IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received July 31, 2019, accepted August 19, 2019, date of publication August 29, 2019, date of current version September 26, 2019.
Digital Object Identifier 10.1109/ACCESS.2019.2938394

Iterative ADP Optimal Control for GGBS Production Based on Dynamic Target Optimization

KANG WANG[1] AND XIAOLI LI[1,2,3]
[1]Beijing Key Laboratory of Computational Intelligence and Intelligent System, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China
[2]Beijing Advanced Innovation Center for Future Internet Technology, Beijing 100124, China
[3]Engineering Research Center of Digital Community, Ministry of Education, Beijing 100124, China

Corresponding author: Xiaoli Li (lixiaolibjut@bjut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61873006, Grant 61673053, and Grant 61473034, in part by the Beijing Major Science and Technology Special Projects under Grant Z181100003118012, and in part by the National Key Research and Development Projects under Grant 2018YFC1602704 and Grant 2018YFB1702704.

ABSTRACT In this paper, a novel optimal control scheme for ground-granulated blast-furnace slag (GGBS) production process is proposed by using the iterative adaptive dynamic programming (ADP) method and dynamic optimization of desired values. To handle the characteristic of changing operation modes and to obtain a practical optimal control result, this paper formulates the GGBS optimal control problem based on an intensive study of the production technique, and constructs the optimal control scheme based on the repetitive optimization strategy — ''target optimization'' and ''process optimization''. To obtain the optimal targets matching with changing operation modes, SVM modeling and multi-objective optimization method are utilized to solve the multi-objective optimization problem. In order to minimize the performance index function further and faster, iterative tuning technology is adopted to design the ADP based optimal control method. Finally, simulation is conducted to verify the effectiveness of the proposed approach.

INDEX TERMS Iterative adaptive dynamic programming, GGBS production process, multi-objective optimization.

I. INTRODUCTION

In the past decades, ground-granulated blast-furnace slag (GGBS) production has been caused widely attention. On one hand, derived from blast furnace slag – a kind of waste material from steel and iron making, GGBS provides a friendly and economic beneficial way to handle the waste. On the other hand, thanks to its unique physical and chemical structure, GGBS can be a kind of supplementary material into cement and improves many characteristics of concrete, such as durability, strength and corrosion resistance [1], [2]. However, in spite of relative high production yield in industrial field, due to the enclosed mill and changing operation modes, it is still difficult to design a stable control process, and to obtain uniform and high product quality.

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

Due to the characteristics of multiple variables, strong coupling and highly nonlinear, it is a open and hard problem to optimize and control the GGBS production. Typically, PID method is adopted to control the mill differential pressure to a desired value by tuning the feed material, so that the production works stably [3]. In recent years, intelligent control methods are widely utilized [4]–[7]. Based on the identified mathematical model, expert system and fuzzy control were combined to optimize the material layer thickness [5]. He et al introduced a domain adaptive random weight neural network to solve the problem of soft sensor for wet ball mill load parameters under multiple loading conditions [6], [7]. Though above methods realized stable control result of GGBS production, further optimization is needed to explore the ability of plant and obtain better quality and yields, at the same time, more stable operations. Combing neural network, dynamic programming, and reinforcement learning, adaptive dynamic programming (ADP) can approximate the optimal control signal of nonlinear system, meanwhile, avoid ''curse of dimensionality'' problem. In recent years, many scholars have conducted in-depth theoretical research and extensive application of ADP. Using the ELM based heuristic dynamic programming optimal control method, Li et al proposed a raw meal fineness control system of vertical mill grinding process [8]. Taking this strategy, material fitness can track the desired value at the same time minimize the performance index function. To explore the extreme production capacity under control constrains, Wang [9] and Shen [10] established and solved the multi-objective problem of GGBS production, selected optimal solutions were considered as the reference values to tune the procedure parameters by workers. However, there are still three main problems that need further consideration. First, it lacks an integrated scheme from the optimization of control target to the optimization of the production progress. Meanwhile, repetitive changes of operation modes always lead to abrupt deterioration of the control performance and quality of products. Last, through paper [11] optimized the production process using online ADP method, the performance index is not ideal and needs further decreasing.

To solve these problems, an integrated ''target optimization'' and ''process optimization'' scheme is developed. The repetitive operation modes mechanism is studied by an intensive analyze of the production technique. Then, the constrained multi-objective optimization problem is formulated and solved, combing the optimization objective and the curve of solution set, selected solution is considered as the control target. Further, at each time interval, the performance index function is iteratively reduced so that the ADP optimal control method can track the optimized target with less control and state energy consumption. The simulation studies are conducted to show the effectiveness of the proposed method.

The rest of this paper is arranged as follows. Section II describes the optimal control problem of GGBS production process. In Section III, iterative ADP controller is designed based on optimized target. Section IV provides the simulation results and Section V concludes this paper.

II. DESCRIPTION OF GGBS PRODUCTION PROCESS OPTIMAL CONTROL

In this section, the main control variables and corresponding constrains will be analysed. The multi-objective optimization problem and optimal tracking control program are formulated. Further, operation modes changing mechanism is analysed.

A. INTRODUCTION OF GGBS PRODUCTION PROCESS

Workflow of GGBS production is shown in Fig. 1. This production process consists of two parts, under the vertical mill, wasted slag from iron and steel making is ground into particles, the hot gas dries and blows up particles to the top of mill for selection. In the selector, materials are classified that particles with qualified size are transformed into the

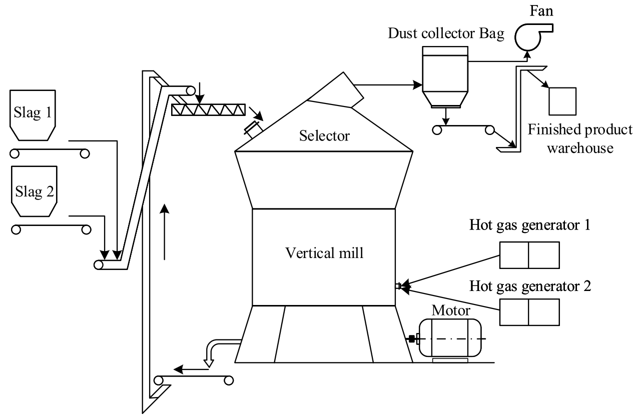

FIGURE 1. Workflow of GGBS production process.

warehouse, and particles with relative big size will drop back onto the millstone for further grinding.

Stability is an essential consideration for GGBS production. Over high mill outlet temperature will burn the filter bag. Dramatically changing mill differential pressure or unappropriate material bed thickness will cause mill vibration. All these factors will lead to production halts or even disasters.

1) MATERIAL BED THICKNESS

Material bed thickness (MBT) is the mean thickness of materials on the grinding bed. MBT is the key factor for a stable GGBS operation. Under the constant grinding pressure, thicker or thinner MBT will increase the rigid contact between grinding roller and grinding bed, leading to strenuous vibration of mill and deterioration of product quality and production effectiveness. In practice, MBT is limited between 5mm and 18mm.

2) MILL DIFFERENTIAL PRESSURE

Mill differential pressure (MDP) is the pressure difference between the pressure at the outlet of mill and pressure at the inlet of mill. MDP reveals the operation state of GGBS production process. In normal cases, MDP is a stable value which means the dynamic balance between the input and output materials. Decreased MDP implies materials input mill are less than that out of mill, and vice versa. In practice, MDP is not only in related with feed material and classifier motor speed, but also hot gas temperature. In engineering, MDP is restricted from 20mbar to 30mbar.

3) MILL OUTLET TEMPERATURE

Mill outlet temperature (MOT) is the temperature at the outlet of selector. Generated from the hot gas dove, hot gas flows through hot gas valve, recirculation gas valve, goes into mill from the bottom, dries the materials and blows them up to the top for selection. Thus, MOT is directly related with the hot gas generator temperature, rotary valve opening of hot gas, and rotary valve opening of recirculation gas. At the same time, it is decided by the materials go in and out of mill, i.e. feed rate and classifier motor speed.

TABLE 1. Admissible area of variables.

Name	Variable	Min	Max	Unit
Feed	u_1	0	115	T/Hr
Motor speed	u_2	900	1250	r/min
Hot gas Temp.	u_3	550	750	°C
Valve opening I	u_4	40	80	%
Valve opening II	u_5	90	100	%
MBT	x_1	5	18	mm
MDP	x_2	20	30	mbar
MOT	x_3	90	120	°C

MOT should be within a certain range. If MOT is too low, GGBS particles can not be fully dried. Otherwise, if MOT is too high, the filter bag is easily to be burnt down. Fitter bag burn is a sever failure which may not only hurt the production yield but also threaten the safety due to the increasing possibility of causing fire disaster. In practical production, MOT should be limited from 90°C to 120°C. Meanwhile, MOT should be as close to the lower limit as possible.

4) SWITCHING OPERATION MODES

In the GGBS production process of Luxin mill line 3 as shown in Fig. 1, except for normal operation mode, hot gas generators will alternate for maintenance, and the feeding materials will change due to different slag sources. These kinds of operation modes take place at 8:00 am on an unknown day. Every time the operation mode changes, mismatch between control signal and controlled plant will cause the abrupt change of states and deteriorate the control performance. For the repetitive changing at fixed time of the day, an effective optimal method is needed to improve the control results.

B. FORMULATION OF GGBS OPTIMAL CONTROL PROBLEM

From above analyse of the focused factors on stability, we extract controlled variables: x_1 is material bed thickness, x_2 is mill differential pressure, x_3 is mill outlet temperature. Control variables: u_1 is feed rate, u_2 is classifier motor speed, u_3 is the hot gas generator temperature, u_4 is rotary valve opening of recirculation gas, u_5 is rotary valve opening of hot gas. Admissible area of variables are listed as in Table 1.

Meanwhile, two optimization problems are formulated as "target optimization" and "process optimization", target optimization provides the optimal set value for the control process, process optimization realizes the optimization of performance index.

1) OPTIMIZATION PROBLEM OF SET VALUES FOR DIFFERENT OPERATING MODES

The objectives are to drive the MBT as close to $x_1^m = 11$ as possible, drive MOT close to x_3^l from the right side, at the same time, all variables should be within admissible ranges.

$$\min\ |x_1 - x_1^m|, \quad x_3 - x_3^l$$
$$s.t.\ \begin{cases} x_i^l \le x_i \le x_i^u \\ u_j^l \le u_j \le u_j^u \\ (x_1, x_2, x_3) = F\ (u_1, u_2, \cdots, u_5) \end{cases} \tag{1}$$

where $i = 1, 2, 3$, $j = 1, 2, \cdots, 5$, x_i^l and u_j^l are the lower bounds of x_i and u_j, x_i^u and u_j^u are the upper bounds of x_i and u_j, F is the static model.

2) OPTIMIZATION PROBLEM OF THE CONTROL PROCESS

Given desired trajectories x_d, the optimal control objective of GGBS production process is to track the x_d in the optimal way with control input varying in admissible range.

$$x \to x_d$$
$$s.t.\ \begin{cases} x_i^l \le x_i \le x_i^u \\ u_j^l \le u_j \le u_j^u \\ u = \arg_u \min J \\ x(k+1) = f(x(k), u(k)) \end{cases} \tag{2}$$

where J is the performance index function about u and x, f is the dynamic model.

III. ITERATIVE ADP TRACKING CONTROL BASED ON DYNAMIC OPTIMIZATION OF SET VALUES

To optimize the GGBS production process, optimal target is obtained by solving the multi-objective problem. Meanwhile, an iterative ADP method is designed to optimize the control process further. Finally, the adaptive control scheme is constructed to deal with the repetitive changing working conditions.

A. DYNAMIC OPTIMIZATION OF SET VALUES

For the multi-objective problem in the form of (1), the static model F should be established first. In this paper, we adopt the PSO-based LS-SVM algorithm to model the three output variables.

For a given sample set $\{u^j, x^j\}_{j=1}^M$, model output by LS-SVM can be obtained as

$$\hat{x}_i(u) = \sum_{j=1}^{M} \Lambda_j K(u, u^j) + b \tag{3}$$

where $i = 1, 2, 3$, the kernel function is given as the radial basis function

$$K(u, u^j) = \exp(||u - u^j||^2/\sigma^2) \tag{4}$$

Parameters Λ_j and b can be calculated by solving the following optimization problem

$$\min_{\omega,b,\zeta} = \frac{1}{2}||\omega||^2 + \frac{1}{2}\gamma \sum_{j=1}^{M} \zeta_j^2$$
$$\text{s.t. } x_i^j = \omega^{\mathrm{T}} K(u^j) + b + \zeta_j, \quad j = 1, \cdots, M \tag{5}$$

Combining PSO and SVM, parameter σ and ζ are optimized and corresponding models (3) of MBT, MDP and MOT can be obtained. Detailed PSO-SVM method can be referred from [9]. Evolutionary algorithms are regarded as effective method to solve multiple competitive objectives in (1), NSGA-II algorithm is adopted to obtain the optimal solution set, where no one solution is better alone with both two objectives than other solutions. It is necessary to select one optimal solution as the set value or control target. Detailed selection mechanism is described in the experiment section.

B. ITERATIVE ADAPTIVE DYNAMIC PROGRAMMING DESIGN

To establish the dynamic model in the optimization problem of the control process described as in (2), following recurrent neural network (RNN) is utilized.

$$\hat{x}(k+1) = A\hat{x}(k) + \hat{W}_1^{\mathrm{T}}(k)\,\phi\left(\hat{V}_1(k)x(k)\right) + \hat{W}_2^{\mathrm{T}}(k)\,\varphi\left(\hat{V}_2(k)x(k)\right)u(k) \tag{6}$$

where $A \in \mathbb{R}^{n\times n}$ is the design matrixes, $\hat{W}_1 \in \mathbb{R}^{n\times n}$, $\hat{V}_1 \in \mathbb{R}^{n\times n}$, $\hat{W}_2 \in \mathbb{R}^{n\times n}$ are the estimated weight matrix. Function $\phi(\hat{V}_1(k)x(k)) \in \mathbb{R}^n$ is a vector with the elements increasing monotonically. The matrix function $\varphi(\hat{V}_2x) \in \mathbb{R}^{n\times m}$ is defined as $\varphi(\hat{V}_2x) = (\varphi_1(\hat{V}_{21}x), \cdots, \varphi_n(\hat{V}_{2n}x))^{\mathrm{T}}$, where $\hat{V}_{21} \in \mathbb{R}^{m\times n}$, φ_i is nondecreasing function. We assume $\hat{V}_1$ and $\hat{V}_2$ are given constant matrix, and only the output weights $\hat{W}_1$ and $\hat{W}_1$ are tuned.

Update RNN weights according to the following law

$$\begin{aligned}
\hat{W}_1(k+1) &= \hat{W}_1(k) - \alpha_1\phi(V_1x(k))x_e^{\mathrm{T}}(k+1) - \alpha_1^2\phi(\bar{x}(k))\\
\hat{W}_2(k+1) &= \hat{W}_2(k) - \alpha_2\varphi(V_2x(k))u(k)x_e^{\mathrm{T}}(k+1)\\
x_e &= \hat{x}(k) - x(k)
\end{aligned} \tag{7}$$

After a period of tuning, neural network weights come to be convergent, and the GGBS production process can be described by the following dynamic [12]

$$x(k+1) = Ax(k) + W_1^{\mathrm{T}}\phi(V_1x(k)) + W_2^{\mathrm{T}}\varphi(V_2x(k))\,u(k) \tag{8}$$

Given the dynamically optimized states $x_d(k)$, desired control input can be derived as

$$u_d(k) = B(x_d(k+1) - Ax_d(k) - W_1^{\mathrm{T}}\phi(V_1x_d(k))) \tag{9}$$

where $B = (W_2^{\mathrm{T}}\varphi(V_2x_d(k)))^{\dagger}$ is the pseudo inverse. Define the state error $e(k)$ and control input error $\upsilon(k)$

$$e(k) = x(k) - x_d(k) \tag{10}$$

$$\upsilon(k) = u(k) - u_d(k) \tag{11}$$

One can obtain

$$\begin{aligned}
e_{k+1} &= Ae_k + W_1^{\mathrm{T}}\phi(V_1x_k) - W_1^{\mathrm{T}}\phi(V_1x_{dk})\\
&\quad + W_2^{\mathrm{T}}\varphi(V_2x_k)u_k - W_2^{\mathrm{T}}\varphi(V_2x_{dk})u_{dk}\\
&= Ae_k + W_1^{\mathrm{T}}\phi(V_1(e_k + x_{dk})) + W_2^{\mathrm{T}}\varphi(V_2(e_k + x_{dk}))u_{dk}\\
&\quad - W_1^{\mathrm{T}}\phi(V_1x_{dk}) - W_2^{\mathrm{T}}\varphi(V_2x_{dk})u_{dk}\\
&\quad + W_2^{\mathrm{T}}\varphi(V_2(e_k + x_{dk}))\upsilon_k
\end{aligned} \tag{12}$$

which can be rewritten as the following dynamic about state error and control error

$$e_{k+1} = f(e_k) + g(e_k)\upsilon_k \tag{13}$$

where $f(e_k) = Ae_k + W_1^{\mathrm{T}}\phi(V_1(e_k + x_{dk})) + W_2^{\mathrm{T}}\varphi(V_2(e_k + x_{dk}))u_{dk} - W_1^{\mathrm{T}}\phi(V_1x_{dk}) - W_2^{\mathrm{T}}\varphi(V_2x_{dk})u_{dk}$, $g(e_k) = W_2^{\mathrm{T}}\varphi(V_2(e_k + x_{dk}))$.

Define the following performance index function

$$J(e(k), \upsilon(k)) = \sum_{n=k}^{\infty} U(e(n), \upsilon(n)) \tag{14}$$

where $U(e(k), \upsilon(k)) = e(k)^{\mathrm{T}}Qe(k) + \upsilon(k)^{\mathrm{T}}R\upsilon(k)$, Q and R are diagonal positive definite matrix of corresponding dimension. To handle the control constraints, we define the following state feedback control,

$$\upsilon(x) = \bar{U}\tanh(\bar{U}^{-1}\nu(x)) \tag{15}$$

where $\bar{U} = \mathrm{diag}\{\tau_1, \cdots, \tau_i, \cdots, \tau_m\}$

$$\tau_i = \min(u_i^u - \max_k u_{dk}, \min_k u_{dk} - u_i^l)$$

System dynamic (13) can be rewritten as

$$e_{k+1} = f(e_k) + g(e_k)\bar{U}\tanh(\bar{U}^{-1}\nu_k) \tag{16}$$

and the performance index function (14) can be rewritten as

$$J(e(k), \nu(k)) = \sum_{n=k}^{\infty} U(e(n), \nu(n)) \tag{17}$$

where $U(e(k), \nu(k)) = e(k)^{\mathrm{T}}Qe(k) + \tanh^{\mathrm{T}}(\bar{U}^{-1}\nu_k)\bar{U}^{\mathrm{T}} \times R\bar{U}\tanh(\bar{U}^{-1}\nu_k)$. Thus, regulating problem (8) with constrained control input has been turned into the tracking problem (16) with unconstrained control signal ν. The ideal optimal control is defined as

$$\nu^*(k) = \arg\min_{\nu(k)}\{J(e(k), \nu(k))\} \tag{18}$$

According to Bellman optimal principle, optimal performance index function satisfies the following HJB function

$$J^*(e(k)) = \min_{\nu(k)}\{U(e(k), \nu(k)) + J^*(e(k+1))\} \tag{19}$$

Optimal control $\nu^*(k)$ satisfies

$$\nu^*(k) = \arg\min_{\nu(k)}\{U(e(k), \nu(k)) + J^*(e(k+1))\} \tag{20}$$

Let the partial derivative of $J^*(e(k))$ with respect to $\nu(k)$ be zero,

$$\begin{aligned}
\frac{\partial J^*(e_k)}{\partial \nu_k} &= \frac{\partial U(e_k, \nu_k)}{\partial \nu_k} + \frac{\partial e_{k+1}^{\mathrm{T}}}{\partial \nu_k} \times \frac{\partial J^*(e_{k+1})}{\partial e_{k+1}}\\
&= 2\Pi\bar{U}^{\mathrm{T}}R\bar{U}\tanh(\bar{U}^{-1}\nu_k) + \Pi^{\mathrm{T}}\bar{U}g^{\mathrm{T}} \times \frac{\partial J^*(e_{k+1})}{\partial e_{k+1}}\\
&= 0
\end{aligned} \tag{21}$$

where

$$\Pi = \bar{U}^{-1}(I_{m\times m} - \tanh(\bar{U}^{-1}\nu_k)\tanh^{\mathrm{T}}(\bar{U}^{-1}\nu_k))$$

Optimal control signal ν_k^* is obtained as

$$\nu_k^* = -\bar{U}\mathrm{artanh}(\frac{1}{2}(\bar{U}R)^{-1}g^{\mathrm{T}}\frac{\partial J^*(e_{k+1})}{\partial e_{k+1}}) \tag{22}$$

C. ADP PRINCIPE

The ADP scheme is implemented by iterating between the control function

$$\nu_i(k) = \arg\min_{\nu(k)}\{U(e(k), \nu(k)) + J_i(e(k+1))\} \tag{23}$$

and performance index function

$$J_{i+1}(e(k)) = U(e(k), \nu_i(k)) + J_i(e(k+1)) \tag{24}$$

where i indicates the iteration number.

1) CRITIC NETWORK

Critic network is designed to approximate the performance index function,

$$\hat{J}_i(e(k)) = \hat{W}_{ci}^{\mathrm{T}}(k)\phi_c(V_c^{\mathrm{T}}e(k)) \tag{25}$$

For simplification, assume that the weight V_c between the input layer and hidden layer is a constant matrix and the weight $\hat{W}_{ci}$ between the hidden layer and the output layer is tuned online.

Objective function of the critic function is

$$J_{i+1}(e(k)) = U(e(k), \nu_i(k)) + \hat{J}_i(e(k+1)) \tag{26}$$

Error function of critic network is defined as

$$e_{ci}(k) = \hat{J}_i(e(k)) - J_i(e(k)) \tag{27}$$

where $J_i(e(k))$ is calculated by (24). The destination of critic network is to minimize the following function

$$E_{ci}(k) = \frac{1}{2}e_{ci}^{\mathrm{T}}(k)e_{ci}(k) \tag{28}$$

According to the gradient decent method, update law of critic network can be derived as

$$\begin{aligned}\hat{W}_{c_{i+1}} &= \hat{W}_{c_i} - \alpha_c(\frac{\partial E_{ci}(k)}{\partial \hat{W}_{c_i}}) \\ &\quad \hat{W}_{c_i} - \alpha_c\phi_c(V_c^{\mathrm{T}}e(k))e_{ci}(k)\end{aligned} \tag{29}$$

where α_c is the learning rate of critic network.

2) ACTOR NETWORK

To approximate the optimal control law, following actor network is designed

$$\hat{\nu}_i(k) = \hat{W}_{ai}^{\mathrm{T}}(k)\phi_a(V_a^{\mathrm{T}}e(k)) \tag{30}$$

Error function of actor network is defined as

$$\begin{aligned}e_{ai}(k) &= \hat{\nu}_i(e(k)) - \nu_i(k) \\ &= \hat{W}_{ai}^{\mathrm{T}}(k)\phi_a(V_a^{\mathrm{T}}e(k)) \\ &\quad + \bar{U}\operatorname{artanh}(\frac{1}{2}(\bar{U}R)^{-1}g^{\mathrm{T}}\frac{\partial\phi_c^{\mathrm{T}}(V_c^{\mathrm{T}}e_{k+1})}{\partial e_{k+1}}\hat{W}_{ci}(k+1))\end{aligned} \tag{31}$$

where

$$\frac{\partial\phi_c^{\mathrm{T}}(V_c^{\mathrm{T}}e_{k+1})}{\partial e_{k+1}} = V_c(I_{m\times m} - \tanh(V_c^{\mathrm{T}}e_{k+1})\tanh^{\mathrm{T}}(V_c^{\mathrm{T}}e_{k+1}))$$

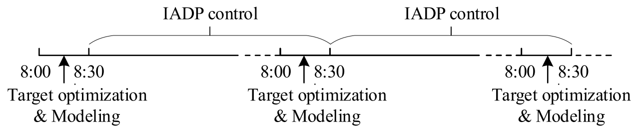

FIGURE 2. Control structure based on target optimization and iterative ADP.

Using the gradient decent method, tuning law of the actor network is obtained as

$$\hat{W}_{a(i+1)}(k) = \hat{W}_{ai}(k) - \frac{\alpha_a\phi_a(V_a^{\mathrm{T}}e(k))e_{ai}^{\mathrm{T}}}{\phi_a^{\mathrm{T}}(V_a^{\mathrm{T}}e(k))\phi_a(V_a^{\mathrm{T}}e(k)) + 1} \tag{32}$$

where α_a is the learning rate of the actor network.

Given initial state error, initial weights of critic and actor network, and maximum iteration times $i_{\max}$, for the ith iteration, flow chart of iterative ADP algorithm is as follows,

1) Calculate the estimated control $\hat{\nu}_i(k)$ according to the actor network (30). Obtain the estimated performance index function $\hat{J}_i(e(k))$ by (25).
2) Substitute $\hat{\nu}_i(k)$ for $\nu_i(k)$ in (16) to obtain the next state error $e(k+1)$.
3) According to critic network I (25), obtain $\hat{J}_i(e(k+1))$ and transfer it to the actor network, update $\hat{W}_{a(i+1)}(k)$ by (32).
4) Update the critic network II's weight $\hat{W}_{c_{i+1}}(k)$ by (29), and assign it to the critic network I (25) to obtain $\hat{J}_{i+1}(e(k))$.

After the ith iteration, $i = i + 1$, go to step 1) if $i \leq i_{\max}$. Otherwise, end the iteration.

D. CONTROL STRUCTURE FOR GGBS PRODUCTION PROCESS

As the operation modes change at 8:00 of some unknown day, once the production process changes, control performance deteriorates due to the mismatch between controlled plant and the controller. In this paper, we adopt the "target optimization & modeling" and "process optimization" scheme as shown in Fig. 2. To decrease the effect of changing modes and improve the control performance, from 8:00 to 8:30 of each day, data of the first thirty minutes are utilized to establish static and dynamic models, and obtain the optimal targets in accordance with current operation mode. Then, based on the targets and dynamic models, iterative ADP controller is applied for 24 hours from 8:30 to next 8:30. It should be noted that it will cause mismatch between plant and controller if modes change at 8:00. However, as the control signal is strictly constrained in given range as (15), above mismatch for thirty minutes is acceptable and will not cause big overshoot.

IV. EXPERIMENTAL RESULTS

A. SIMULATION PLATFORM AND MODELING

As the GGBS production process runs among three typical operation modes, it is necessary to model the three operation

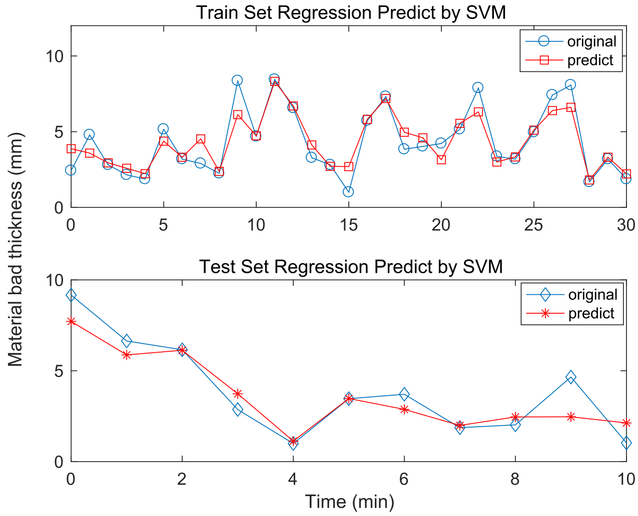

FIGURE 3. Modellng and prediction result by PSO-SVM.

modes and to establish the simulation platform before the proposed scheme is applied into real production process.

According to the experience of engineering, we abstract data from three operation modes, and assume that the modes run and change in the following way

$$\text{Operation mode} = \begin{cases} \text{mode } 1, & t \in [0-90) \\ \text{mode } 2, & t \in [90-180) \\ \text{mode } 3, & t \in [180-270) \end{cases} \quad (33)$$

As explained in Fig. 2, the first 30 minutes are utilized to establish system model. Practical production demands a fast convergence time, thus, to simulate the real system, every half and an hour is regarded as a running period, we assume that system change happens every 90 minutes. As the control input only change within given constraints, mismatch between control input and the real system will not cause the output change abruptly in a short time. For each period, during the first half an hour the system operates with the control signal inherited from the former period, so that enough data can be obtained to model current system. After the modeling performance is satisfied, proposed iterative ADP control method is applied to obtain good control performance.

For each operation modes, three SVM models in respect with the essential state variables MBT, MDP and MOT are established. Modeling and prediction error represented by mean square (MSE) and squared correlation coefficient (SCC) are listed in Table 2. Due to the limit of space, only the modeling and prediction result of MBT in mode 1 is illustrated in Fig. 3. As in above mentioned table and figure, the PSO-SVM method shows satisfying performance to establish the static models of GGBS production process.

Using the RNN method, dynamic models of the three considered state are built for each operation mode. Modeling result of mode 1 is shown in Fig. 4. Given the same design

TABLE 2. Modeling and prediction error by PSO-SVM.

Mode	Model	modeling error		Prediction error	
		MSE	SCC	MSE	SCC
Mode 1	MBT	0.01050	0.80425	0.01394	0.73095
	MDP	0.00068	0.98145	0.00237	0.86290
	MOT	0.00436	0.85145	0.00720	0.79168
Mode 2	MBT	0.00809	0.84792	0.01050	0.77526
	MDP	0.00074	0.97963	0.00284	0.82457
	MOT	0.00586	0.79685	0.00757	0.79115
Mode 3	MBT	0.01076	0.77728	0.01307	0.68844
	MDP	0.00068	0.98145	0.00286	0.80058
	MOT	0.00310	0.86516	0.00634	0.80062

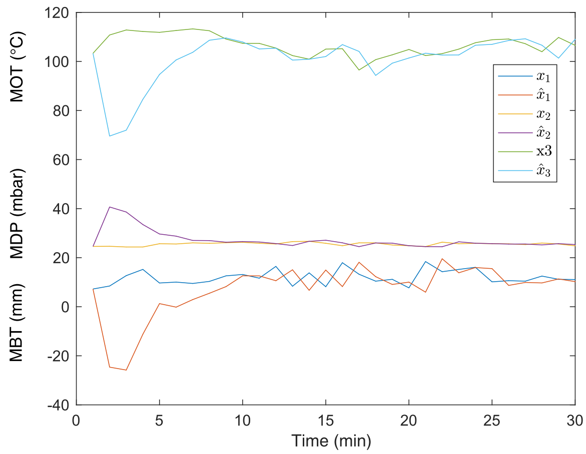

FIGURE 4. Modeling result by RNN.

matrix $A = -\mathrm{diag}([0.8, 0.8])$, constant random V_1 and V_2, RNN models (8) of three operation modes are obtained

$$x_{k+1} = Ax_k + W_{1i}^{\mathrm{T}}\phi\left(V_1 x_k\right) + W_{2i}^{\mathrm{T}}\varphi\left(V_2 x_k\right) u_k \quad (34)$$

where $i = 1, 2, 3$, convergent weights of different models are

$$W_{11} = \begin{bmatrix} -0.2786 & -0.0068 & 0.2843 \\ 0.5547 & 0.2372 & -0.4996 \\ -0.1734 & -0.0721 & 0.3785 \end{bmatrix}$$

$$W_{21} = \begin{bmatrix} 0.0062 & -0.5725 & 0.3540 \\ -0.1324 & 0.8883 & -0.2263 \\ 0.1182 & -0.3240 & -0.1686 \end{bmatrix}$$

$$W_{12} = \begin{bmatrix} -0.2992 & -0.0465 & 0.2427 \\ 0.5341 & 0.1974 & -0.5412 \\ -0.1940 & -0.1118 & 0.3369 \end{bmatrix}$$

$$W_{22} = \begin{bmatrix} 0.0122 & -0.5604 & 0.3663 \\ -0.1263 & 0.9004 & -0.2141 \\ 0.1243 & -0.3118 & -0.1564 \end{bmatrix}$$

$$W_{13} = \begin{bmatrix} -0.2384 & 0.0221 & 0.3332 \\ 0.5949 & 0.2661 & -0.4507 \\ -0.1332 & -0.0432 & 0.4274 \end{bmatrix}$$

$$W_{23} = \begin{bmatrix} -0.0053 & -0.5811 & 0.3410 \\ -0.1439 & 0.8797 & -0.2393 \\ 0.1067 & -0.3325 & -0.1817 \end{bmatrix}$$

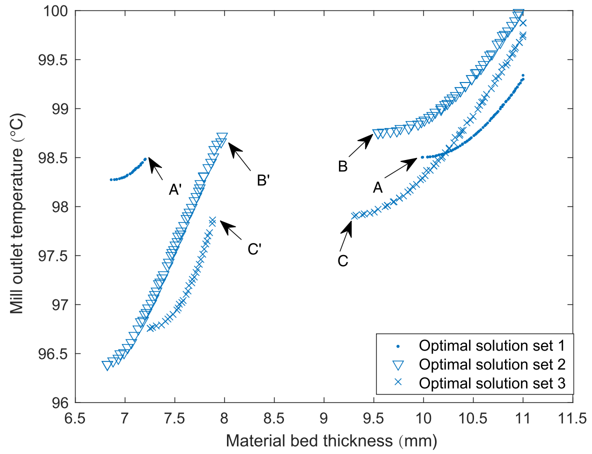

FIGURE 5. Optimal solution set.

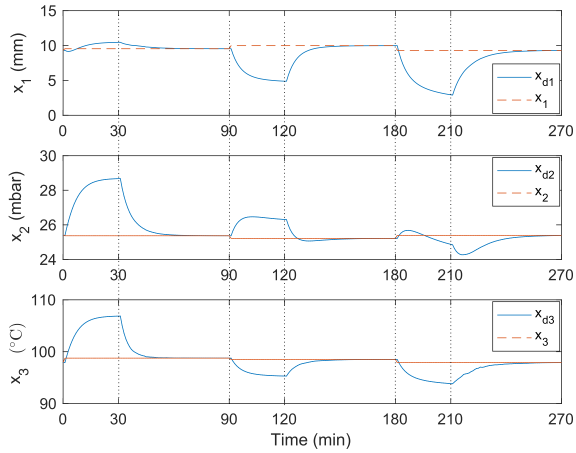

FIGURE 6. The trajectories of x_d and x.

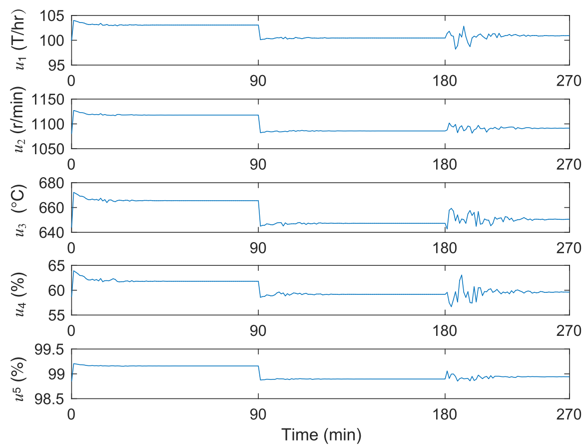

FIGURE 7. Optimal control input.

B. APPLICATION RESULT

Based on the PSO-SVM models, the constrained multi-objective optimization problems (1) of different modes are solved by the NSGA-II method, corresponding Pareto sets are shown in Fig. 5.

Due to control constraints, each optimal solution consists of curves of two parts–the left part and the right part. As the production process demands the material bed thickness to be as close to 11mm as possible and the mill outlet temperature to be as close to 90°C as possible, no point can beat all the other solutions towards both objectives in corresponding optimal set. However, considering the two adjacent points of two parts, for example, A and A′, the temperature increases slightly but the thickness rises greatly from A to A′. That is, great thickness improvement can be achieved with small loss of temperature. By this means, solutions at point A, B and C are selected as optimal solutions of three modes, and the desired set values x_{d1}, x_{d2}, and x_{d3} are obtained as

$$\begin{aligned} x_{d1} &= [9.9863 \quad 25.2169 \quad 98.5031]^T \\ x_{d2} &= [9.3052 \quad 25.3912 \quad 97.9022]^T \\ x_{d3} &= [9.5428 \quad 25.3688 \quad 98.7536]^T \end{aligned} \tag{35}$$

From Fig. 6, it can be seen that due to operation mode change, the output states deteriorate in the first 30 minutes and then track the desired trajectories in the next hour. Fig. 7 shows the control signals which all fall within given constraints as listed in Table 1.

Comparative result of proposed iterative ADP and online ADP applied to GGBS production process in [11] is shown in Table 3. For above three time intervals, maximum performance index calculated by (14), convergence time when the performance index is less than $1e^{-5}$, and simulation time of both methods are listed. It can be seen that by taking the iterative learning scheme, proposed iterative ADP obtains better control performance and less convergence time. However, because of the iterative calculation at each time k, running time increases greatly from about 0.1s to more than 7s. From the respective of engineering, calculating in 7s satisfies the real-time requirement as the system runs for an hour. At the same time, with the rapid progress of computer technology, computation ability will be further improved.

TABLE 3. Compare of online ADP and proposed iterative ADP.

Index	Interval 1	Interval 2	Interval 3
OADP max index	0.3386	0.2275	0.9587
IADP max index	0.2855	0.1088	0.1453
OADP convergence time	48	41	47
IADP convergence time	42	39	42
OADP running time	0.011	0.011	0.013
IADP running time	7.401	7.759	8.601

V. CONCLUSION

In this paper, combing iterative ADP and dynamic target optimization, an effective optimal control scheme for GGBS production process is proposed. Considering the repetitive operation modes change, a repetitive "target optimization" and "process optimization" scheme is constructed. At the beginning of each day, control target is optimized by solving multi-objective problem, during the rest time, an iterative ADP strategy is applied to track the control target in the optimal way. Combing iterative calculation at every interval, the iterative ADP reduces the performance index and makes it converge faster. By using this optimization method, the operational stability of GGBS production is greatly improved. In the future, to further decrease the optimal index and reduce the calculation caused by iterative calculation, appropriate ceasing condition based on the convergent error can be introduced.

REFERENCES

[1] S. Saluja, S. Goyal, and B. Bhattacharjee, ''Strength properties of roller compacted concrete containing GGBS as partial replacement of cement,'' *J. Eng. Res.*, vol. 7, no. 1, pp. 1–17, 2019.

[2] P. Saranya, P. Nagarajan, and A. P. Shashikala, ''Eco-friendly GGBS concrete: A state-of-the-art review,'' in *Proc. IOP Conf. Ser., Mater. Sci. Eng.*, Jun. 2017, pp. 1–5.

[3] Y. Ning, M. Yuan, and Z. Wang, ''Generalized predictive PID controller of differential pressure in vertical roller mill for cement-raw grinding,'' *Inf. Control*, vol. 41, no. 3, pp. 378–390, 2012.

[4] S. Ogonowski, Z. Ogonowski, and M. Pawelczyk, ''Multi-objective and multi-rate control of the grinding and classification circuit with electromagnetic mill,'' *Appl. Sci.*, vol. 8, no. 4, pp. 1–22, 2018.

[5] X. Wang, Z. Liu, Z. Wang, and M. Fan, ''Research and application of intelligent control system for slag grinding,'' *Control Eng.*, vol. 19, no. 2, pp. 240–244, 2012.

[6] M. He, J. Tang, X. Guo, and G. Yan, ''Soft sensor for ball mill load using DAMRRWNN model,'' *Acta Autom. Sinica*, vol. 45, no. 2, pp. 398–406, 2019.

[7] L.-Y. Zhang and T.-Y. Chai, ''Intelligent mill-load control for pulverizing system in alumina sintering process,'' *Control Theory Appl.*, vol. 27, no. 11, pp. 1471–1478, 2010.

[8] X. Lin and W. Kong, ''Adaptive dynamic programming in raw meal fineness control of vertical mill grinding process based on extreme learning machine,'' *J. Syst. Simul.*, vol. 28, no. 11, pp. 2764–2770, 2016.

[9] K. Wang, X. Li, C. Jia, S. Yang, M. Li, and Y. Li, ''Multiobjective optimization of the production process for ground granulated blast furnace slags,'' *Soft Comput.*, vol. 22, no. 24, pp. 8177–8186, 2018.

[10] S. Shen, X. Li, Y. Li, and J. Gao, ''Modeling and analysis of material layer thickness in slag powder process,'' in *Proc. Chin. Control Decis. Conf. (CCDC)*, Shenyang, China, 2018, pp. 4241–4246.

[11] K. Wang, X. Li, C. Jia, and G. Song, ''Optimal tracking control for slag grinding process based on adaptive dynamic programming,'' *Acta Autom. Sinica*, vol. 42, no. 10, pp. 1542–1551, 2016.

[12] X. Zhang, H. Zhang, Q. Sun, and Y. Luo, ''Adaptive dynamic programming-based optimal control of unknown nonaffine nonlinear discrete-time systems with proof of convergence,'' *Neurocomputing*, vol. 91, pp. 48–55, Aug. 2012.

KANG WANG received the B.E. and Ph.D. degrees in control science and engineering from the School of Automation and Electrical Engineering, University of Science and Technology Beijing, in 2012 and 2018, respectively. He is currently a Postdoctoral Researcher with the Faculty of Information Technology, Beijing University of Technology. His research interests include artificial neural networks, optimal control, and intelligent control.

XIAOLI LI received the B.E. and M.E. degrees from the Dalian University of Technology, in 1994 and 1997, respectively, and the Ph.D. degree from Northeastern University, China, in 2000. From 2000 to 2003, he was a Research Fellow of Tsinghua University, China, and the Université Libre de Bruxelles, Belgium. He is currently a Professor with the Faculty of Information Technology, Beijing University of Technology. His research interests include intelligent control, multiple model control, adaptive control, and robust control.