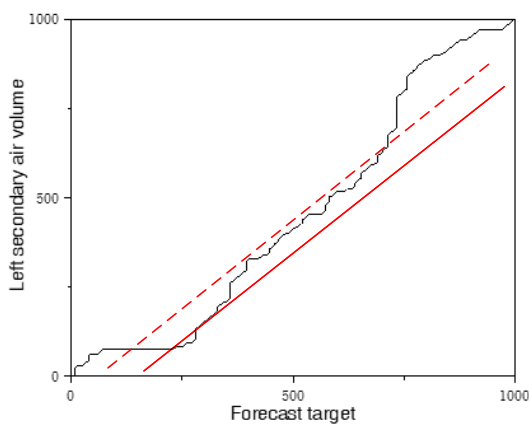
In this paper, a typical 330MW unit is considered in the case study. The data come from JoinLion 1 # 330MW circulating fluidized bed unit in China. In order to facilitate the training of LSTM, Sigmoid function is used to adjust Min-max normalization. The Sigmoid function enlarges data change trend and reduces the influence of deviation point on overall data. In addition, DTW path is used to process the time-data symmetrically to reduce the influence of large inertia links in mechanism objects on model prediction.

**A. CORRELATION ANALYSIS BASED ON DTW**

Correlation analysis is carried out on data variables, and input variables of the model are selected according to the correlation with target variables. The target variable of the model is bed pressure difference, so correlation analysis is carried out on bed pressure and bed pressure difference simultaneously. As long as the variables satisfy one of them, they can be

**TABLE 2. Correlation coefficient of left bed pressure difference.**

Variable	DTW correlation coefficient
Load instruction	360
Total Fuel Quantity Directive	307.5
The corrected total fuel quantity	183
4# Weighing Coal Feeder Instantaneous Flow	253.5
3# Weighing Coal Feeder Instantaneous Flow	148
2# Weighing Coal Feeder Instantaneous Flow	223
1# Weighing Coal Feeder Instantaneous Flow	355
Speed of slag cooler A	159
Speed of slag cooler B	236.5
Speed of slag cooler C	474.5
Speed of slag cooler D	175.5
Speed Setting of 2# Chain Scraper Coal Feeder	534.5
Speed feedback of 4# chain scraper coal feeder	185
Valve Position of Conical Electric Regulating Valve of Left 1# Return Valve	548
Valve Position of Conical Electric Regulating Valve for Right Side 1# Return Valve	215
Lower Left secondary air	152.5
Lower right secondary air	134.5
Upper left secondary air	72.5
Upper right secondary air	124.5
Left primary wind	103.5
Right primary wind	113



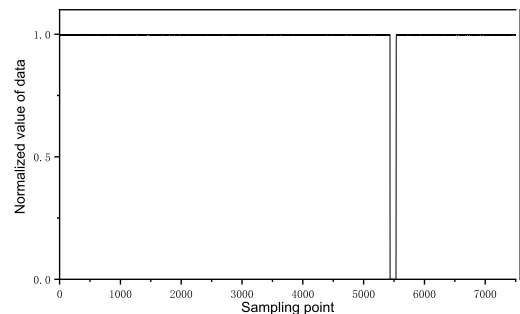
**FIGURE 6. The time matching of the data based on the DTW path.**

used as input variables of the model. The correlation coefficients are shown in Tab.1 and Tab.2. According to experience, the DTW coefficient threshold is set to 260.

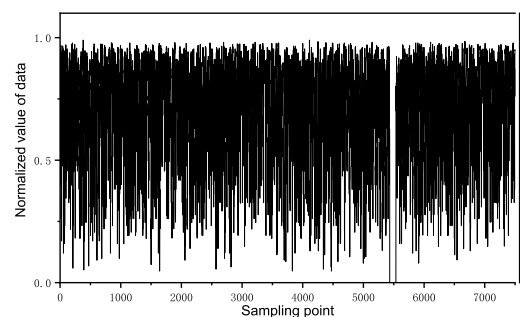
According to the results, we selected 15 related variables such as corrected total fuel quantity, instantaneous flow of 4# weighing coal feeder, instantaneous flow of 3# weighing coal feeder and instantaneous flow of 2# weighing coal feeder.

**B. INERTIA PROCESSING WITH DTW ALGORITHM**

DTW algorithm finds the shortest path in the metric space with the dynamic time adjustment, which also represents the correlation between two groups of data. Besides, the direction



**FIGURE 7. The result of the left primary wind after normalization.**



**FIGURE 8. Data normalization after Sigmoid processing.**

of associated data are presented in the path, describing the relationship among data. In the industrial process, huge inertia exists among variables, which means data out of

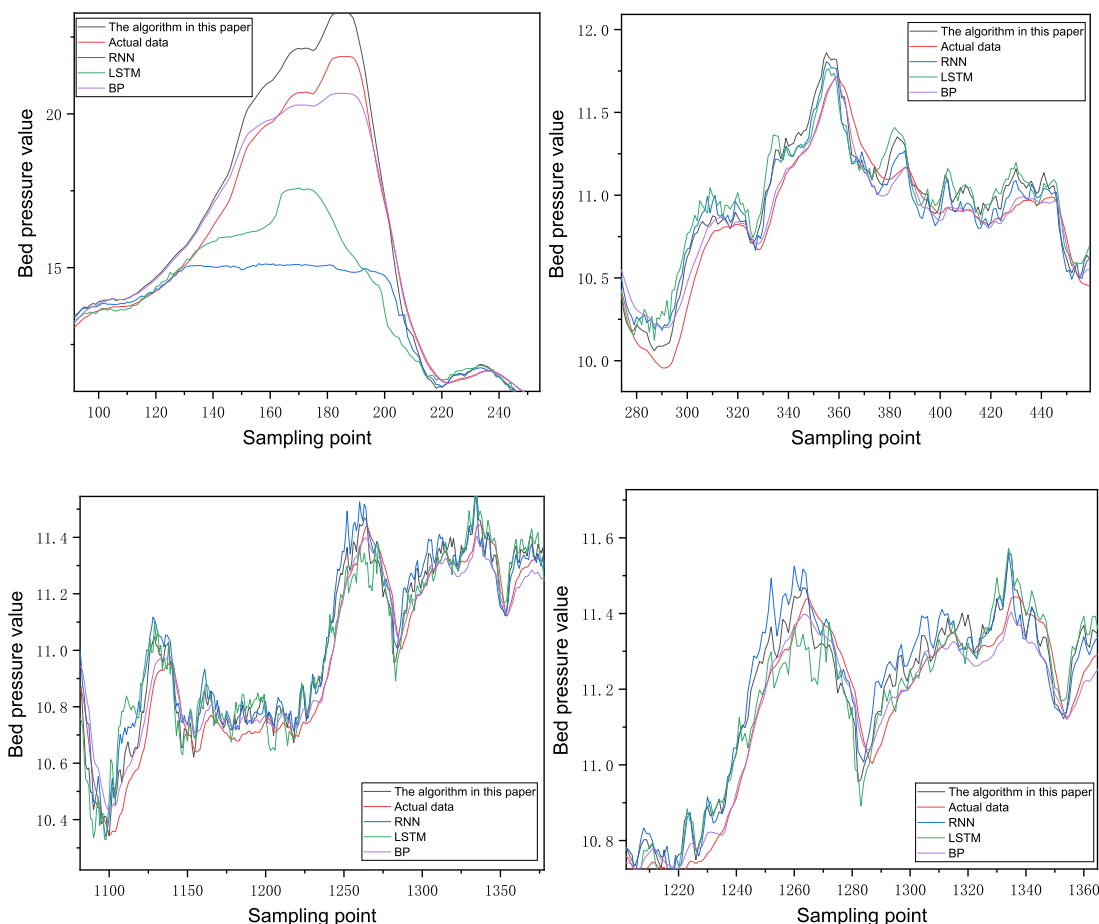


FIGURE 9. Data normalization after Sigmoid processing.

asynchronization should be noticed and arranged. This paper analyses variable correlation with DTW algorithm. What’s more, in order to reduce the influence of inertia, the article adjusts the time matching of the data based on the DTW path. As shown in Fig. 6, line  $y = x + b$  gets close to the DTW path. With this procedure, the input data would combine the data from DTW path and learn the mapping relationship based on the LSTM sequential processing.

Taking the lower left secondary air as an example, since the model chos 100 lookback, the closing line in the DTW image should contain the DTW path as much as possible. The part between the dotted red line and the solid line represents the input region of the model. This part should contain DTW paths as much as possible to ensure that the input model data contains correlation information between variables. After data processing, we could see the prediction from the model. Here the article compares the results from seven different algorithms.

**C. NORMALIZATION WITH SIGMOID FUNCTION**

The framework will normalize the data and perform Sigmoid nonlinear conversion according to the normalization result. The result of the left primary wind after normalization is illustrated in Fig. 7, in which the data almost presents a straight line with the occurrence of few abnormal points data

in the whole data set. According to the result, we select 0.996 and 1 as the values of a and b respectively. After Sigmoid processing, the data are shown in Fig. 8. It is obvious that the data trend has been enlarged, enabling it to be clearer and conducive to model training. Similarly, all the screened variables are processed.

**D. COMPARISON AND DISCUSSION OF THE RISK PREDICTION PERFORMANCES**

To develop the model, we choose Mean Square Error (MSE) as the loss function of the model, train the model with 20000 sets of data, and test with 1440 sets of data. Three common neural networks models are selected for comparison, including BP neural network, RNN neural network and LSTM neural network model. The comparison results are shown in the Fig. 9.

As illustrated in these figures, best performance, meaning the fastest speed to adjust trend when breakdown happens, occurs in the approach presented in this paper, the bed pressure drop prediction model. The presented model can effectively predict bed pressure, with an average prediction advance time of about 25s. In the normal status, since the normal data take up a large proportion, the advantage of differential prediction is not obvious. As shown in the chart, the prediction accuracy of RNN and LSTM algorithm is



TABLE 3. Average advance value.

Algorithm	Failure condition	Normal operating conditions
The algorithm in this paper	5.23	5.56
RNN	4.13	5.61
LSTM	4.02	5.45
BP	4.75	1.98

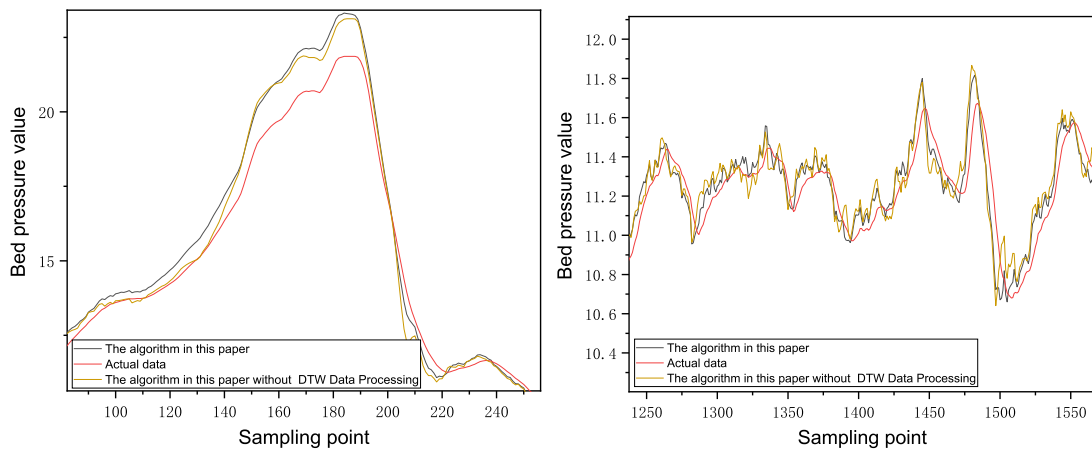


FIGURE 10. The comparison of processed data and unprocessed data.

fluctuated while normal running, and the accuracy is low when the breakdown happens. BP algorithm has high accuracy at the fault point. However, it behaves worse than other algorithms while normal running. The algorithm raised by the article could guarantee the accuracy while normal running and response quickly when the fault happens at the same time. The stability of its prediction is very helpful to the operation of the field person.

Compared with the conventional error function, which employs the Euclidean distance among data to obtain accuracy, the advance time weights more in practical application. Hence, we use the concept of advance degree to characterize the result of model prediction. The formula for the advance is as follows:

$$n = \max_{p \in (t-N, t)} \{t_r - t_p\}, |Y_p - Y_r| \leq \varepsilon \quad (14)$$

According to the formula, the four algorithm models are calculated, and the result is shown in the Tab.3

As illustrated in Tab.3, it can be seen that the algorithm in this paper has the best advance at the fault and can ensure a certain advance during normal operation. This is consistent with the results of previous analysis of the resulting images.

In this paper, the data processing also plays a very important role in the prediction accuracy of the model.

Fig. 10 shows the comparison of processed data and unprocessed data. It is observed that the prediction under normal operation is smoother after the data processing with DTW algorithm, and the prediction for faults is much more accurate than the model without data processing.

## VI. CONCLUSION

A data-driven framework for time series risk prediction has been proposed and validated with a case of bed inventory overturn prevention in a pant-leg CFB boiler. The underlying principle in the proposed framework is to excavate the time series characteristics between variables. Instead of conventional Euclidean metric, DTW employs the data-difference-discrimination pattern matching to reduce the calculation and improve the adaptability of variables in different dimensions. After variable processing of the time series data by the variant DTW algorithm, the bed pressure drop prediction model is established based on the LSTM structure. The framework performance is discussed and validated with real operational data. The tested variant DTW approach found consistent correlations between variables and the input model data. The LSTM model can effectively predict bed pressure, with an average prediction advance time of about 25s. Compared with other common neural network models, including BP, RNN and LSTM, the bed pressure drop

prediction model has the best performance both during normal running and when fault happens. Therefore, the proposed framework does have the ability to leave enough time for boilers to adjust the primary air volume prevent the occurrence of bed inventory overturn risks, thus protecting CFB boilers to operate stably.

## REFERENCES

- [1] S. Rahiala, K. Myöhänen, and T. Hyppänen, "Modeling the behavior of limestone particles in oxy-fuel CFB processes," *Fuel*, vol. 127, no. 7, pp. 141–150, Jul. 2014.
- [2] A. Nikolopoulos, I. Malgarinos, N. Nikolopoulos, P. Grammelis, S. Karrelas, and E. Kakaras, "A decoupled approach for  $\text{NO}_x$ - $\text{N}_2\text{O}$  3-D CFD modeling in CFB plants," *Fuel*, vol. 115, no. 1, pp. 401–415, Jan. 2014.
- [3] R. Ding, J. Dong, M. Zhang, H. Yang, and J. Lv, "Field test of coal type adaptability on a 300 MW CFB boiler," *Powder Technol.*, vol. 274, pp. 180–185, Apr. 2015.
- [4] M. M. Marchetti and T. S. Czarnecki, "Alstom's large CFBs and results," in *Proc. 15th Int. Conf. Fluidized Bed Combustion*, Jacksonville, FL, USA, May 1999, pp. 673–683.
- [5] Q. Wang, Z. Luo, M. Fang, M. Ni, and K. Cen, "Development of a new external heat exchanger for a circulating fluidized bed boiler," *Chem. Eng. Process.*, vol. 42, pp. 327–335, Apr. 2003.
- [6] H. Wang, Y. Li, G. Qiu, G. Song, and W. Yang, "Measurement of gas–solids flow in loop seal and external heat exchanger in a circulating fluidized bed," *Powder Technol.*, vol. 266, no. 6, pp. 249–261, Nov. 2014.
- [7] F. Xiao et al., "Study on the characteristic of CFB boiler bed and steam temperature when using FBHE," *Boiler Technol.*, pp. 4–7, 2003.
- [8] J. Li, W. Wang, H. Yang, J. Lv, and G. Yue, "Bed inventory overturn in a circulating fluid bed riser with pant-leg structure," *Energy Fuels*, vol. 23, no. 5, pp. 2565–2569, May 2009.
- [9] J. W. Wang, "The processing methods of turning-over for two beds in 300 MW circulating fluidized bed boiler," *Energy Eng.*, vol. 5, 2007.
- [10] J. Liu, Z. Q. Dong, and J. S. Zhu, "Cause analysis of losing stability and turning over for two beds in a 300 MW CFB boiler," *Thermal Power Gener.*, vol. 36, no. 4, 2007.
- [11] P. Basu and L. Cheng, "An analysis of loop seal operations in a circulating fluidized bed," *Chem. Eng. Res. Des.*, vol. 78, no. 7, pp. 991–998, Oct. 2000.
- [12] X.-B. Qi, W.-X. Huang, and J. Zhu, "Comparative study of flow structure in circulating fluidized bed risers with FCC and sand particles," *Chem. Eng. Technol.*, vol. 31, no. 4, pp. 542–553, Apr. 2008.
- [13] J. Li, H. Zhang, H. Yang, Q. Liu, and G. Yue, "The mechanism of lateral solid transfer in a CFB riser with pant-leg structure," *Energy Fuels*, vol. 24, no. 4, pp. 2628–2633, Apr. 2010.
- [14] G. B. Zhao, Y. F. Shi, and H. R. Yu, "Chaotic prediction of pressure fluctuation in fluidized bed," *J. Chem. Ind. Eng. China*, vol. 21, no. 5, 2000.
- [15] A. Gungor, "Predicting axial pressure profile of a CFB," *Chem. Eng. J.*, vol. 140, nos. 1–3, pp. 448–456, Jul. 2008.
- [16] Y. Lv, F. Hong, T. Yang, F. Fang, and J. Liu, "A dynamic model for the bed temperature prediction of circulating fluidized bed boilers based on least squares support vector machine with real operational data," *Energy*, vol. 124, pp. 284–294, Apr. 2017.
- [17] B. Li, J. Zhang, Y. He, and Y. Wang, "Short-term load-forecasting method based on wavelet decomposition with second-order gray neural network model combined with ADF test," *IEEE Access*, vol. 5, pp. 16324–16331, 2017, doi: [10.1109/ACCESS.2017.2738029](https://doi.org/10.1109/ACCESS.2017.2738029).
- [18] Z. Cao, C. Wan, Z. Zhang, F. Li, and Y. Song, "Hybrid ensemble deep learning for deterministic and probabilistic low-voltage load forecasting," *IEEE Trans. Power Syst.*, vol. 35, no. 3, pp. 1881–1897, May 2020, doi: [10.1109/TPWRS.2019.2946701](https://doi.org/10.1109/TPWRS.2019.2946701).
- [19] Z. Que and Z. Xu, "A data-driven health prognostics approach for steam turbines based on XGBoost and DTW," *IEEE Access*, vol. 7, pp. 93131–93138, 2019, doi: [10.1109/ACCESS.2019.2927488](https://doi.org/10.1109/ACCESS.2019.2927488).
- [20] G. Li, H. Wang, S. Zhang, J. Xin, and H. Liu, "Recurrent neural networks based photovoltaic power forecasting approach," *Energies*, vol. 12, no. 13, p. 2538, Jul. 2019.
- [21] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [22] M. Kim, B. Cao, T. Mau, and J. Wang, "Speaker-independent silent speech recognition from flesh-point articulatory movements using an LSTM neural network," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 25, no. 12, pp. 2323–2336, Dec. 2017, doi: [10.1109/TASLP.2017.2758999](https://doi.org/10.1109/TASLP.2017.2758999).
- [23] Y. Xie, R. Liang, Z. Liang, C. Huang, C. Zou, and B. Schuller, "Speech emotion classification using attention-based LSTM," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 27, no. 11, pp. 1675–1685, Nov. 2019, doi: [10.1109/TASLP.2019.2925934](https://doi.org/10.1109/TASLP.2019.2925934).
- [24] Y. Cui, S. Wang, and J. Li, "LSTM neural reordering feature for statistical machine translation," *Comput. Sci.*, vol. 2, 2015.
- [25] M. Gao, J. Li, F. Hong, and D. Long, "Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM," *Energy*, vol. 187, no. 15, pp. 115–128, 2019.
- [26] A. Rahman, V. Srikumar, and A. D. Smith, "Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks," *Appl. Energy*, vol. 212, pp. 372–385, Feb. 2018.
- [27] F. Hong, D. Long, J. Chen, and M. Gao, "Modeling for the bed temperature 2D-interval prediction of CFB boilers based on long-short term memory network," *Energy*, vol. 194, no. 1, pp. 116–131, 2020.
- [28] B. Xiong, X. Lu, R. S. Amano, and H. Liu, "Gas–solid flow in an integrated external heat exchanger for CFB boiler," *Powder Technol.*, vol. 202, nos. 1–3, pp. 55–61, Aug. 2010.



**FENG HONG** received the B.S. degree in automation and the Ph.D. degree in control theory and control engineering from North China Electric Power University, Beijing, in 2014 and 2019, respectively.

He was a Research Assistant with the State Key Laboratory of Alternate Electric Power System with Renewable Energy Sources. Since 2019, he has been a Lecturer with the School of Control and Computer Engineering, North China Electric Power University, Beijing. He is the author of more than 20 articles. His research interest includes the optimal control and engineering, operation condition monitoring and fault diagnosis of power generation systems. He has hosted two funds as the Fundamental Research Funds for the Central Universities (2020MS013), and has participated five funds as the National Key Research and Development Program of China.



**JIYU CHEN** received the B.S. degree in automation from North China Electric Power University, Beijing, in 2018, where he is currently pursuing the Ph.D. degree in control theory and control engineering. His main research field is the application of artificial intelligence algorithm in industrial process.



**ZHIYU ZHANG** is currently pursuing the B.S. degree in telecommunications with management with the Beijing University of Posts and Telecommunications (BUPT), China.

From 2018 to 2019, he was a member of JP Innovation Project. His research interest includes the application of deep learning in thermal power generation and the network communication.

Mr. Zhang is a member in the joint-program of BUPT and NCEPU. He has participated some of the researches in the group.



**RUI WANG** was born in Shandong, China, in April, 1999. She is currently pursuing the degree in telecommunications engineering with management with the International School, Beijing University of Posts and Telecommunications (BUPT).

From 2018 to 2020, she was selected into Yepeida Innovation and Entrepreneurship College of BUPT, getting education in innovation and artificial intelligence. In Summer 2019, she went to the University of Cambridge for Summer School, learning artificial intelligence and entrepreneurship. For the three years in BUPT, she has won second-class scholarship and first-class scholarship. Under the guidance of her mentor, her research interest is employing some artificial intelligence algorithms in evaluating and adjusting the performance of boilers in electric grids.



**MINGMING GAO** was born in Shanxi, China, in 1979. He received the B.S. degree in computer science and technology from Central South University, Changsha, in 2002, the M.S. degree in computer software and theory from Central South University, Changsha, in 2005, and the Ph.D. degree in control theory and control engineering from North China Electric Power University, Beijing, in 2013.

He is currently an Associate Professor with the School of Control and Computer Engineering, North China Electric Power University, Beijing. He is the author of more than 40 articles. His research interest includes the optimal control and engineering and operation condition monitoring of thermal power generation systems.

• • •