Data Analytics Based Dual-Optimized Adaptive Model Predictive Control for the Power Plant Boiler

Zhenhao Tang,1 Haiyang Zhang,1 Ping Che,2 Shengxian Cao,1 and Zhiyong Zhao3

1School of Automation Engineering, Northeast Electric Power University, Jilin 132012, China
2Department of Mathematics, College of Sciences, Northeastern University, Shenyang 110819, China
3Changchun Xinbocheng Technology Company, Changchun 130000, China

Correspondence should be addressed to Zhenhao Tang; zhenhao_tang@126.com

Received 22 September 2016; Accepted 12 January 2017; Published 13 February 2017

To control the furnace temperature of a power plant boiler precisely, a dual-optimized adaptive model predictive control (DoAMPC) method is designed based on the data analytics. In the proposed DoAMPC, an accurate predictive model is constructively by the hybrid algorithm of the least squares support vector machine and differential evolution method. Then, an optimization problem is constructed based on the predictive model and many constraint conditions. To control the boiler furnace temperature, the differential evolution method is utilized to decide the control variables by solving the optimization problem. The proposed method can adapt to the time-varying situation by updating the sample data. The experimental results based on practical data illustrate that the DoAMPC can control the boiler furnace temperature with errors of less than 1.5% which can meet the requirements of the real production process.

1. Introduction

Fossil-fuel-fired power plants can produce stable and controllable energy. Despite the quick development of sustainable power generation methods, such as solar power systems and wind power systems, the fossil-fuel-fired power generation system is and will be an important part of the power system. In the meantime, more effort must be made to reduce the waste gas emission of the thermal power production process in order to protect the environment. To realize this target, the boiler furnace temperature should be controlled to follow certain curves. However, there are several challenges that need to be overcome.

To start with, there is a lack of a dynamic and accurate model of the boiler furnace temperature. There are several mechanism models constructed for the boiler. Gao and Dai [1] extended a new linear model of the steam unit. The model showed a good performance in the dynamic analysis of the power system. Alobaid et al. [2] mentioned numerical models at different steady-state operation points. The models obtain relative error of less than 1% at the steady-state situation. They perform well in a certain situation, but when the situation changes, the accuracy of the model will drop significantly. Wu et al. [3] presented data-driven modeling based on subspace method. It shows good performance in the simulation, but the selection of the modeling method needs to be done by experience before the application. There were many other nonlinear models constructed for the boiler-turbine unit, such as [4, 5]. They considered the unit as a whole system, and there was no boiler model mentioned alone. Motivated by the successful usage of data-mining method in boiler wall temperature prediction [6], a predictive model is constructed based on the least square support vector machine in this paper. To improve the predictive accuracy, differential evolution algorithm is employed to optimized parameters for different problems dynamically.

Second, a modern control method should be tailored to fit the boiler furnace temperature control problem. The widely used control method in thermal power plants is PID control method. Although this method is easy to utilize and simple to understand, the control accuracy cannot meet the requirement. Besides the PID method, many other methods are utilized in boiler control. Park et al. [7] presented robust control for cogeneration plants. Keshavarz et al. [8] provided...
2 Mathematical Problems in Engineering

Figure 1: The production process of the boiler.

a discrete-time piecewise affine model and model predictive control method for the boiler-turbine unit. Heo et al. [9] and Li et al. [10] proposed intelligent control methods based on particle swarm optimization and genetic algorithm relatively. These methods mainly focus on the application of these modern methods, and little attention is paid to high accuracy modeling. Wu et al. [3] presented a predictive control method and demonstrated the effectiveness of the methods. However, the rolling optimization problem in model predictive control method was solved by a simple mathematic method which is inefficient in solving the nonlinear problem. In [11, 12], model predictive control methods like dynamic matrix control and nonlinear model predictive control were utilized relatively. Liu et al. [13] presented two alternative methods of exploiting the performance of model predictive control. These methods obtained good performance, but they paid little attention to the optimization of the predictive model. With the development of the model predictive method, a generalized predictive control method was proposed and utilized in boiler steam temperature [14]. The generalized predictive control method provides a more flexible structure and fewer constraints of the predictive model.

Motivated by the two mentioned issues, we proposed a dual-optimized adaptive model predictive control (DoAMPC) method. The proposed DoAMPC has the following main features compared with the previous researches:

1. The DoAMPC adopts LSSVM to construct the predictive model for the boiler furnace temperature with high accuracy rapidly. This model considered the main state variables and control variables. Additionally, to improve the prediction accuracy, differential evolution (DE) is utilized to optimize the parameters of LSSVM for each different problem.

2. The DoAMPC includes a rolling optimization problem with a black-box model and bound constraint conditions. Furthermore, DE is employed to solve this nonlinear problem to get optimized control variable values.

3. The DoAMPC presents a new way to realize model predictive control in practical problems by utilizing data-driven algorithm and intelligent optimization algorithms. Instead of the canonical linear predictive model and linear optimization problem, the nonlinear predictive model and optimization problem is constructed and solved in this method.

The remainder of this paper is organized as follows. Section 2 describes some background knowledge of the boiler control problem, LSSVM, and DE briefly. The details of the proposed DoAMPC are provided in Section 3. In Section 4, some practical cases are employed to testify the performance of the proposed DoAMPC by comparing with other common methods. The conclusions based on the present studies are drawn in Section 5.

2. Background Knowledge

2.1. Boiler Temperature Control Problem. The production process is shown in Figure 1. The boiler turns the fossil energy into thermal energy. And the steam generated by the boiler is transferred to the engine for further usage. There are dozens of variables in this process. Most of these variables can be separated into two types briefly. The control variables, such as feedwater flow, are variables that can be controlled by the operator directly. And to set these control variables properly is the main purpose of this paper. The state variables, such as oxygen content, are the variables detected to monitor the production process. These state variables can reflect the situation of the process. Actually, all the variables contain the information that influences the boiler furnace temperature. So, these variables are employed as the input variables in the proposed DoAMPC.

2.2. LSSVM. Suykens and Vandewalle [15] proposed LSSVM to improve the calculation efficiency of SVM. LSSVM is a supervised learning method. So, a training dataset is needed to construct a model. Assume that \( T = \{ (\hat{x}_{\text{num}}, y_{\text{num}}) \} \) \( \text{num} = 1, 2, \ldots, n \) is the training dataset, where \( n \) denotes the number of samples; \( \hat{x}_{\text{num}} \in \mathbb{R}^{nf} \) denotes the input and \( y_{\text{num}} \in \mathbb{R} \) denotes output. \( nf \) is the number of input features. Then, the predictive model can be described as formula (1). Hence,

\[
\hat{f}(\vec{x}) = \sum_{\text{num}=1}^{n} \alpha_{\text{num}} K(\hat{x}_{\text{num}}, \vec{x}) + b,
\]  

(1)
where \( f(\vec{x}) \) means the predictive value of a new input \( \vec{x} \). \( K(\vec{x}_{\text{num}}, \vec{x}) \) means the kernel function which provides outstanding ability to deal with nonlinear problems. The main target of LSSVM is to obtain the coefficients \( \alpha_{\text{num}} \) and \( b \). Then, the predictive model is established.

The selection of the kernel parameters influences the predictive accuracy significantly. To maintain the predictive accuracy in a different situation, DE is employed to optimize the parameters dynamically.

2.3. DE. DE was proposed by Storn and Price [16]. To solve an optimization problem with DE, the search area should be specified. Then, the swarm is initialized randomly. Assume that \( \vec{p}_{\text{g}}(i = 1, 2, \ldots, np) \) is the \( i \)th particle in \( g \) iteration. Then, the fitness value of \( \vec{p}_{\text{g}} \) is calculated considering the objective function of the optimization problem. Then, the particles are updated following certain rules to improve the quality of the swarm. After certain iterations, which depend on the optimization problem, the swarm should converge to a relatively small area. Finally, the particle information with the best fitness value is considered to be the solution of the optimization problem.

3. Dual-Optimized Adaptive Model Predictive Control (DoAMPC) Method

The dual-optimization model predictive control method is proposed in this section. In Section 3.1, the relationship of differential evolution algorithm and LSSVM and the main procedure are shown. Then, the data preparation method used in this paper is given in Section 3.2. Consequently, the predictive model based on optimized LSSVM is described in Section 3.3. Additionally, the optimization model is constructed and solved in Section 3.4. Finally, the feedback correction method is given in Section 3.5.

3.1. The Main Procedure of DoAMPC. The main procedure of DoAMPC is shown in Figure 2.

Firstly, the practical data (PD) collected from power industry are preprocessed to reduce the noise and guarantee the effectiveness of the sample data (SD).

Then, to realize the predictive control of boiler furnace temperature, a predictive model of boiler furnace temperature must be constructed. Due to the nonlinearity and time-varying boiler furnace temperature, a black-box model based on data analytics is conducted. The control variables \( u(t) \) and process variables \( x(t) \) are utilized as inputs to obtain the boiler furnace temperature (bt(t)). \( t \) used as the subscript means the time index of each variable. In this model, DE algorithm is utilized to improve the model accuracy by solving an optimization problem.

Thirdly, an optimization problem based on the black-box model and several constraint conditions is given and solved by DE. The solution is the optimized set value of control variables \( u(t) \).

Finally, a model correction method is proposed to improve the control accuracy.

3.2. Preprocess the Practical Data. The practical data are read from the database which stores the real process variables. These data can reflect the state of the boiler which is precious to construct the prediction model. However, due to the tough process environment, there are data missing, outliers in practical data. To ensure the accuracy of the prediction model, the practical data must be preprocessed. In this paper, we utilize two different methods to deal with the missing data and outliers, respectively.

First, to deal with the data missing, the cubic spline interpolation is employed. Actually, the state variables of the boiler are continuous. The cubic spline interpolation is good at solving this kind of question. It can recover the missing data with relatively smooth data. Additionally, the cubic spline interpolation is easy to program with good stability and convergence property.

Second, the outliers are deleted and replaced by the cubic spline interpolation. The 3σ criterion is utilized to find out the outliers. 1000 continuous samples are used to calculate
the mean ($\bar{m}$) and variance ($\delta$) of every variable. If a variable value exceeds $\bar{m} \pm 3\delta$, then this variable value is considered as an outlier. The reason for choosing this method is that this method can perform calculation with high efficiency to deal with a large amount of data. To ensure the effectiveness of this method, the amount of the data should be large. In fact, there is a large amount of practical data collected from the production process every minute.

After the preprocessing of the practical data, the sample data are utilized in the following procedure. The preprocessing is done repetitively when there are new data collected.

3.3. Construct the Predictive Model. As mentioned before, the traditional boiler furnace temperature models are not good at dealing with the dynamic situation and cannot be utilized in the model predictive control method directly. So, the black-box model based on DE and LSSVM is constructed.

DE is a swarm-based optimization algorithm which was proposed by Storn in 1995. In this algorithm, several particles made up of parameters that need to be determined are initialized with random values within a constraint. Then, the fitness of each particle is calculated. And the particles are updated following certain formulas considering their fitness. Finally, the swarm converges to a global best location which is the solution of the optimization problem.

LSSVM was proposed in 1995 by Suykens as a variant of SVM. A linear programming problem instead of the quadratic programming problem is solved in LSSVM, which accelerates the calculation while inheriting the ability to solve nonlinear problems. Additionally, LSSVM shows an outstanding generalization ability. The selections of kernel function parameters and regularization parameter are the keys to maintain the predictive accuracy. Due to the complexity of the selection, an optimization problem is constructed and then solved by DE.

Assume that the sample data used to train the predictive model is SD = \{(\bar{u}_t, b_t)\}, $t = 1, 2, \ldots, ns$, where $\bar{u}_t$ denotes the input of prediction model made up with the control variables $u_t$ and the state variables $x_t$. $b_t$ denotes the boiler furnace temperature. ns denotes the number of samples. Then, the predictive model obtained by LSSVM is $f(\bar{t}_{in}) = \sum_{i=1}^{ns} a_i K(\bar{in}, \bar{in}_t) + b$, where $\bar{in}$ denotes new input data, $f(\bar{t}_{in})$ means the prediction value of $\bar{in}$, $\bar{a}_i$ and $b$ are the coefficients calculated by LSSVM, and $K(\bar{in}, \bar{in}_t)$ is the kernel. In this paper, we utilize radial basis function as the kernel. The formula is shown in (2). Whenever the LSSVM is utilized, the kernel parameter $\sigma^2$ and the penalty factor $\gamma$ influence the prediction accuracy significantly. The best parameters for the different problem are verified. To obtain the best parameters dynamically, DE is employed. Hence,

$$K(\bar{in}, \bar{in}_t) = \exp \left(-\frac{\|\bar{in} - \bar{in}_t\|^2}{2\sigma^2}\right). \quad (2)$$

In the hybrid of DE and LSSVM, DE calculates the candidates of parameters while the LSSVM offers the fitness value of each candidate. To testify the accuracy of the predictive model, new data which are different from SD are prepared. Assume that TD = \{(\bar{u}_t, b_t), k = 1, 2, \ldots, nt\}, is the new database, where nt denotes the number of test data. The optimization problem to select the best parameters is to choose the parameters which can allow the prediction accuracy to minimize. The mathematical representation is given as

$$\min \sum_{k=1}^{nt} |b_t - f(\bar{in}_k)|$$

s.t. $f(\bar{in}_k) = \sum_{i=1}^{ns} a_i K(\bar{in}_k, \bar{in}_t) + b$

$$K(\bar{in}_k, \bar{in}_t) = \exp \left(-\frac{\|\bar{in}_k - \bar{in}_t\|^2}{2\sigma^2}\right) \quad (3)$$

$\sigma^2 \geq 0$;

$\gamma \geq 0$;

$k = 1, 2, \ldots, nt$.

To solve this problem with DE, the following steps should be done.

Step 1 (initialize DE swarm and parameters). Assume that $p_i^0$ ($i = 1, 2, \ldots, np$) is the ith particle. np denotes the total number of particles in the gth iteration. g means the iteration times. Set $g = 1$ in this step. $p_{i,1}^g$ and $p_{i,2}^g$ represent the kernel parameter $\sigma^2$ and the penalty factor $\gamma$, respectively. $p_{i,1}^g$ and $p_{i,2}^g$ are initialized randomly by the following equation:

$$p_{i,1}^1, p_{i,2}^1 = 5000 + \text{rand}(0, 1) \ast 5000, \quad (4)$$

where $\text{rand}(0, 1)$ means a random number obeying normal distribution.

Step 2 (calculate the fitness of particles). To calculate the fitness of particles, the kernel parameter $\sigma^2$ and the penalty factor $\gamma$ are set based on each particle firstly. Then, a predictive model based on sample data is constructed. Calculate the objective function in (3). Next, the value is considered as the fitness of the corresponding particle. Finally, let the global best location $p^g_{gb}$ be equal to the best particle.

Step 3 (exchange the information among different particles). To update particles, the particles change their location by the following equation:

$$\bar{u}^{g+1}_i = p^g_i + c1 \ast r1_g \ast \left(\bar{p}^{g+1}_{gb} - p^g_i\right) + c2 \ast r2_g \ast \left(p^g_{gb} - \bar{p}^{g+1}_i\right), \quad (5)$$

where $\bar{u}^{g+1}_i$ denotes a new candidate particle; $c1$ and $c2$ denote the scale factor to control the update velocity;
integer in the range $[0,1]$ mean random number in the range $[0,1]$ obeying uniform distribution. $r3$ and $r4$ are a random integer in the range $[1, np]$ which is different from $i$. These numbers are generated every iteration.

**Step 4** (select the particles). After the update, run Step 2 to calculate the fitness of candidate particles. Compare the fitness of $u_i^{gr+1}$ and $p_i^g$ and let the better particle be $p_i^{gr+1}$.

**Step 5.** If one of the termination conditions is met, stop the algorithm and output the predictive model. If not, go to Step 3.

The termination conditions utilized in this paper are as follows: (1) the fitness of $p_{tg}$ is smaller than 0.1% and (2) $g$ reaches a specified integer $G$.

The outputs stored in a specified file for further application include the best parameters, the coefficients $\alpha_\text{ij}$ and $b_j$, and the sample data.

### 3.4. Construct and Solve a Rolling Optimization Problem

To control the boiler furnace temperature, we attempt to control the temperature to change following a certain curve. This curve named reference curve is given considering the practical requirement. In other words, the objective target is to minimize the error between the outputs and reference curve.

Then, the objective function of the rolling optimization problem is set as the following formula:

$$
\min \sum_{j=1}^M |f_j(\overrightarrow{\text{in}}) - r_j|,
$$

where $M$ denotes the predictive length and $f_j(\overrightarrow{\text{in}})$ denotes the predictive value of $\overrightarrow{\text{in}}$. To realize the prediction of $M$ steps, $M$ predictive models need to be constructed. In the input $\overrightarrow{\text{in}}$, there are two parts. The first part includes the proposal control variables $\overrightarrow{u}$, and the second part includes the state variables at the prediction moment. $\overrightarrow{u}$ is the solution which needs to be optimized. Considering the predictive model constructed in Section 3.3 and the bound constraint of $\overrightarrow{u}$, the rolling optimization problem is shown in the following formula:

$$
\min \sum_{j=1}^M |f_j(\overrightarrow{\text{in}}) - r_j|
$$

s.t. $f_j(\overrightarrow{\text{in}}) = \sum_{i=1}^{\text{in}} \alpha_{ji} \overrightarrow{K}(\overrightarrow{\text{in}}, \overrightarrow{\text{in}}) + b_j$;

$$
\overrightarrow{u}^L \leq \overrightarrow{u} \leq \overrightarrow{u}^U; \\
j = 1, 2, \ldots, M; \\
t = 1, 2, \ldots, \text{ns},
$$

where $\alpha_{ji}$ and $b_j$ denote the coefficients of the $j$th predictive model; $\overrightarrow{u}^U$ denotes the upper bound of the control variables; $\overrightarrow{u}^L$ denotes the lower bound of the control variables.

To solve the problem, DE is utilized to obtain the optimized control variables. The flowchart of the procedure is shown in Figure 3. The content of each step is given in Figure 3.

**Initialize the swarm.** As in Section 3.3, assume that $p_{tg}^i$ $(i = 1, 2, \ldots, np)$ is the particle. However, $p_{tg}^k$ represents the $k$th control variable which is different from Section 3.3. $p_{tg}^k$ is initialized randomly by the following equation:

$$
p_{tg}^k = u_k^L + \text{rand}[0,1] \ast (u_k^U - u_k^L),
$$

where $\text{rand}[0,1]$ means a random number obeying uniform distribution. This number is different for each $k$.

To calculate the fitness of each particle, the predictive values are calculated. Then, the error between the predictive outputs and reference curve is calculated by formula (6).

The terminal conditions and the method to update the particle are the same as the ones in Section 3.3.

The result outputted by DE includes the optimized control variables and the objective value for further analysis.

#### 3.5. Model Correction

The state of the boiler is dynamic. And the boiler load should change with the requirement. In this case, the prediction model needs to be corrected if the predictive error exceeds the acceptable range. In this paper, considering that the modeling can be done in seconds, a reconstruction strategy is employed.

If the relative predictive error is bigger than 1% for continuous $T$ sample period, the prediction model is reconstructed with updated sample data.
4. Computational Results

To testify the performance of the proposed algorithm, experiments based on the practical production data are carried out. Both the accuracy of the predictive model and the effectiveness of the DoAMPC method are testified. All the algorithms are implemented using VC++. All the following experiments are run on a PC with Intel Core i7-4712MQ CPU (2.30 GHz), 4.00 GB RAM, and Windows 10 operating system.

4.1. Experiment Design. The data utilized in the following experiments are collected from a Chinese thermal power plant. The details of the data are shown in Table 1. The inputs include the control variables (such as feedwater flow, pulverized coal flow, and wind flow) and the state variables (the boiler furnace temperature, oxygen content, etc.). The control object is the temperature of the boiler. Case 1 includes a step change in the boiler furnace temperature which can illustrate the sudden change in practical production. Case 2 and case 3 represent the frequent continuous change of boiler furnace temperature.

To utilize DE to solve the two optimization problems, some of the DE parameters need to be set up. By the commonly used trial-and-error method, the parameters are set up as follows: \( np = 100, c_1 = 1.4, c_2 = 1.9, \) and \( G = 1000 \).

4.2. The Experimental Results of the Predictive Model. To verify the accuracy of the proposed black-box model, several experiments based on the datasets mentioned in Section 4.1 are provided. And the other 3 widely used methods, radial basis function neural network (RBFNN), multilayer perception (MLP), and linear programming (LP), are utilized as comparison methods. The 5-fold method is employed. In this method, the dataset is divided into 5 subsets randomly, and 4 of the 5 subsets (including 400 instances) are utilized as the training data while the rest of the subsets (including 100 instances) are utilized as test data. The predictive results of the test data are shown in Figure 4. There are six subplots in the figure. Subplots (a)–(c) show the predictive errors of 4 different methods. Subplots (d)–(f) show the boxplot of the prediction errors of the methods in the 3 different applications.

The results in subplots (a)–(c) illustrate that (1) the prediction errors of the proposed method are smaller than other comparative methods. This shows the effectiveness of the proposed predictive method. (2) The predictive errors of the proposed method are smaller than 10 which can meet the requirement of the practical control. We also note that when the load of the unit changes, the prediction error becomes bigger than that in a stable situation. This may be caused by lack of the training data which can reflect the load change. In other words, most of the training data are collected in a stable situation. The number of samples which contain the information of load changing is smaller than that of stable samples. The results in subplots (d)–(f) indicate that (1) the proposed method can obtain stable high accuracy in three cases and (2) the predictive relative errors of the proposed method are below 1.0%.

All the results illustrate that the proposed modeling method can construct an accurate model for nonlinear problems.

4.3. The Experimental Results of DoAMPC. The control results of the proposed DoAMPC method and the widely used PID control method are shown in Figure 5. Three different datasets are utilized in the experiment. Scene 1 represents the situation of decreasing boiler temperature. Scene 2 represents the situation of normal production. Scene 3 represents the situation of increasing boiler temperature. 495 samples are utilized in each scene. There are also 4 subplots in this figure. Subplots (a)–(c) show the control results of 2 different methods and the reference curves. Subplot (d) shows the boxplot of the control errors of the proposed method in the 3 different applications.

From subplots (a)–(c), the following results can be concluded. (1) The control results of the DoAMPC method in three different situations follow the reference curves dynamically, which means the DoAMPC is effective. (2) Compared with the PID method, the control accuracy of the proposed method is higher in all different datasets. From subplot (d), we can find out that the control errors of the proposed method are less than 1%, and the average error is below 1.5%.

Similar to the predictive results, the control errors are bigger in the load change point than in a stable situation. This may be caused by the lag of model reconstruction. That is, the predictive model needs several sample times to reconstruct the model. However, the good performance of the reconstructed model demonstrates the effectiveness of the model correction strategy.

All the results demonstrate that the DoAMPC can handle nonlinear and dynamic problems with good performance. The generalization ability of the proposed method is outstanding. With the update of the sample data, this method can be utilized in other boilers in other thermal power plants.

5. Conclusion

A dual-optimized adaptive model predictive control method based on black-box model and DE algorithm is proposed in this paper in order to control the boiler furnace temperature accurately. One major feature of the proposed DoAMPC method is that it uses a black-box model based on LSSVM
and DE. The DE optimizes the parameters of the LSSVM to maintain the high predictive accuracy. Another main feature is that it contains a minimization problem with nonlinear model and bound constraint conditions for the boiler furnace temperature control problem. To solve this nonlinear problem, DE is employed to obtain the optimized control variables. The practical data are assessed to test the performance of the proposed methods. And the results show that the proposed strategies in both modeling and controlling are effective. The proposed method is comparative or superior.
to some commonly used algorithms. In the future, we will make more effort in improving the prediction accuracy and trying to utilize other data-mining methods in this structure.

Competing Interests
The authors declare that there are no competing interests regarding the publication of this paper.

Acknowledgments
This work was supported by the National Natural Science Foundation of China under Grants 61503072, 71402021, and 51376042, Northeast Electric Power University BSJXM-201434.

References


Submit your manuscripts at
https://www.hindawi.com