Main Parameters Prediction of the Hot Water Boiler Based on the LSTM Neural Networks

Sijie Liu\textsuperscript{a}, Chengke Guo\textsuperscript{b} and Ning Mei\textsuperscript{c}

Ocean university of China, Qingdao, China

\textsuperscript{a}Corresponding author e-mail:nmei@ouc.edu.cn, \textsuperscript{b}sijjeliu_123@sina.com, \textsuperscript{b}1069376393@qq.com

Abstract. This paper presents a data-driven method for the main parameter’s prediction of a hot water boiler. Principal component analysis is used to compress the input dimensions of the model and reserves the main information of the monitored parameters. The validity of the model is demonstrated by a case study of a coal water slurry circulating fluidized bed hot water oiler belong to a heating company. The historical data of the boiler is employed to establish a deep long short memory cell neural network as the prediction. The prediction results of the main parameters could fulfil the demand of the actual engineering.

1. Introduction

With the increasing safety and the control demands of the industrial process, the whole process are usually gained more attention in recent years.[1] An import control task in the management of the process is the Abnormal Events Management. The AEM contains the following contents: the timely detection of abnormal events, diagnosing its origins and then take appropriate control decisions and actions bring the process back to normal, safe, operating state. For the AEM, there are three main methods: quantitative model-based methods, qualitative model-based methods and data-driven methods. The limitation of the quantitative model-based methods is that it could only handle linear, and a small scale of specific nonlinear cases. [2] Meanwhile, the qualitative model-based methods require a combination of the basic understanding of the physics of the whole process and the corresponding search strategies when the access to a large amount of data is not possible. [3] The data-driven methods are based on a significant amount of data which is now become possible for the model industrial plants that are industrial plants are often installed with a large number of sensors, aiming at monitoring and controlling the process based on those delivered data. [4]

As for the hot water boiler, the abnormal state could be defined that the monitored parameters of the equipment are out of the boundary of the acceptable ranges. [2] The first step of the AEM is to predict the main parameters of the hot water boiler in the next several minutes. An accurate prediction could provide an evidence for the supervised control finished by the operators of the hot water boiler. According to the method of the model predictive control, the prediction is an import input of the control system, which could make the whole plant working at a steady and high efficiency state. Meanwhile, the widely use of the distributed control system (DCS) makes it possible to acquire the monitored parameters and realize the work of supervised control.
For the data-driven method, there are two main types of the method: unsupervised method and supervised method. The supervised method needs to use the historical data which is labeled on the basis of some prior process knowledge to construct a learning model that could handle the classification problems with noisy or incomplete data. The unsupervised method could deal with the data without any prior process knowledge, which could be employed to do the selection of variables.

Inspired by the work of references [6-8], the paper establishes a dynamic model of the hot water boiler aiming at predicting the temperature and pressure of the output water. The principle component analysis is used to select the input variables.

2. Introduction of the basic structure and parameters of the hot water boiler

2.1. The structure and working principles of the hot water boiler

The hot water boiler is a regional heating boiler plant. As shown in Fig.1, the structure of the hot water boiler.

![Diagram of hot water boiler](image)

**Figure 1.** Structure of hot water boiler.

The boiler is coal-fired circulating fluidized bed boiler. The circulation type is forced circulation. The main part contains furnace, adiabatic cyclone separator, self-balance feedback valve and convection pass. The fuel is coal water slurry (CWS) that is stored in a special tank. The CWS is pumped by the fuel supply pump into the furnace. The air for the combustion is supplied by the primary fan and secondary fan. Air from the primary fan is preheated by the primary air preheater and induced into the furnace. Air from the secondary fan enters the furnace through the nozzle distributed on the front and rear furnace wall aiming at reinforcing the mixture and making up more air. The fuel and air are mixed and burned in the fluidized state in the furnace, and heat is exchanged with the heating surface. After the flue gas carrying a large number of unburned carbon particles and containing a large number of materials enters the adiabatic cyclone separator through the furnace outlet, most of the materials are separated and returned to the furnace through the return materials to realize the circular combustion. The separated flue gas is discharged from the tail flue through the steering chamber, high temperature economizer, low temperature economizer and primary and secondary air preheater.

The main goal of the hot water boiler is to provide users with qualified hot water which has the specific temperature and pressure.
2.2. The main parameters of the hot water boiler
As shown in Table 1, the main working parameters of the hot water boiler.

| Number | Parameters                          | value     |
|--------|------------------------------------|-----------|
| 1      | Thermal power rating               | 29MW      |
| 2      | Rated outlet water pressure         | 1.6Mpa    |
| 3      | Rated outlet water temperature     | 70°C      |
| 4      | Rated inlet water temperature      | 130°C     |
| 5      | Design thermal efficiency          | 90%       |
| 6      | Exhaust gas temperature            | 130°C     |
| 7      | Exhaust excess air coefficient     | 1.3       |

For the hot water boiler, there are 224 parameters monitored by the DCS. There are 26 parameters could be controlled by the operators and 3 parameters about the backwater from the users. As shown in table 2, there 29 parameters are chosen as the initial input parameters of the prediction model.

| Number | Name in database | parameters                                      |
|--------|------------------|-------------------------------------------------|
| 1      | PT131            | Primary fan outlet pressure                      |
| 2      | PT132            | Secondary fan outlet pressure                    |
| 3      | PT133            | Primary air pressure after preheater            |
| 4      | PT134            | Secondary air pressure after preheater          |
| 5      | PT135            | Primary hot air pressure                         |
| 6      | PT136            | Circulating air pressure                        |
| 7      | PT137            | Circulating air pressure                        |
| 8      | PT141A           | Wind room pressure A                            |
| 9      | PT141B           | Wind room pressure B                            |
| 10     | PT152            | Induced draft fan pressure                      |
| 11     | FT131            | Flow rate of primary air                        |
| 12     | FR132            | Flow rate of secondary air                      |
| 13     | FT133A           | Flow rate of preheated primary air              |
| 14     | FT133B           | Flow rate of preheated primary air              |
| 15     | FT134A           | Flow rate of circulating air A                  |
| 16     | FT134B           | Flow rate of circulating air B                  |
| 17     | TE131            | Air temperature at outlet of primary preheater  |
| 18     | TE132            | Air temperature at inlet of secondary preheater |
| 19     | TE133            | Air temperature at inlet of primary preheater  |
| 20     | TE134            | Air temperature at outlet of secondary preheater|
| 21     | TE141A           | Wind room temperature A                         |
| 22     | TE142B           | Wind room temperature B                         |
| 23     | PT174A           | CWS pump pressure A                             |
| 24     | PT174B           | CWS pump pressure B                             |
| 25     | FT171A           | Flow rate of CWS pump A                         |
| 26     | FT171B           | Flow rate of CWS pump B                         |
| 27     | PT101            | Boiler return water pressure                    |
| 28     | FT101            | Flow rate of Boiler return water                |
| 29     | TE101            | Boiler return water temperature                 |
For the LSTM prediction model, the 29 input variables are too many. A lot of input redundant information contains in the input parameters and it cost too much calculation power to establish such a complete prediction model. So that a dimensionality reduction is necessary for the input parameters, which could select the parameters contains more information.

3. Methodology

3.1. Principal component analysis for the input of the model

Principal component analysis (PCA) is widely used in the dimensionality reduction and feature extraction. It could extract the main information by mapping the high-dimensional original data into a low subspace. The eigenvalue of the centered covariance matrix reflects the amount of the information in the corresponding input variables.

The boiler return water pressure, flow rate of Boiler return water, boiler return water temperature is necessary for the prediction model. The scope of the data selection is the other 26 parameters. While the PCA is based on the basic feature of the data itself, so that the 26 input is divided into 4 groups according to the character of the variables. As shown in Table.3, the PCA results of the pressure parameters. Primary fan outlet pressure and secondary fan outlet pressure contain more than 99% information, these two pressure parameters are selected for the input of the model.

Table 3. PCA results of the pressure parameters.

| Number | parameters                        | eigenvalue | Information ration |
|--------|-----------------------------------|------------|--------------------|
| 1      | Primary fan outlet pressure        | 1328.35    | 84.2504%           |
| 2      | Secondary fan outlet pressure      | 247.353    | 15.6883%           |
| 3      | Primary air pressure after preheater | 0.8335   | 0.0528%            |
| 4      | Secondary air pressure after preheater | 0.1072  | 0.0068%            |
| 5      | Primary hot air pressure           | 0.0104     | 0.0007%            |
| 6      | Circulating air pressure           | 0.0009     | 0.0001%            |
| 7      | Circulating air pressure           | 0.0012     | 0.0001%            |
| 8      | Wind room pressure A               | 0.0054     | 0.0003%            |
| 9      | Wind room pressure B               | 0.0034     | 0.0002%            |
| 10     | Induced draft fan pressure         | 0.0037     | 0.0002%            |

As shown in Table.4, the PCA results of the flow rate parameters. Flow rate of primary air and flow rate of secondary air contain more than 99% information, these two flow rate parameters are selected for the input of the model.

Table 4. PCA results of the flow rate parameters.

| Number | parameters                    | eigenvalue | Information ration |
|--------|-------------------------------|------------|--------------------|
| 1      | Flow rate of primary air      | 506000     | 99.9164%           |
| 2      | Flow rate of secondary air    | 4030       | 7.8989%            |
| 3      | Flow rate of preheated primary air | 204  | 0.0003%            |
| 4      | Flow rate of preheated primary air | 0.00127 | 0.0000%            |
| 5      | Flow rate of circulating air A | 0.000171  | 0.0000%            |
| 6      | Flow rate of circulating air B | 0.000432  | 0.0000%            |

As shown in Table.5, the PCA results of the temperature parameters. Air temperature at outlet of primary preheater, air temperature at inlet of secondary preheater, air temperature at inlet of primary preheater and air temperature at outlet of secondary preheater contain more than 99% information, these four temperature parameters are selected for the input of the model.
Table 5. PCA results of the temperature parameters.

| Number | parameters                                           | eigenvalue | Information ration |
|--------|------------------------------------------------------|------------|--------------------|
| 1      | Air temperature at outlet of primary preheater       | 501.28     | 93.7397%           |
| 2      | Air temperature at inlet of secondary preheater     | 5.2069     | 0.9737%            |
| 3      | Air temperature at inlet of primary preheater       | 8.1040     | 1.5155%            |
| 4      | Air temperature at outlet of secondary preheater    | 19.5546    | 3.6567%            |
| 5      | Wind room temperature A                             | 0.3250     | 0.0537%            |
| 6      | Wind room temperature B                             | 0.2870     | 0.0608%            |

As shown in Table 6, the PCA results of the CWS pump parameters. CWS pump pressure A, flow rate of CWS pump A and flow rate of CWS pump B contain more than 99% information, these three CWS pump parameters are selected for the input of the model.

Table 6. PCA results of the CWS pump parameters.

| Number | parameters                          | eigenvalue | Information ration |
|--------|-------------------------------------|------------|--------------------|
| 1      | CWS pump pressure A                 | 0.048      | 96.2986%           |
| 2      | CWS pump pressure B                 | 0.00007    | 0.1456%            |
| 3      | Flow rate of CWS pump A             | 0.00123    | 2.4720%            |
| 4      | Flow rate of CWS pump B             | 0.00054    | 1.0839%            |

As a result, the input variables is compressed from 29 parameters into 14 parameters for the deep LSTM neural networks.

3.2. Deep LSTM neural networks

The LSTM is the improvement of the Recurrent Neural Network (RNN). RNN could only take the information of a short past time into consideration. The LSTM could forget the useless information and record the useful information in a long time series, and its output contains the information of the long time series and the information of the current input. As shown in Fig.2, the structure of the LSTM cell. The single layer LSTM is realized by connecting the LSTM cells in the time series. As shown in Fig.3, the single layer LSTM.

Figure 2. Structure of LSTM cell.
The first step of the LSTM is the forward calculation. The formula (1)-(6) is the forward calculation of the LSTM.

\[ f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \]  
(1)

The forget gate \( f_t \) decides to keep the information from the last state.

\[ i_t = \sigma(w_i[h_{t-1}, x_t] + b_i) \]  
(2)

The function of the input gate \( i_t \) is to bring the information of the current input variables into the LSTM cell.

\[ \tilde{c}_t = \tanh(w_c[h_{t-1}, x_t] + b_c) \]  
(3)

\[ c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \]  
(4)

The state gate \( c_t \) describes the current state of LSTM cell, it combines the information from long time series and the input at this moment.

\[ o_t = \sigma(w_o[h_{t-1}, x_t] + b_o) \]  
(5)

The output gate \( o_t \) controls the impact from the long time series on the current.

\[ h_t = o_t \circ \tanh(c_t) \]  
(6)

The final output of the LSTM cell combines the information of the output gate \( o_t \) and the information of the state gate \( c_t \).

The LSTM neural network is trained by the error back propagation method. The error propagates in two directions. The first is Back propagation Through Time (BPTT). The second direction that the error back propagates from the output to the input at one moment.

The error propagates to the k moment could be described as follows:

\[ \delta^T_k = \prod_{j=k}^{t-1} \delta^T_{f,j}W_{fh} + \delta^T_{i,j}W_{ih} + \delta^T_{c,j}W_{ch} + \delta^T_{o,j}W_{oh} \]  
(7)

The error propagates from the output to the l-1 layer could be described as follows:

\[ \delta^{l-1}_t = (\delta^T_{f,t}W_{fh} + \delta^T_{i,t}W_{ih} + \delta^T_{c,t}W_{ch} + \delta^T_{o,t}W_{oh})f^{l-1}(W_c x_t + b_o)^{-1}) \]  
(8)
For the hot water boiler, the single layer LSTM is not enough to describe such a complex and highly nonlinear system. As shown in Fig.4, the deep LSTM neural network is appropriate to establish this prediction model. [9].

![Figure 4. Structure of Deep LSTM neural networks](image)

### 3.3. Data pre-process and the selection of hyperparameters

#### 3.3.1. Data pre-process. The dataset is collected from a heating company in Qingdao. The data is the monitored parameters in 30 days and the whole data is a matrix in the size of (97022,224). While the first 80000 is taken as the training data and the 80000-90000 is taken as the testing data.

The pre-process contains two steps: normalization and sliding window processing. The normalization is finished by the mean-variance normalization. The timestep for The LSTM model is 30, which means that the model predicts the future 15 minutes using the data from the past 15 minutes. After the sliding window processing, the training data matrix size is transformed from (80000,16) to (79940,30,16). The size of the testing data matrix is (10000,16)

#### 3.3.2. The selection of hyperparameters. The input is 14 dimensions and the output is 2 dimensions. The timestep is 30. The network has 3 layers and each layer has 100 hidden nodes. The batch size is 500 and learning rate is 0.001.

Meanwhile, the L2 regularization is used and the loss function is the Mean Square Error (MSE) function. As shown in Fig.5, the whole training process.

![Figure 5. The training process of Deep LSTM](image)
4. Results and discussion.

4.1. The prediction of the supply water temperature

As shown in Fig. 6, the tendency of the prediction value is similar to the true value.

![Figure 6. The prediction of supply water temperature](image)

As shown in Fig. 7, the predicted value lags behind the true value prediction. Also, the predicted value is lower than the true value. The reason is clear that when the temperature changes sharply, the prediction is delayed and conservative.

![Figure 7. The prediction of supply water temperature](image)

4.2. The prediction of the supply water pressure

As shown in Fig. 8, when the boiler water supply pressure steps up or down and the range is small, the predicted trend of the boiler water supply pressure is consistent with the predicted value.

![Figure 8. The prediction of supply water pressure](image)
Figure 8. The prediction of supply water pressure

Figure 9. The prediction of supply water pressure

As shown in Fig.9, when the supply water pressure changes frequently, the prediction model is invalid.

4.3. The error analysis

The figures in the former section, is some representative parts of the prediction results. The error of the prediction is list in the Table 7.

| Number | parameters               | Mean absolute error | Mean relative error |
|--------|--------------------------|---------------------|---------------------|
| 1      | Supply water temperature | 0.750               | 0.952%              |
| 2      | Supply water pressure    | 0.0141              | 2.251%              |

The mean absolute error of the predicted supply water temperature is 0.750, and the mean relative error of the predicted supply water temperature is 0.952% The mean absolute error of the predicted supply water pressure is 0.0141, and the mean relative error of the predicted supply water pressure is 2.251%
5. Conclusion
In this work, the prediction model which is used to predict the supply water temperature and the supply water pressure is established. The PCA is used to compress the 29 dimensions into 14 dimensions and reserves the main information in the monitored parameters. A deep LSTM neural network is established to predict the supply water temperature and supply water pressure. The mean relative error of the predicted supply water temperature is 0.952% and the mean relative error of the predicted supply water pressure is 2.251%. It is clear that the accuracy of the prediction model could meet the engineering demand.

Acknowledgments
This work was financially supported by National Natural Science Foundation of China (51679225), Key Basic Research of Shandong Provincial (2017GHY15113).

References
[1] Jiang, Q. and X. Yan, Monitoring multi-mode plant-wide processes by using mutual information-based multi-block PCA, joint probability, and Bayesian inference. Chemometrics & Intelligent Laboratory Systems, 2014. 136 (9): p. 121-137.
[2] Venkat Venkatasubramanian, Raghunathan Rengaswamy b, Surya N. Kavuri, A review of process fault detection and diagnosis: Part I: Quantitative model-based methods. Computers & Chemical Engineering, 2003. 27 (3): p. 293-311.
[3] Venkat Venkatasubramanian, Raghunathan Rengaswamy b, Surya N. Kavuri, A review of process fault detection and diagnosis: Part II: Qualitative models and search strategies. Computers & Chemical Engineering, 2003. 27 (3): p. 313-326.
[4] Venkat Venkatasubramanian, Raghunathan Rengaswamy b, Surya N. Kavuri, A review of process fault detection and diagnosis Part III: Process history based methods [J]. Computers & Chemical Engineering, 2003, 27 (3):293-311.
[5] Ashok, P., et al., Improved Performance of Unsupervised Method by Renovated K-Means. Computer Science, 2013.
[6] Xiangyu Qing Yugang Niu.Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM.Energy [J]2018,148:461-468.
[7] Selvin S , Vinayakumar R , Gopalakrishnan E A , et al. Stock price prediction using LSTM, RNN and CNN-sliding window model[C]// 2017 International Conference on Advances in Computing,Communications and Informatics (ICACCI). IEEE, 2017.
[8] Li bing, zhang yan, liu shi. Research on short-term wind speed prediction based on LSTM [J]. Computer simulation, 2008,35 (11):456-461.
[9] Yoshua Bengio,Ian Goodfellow,Aaron Courville 《 deep learning 》 [M]:http://www.deeplearningbook.org/, 2015 284-290.