Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

A Dynamic Grey-Box Model and its Application in the Sintering Process of Ternary Cathode Material

Jiayao Chen*, Weihua Gui **, Ning Chen*,Jiayang Dai****, Chunhua Yang *****, Xu Li******

* *School of Automation, Central South University, Changsha, 410083 China* (*E-mail: jychen@csu.edu.cn*).
** *School of Automation, Central South University, Changsha, 410083 China* (*E-mail: gwh@csu.edu.cn*)
*** *School of Automation, Central South University, Changsha, 410083 China* (*E-mail: ningchen@csu.edu.cn*)
**** *School of Automation, Central South University, Changsha, 410083 China* (*E-mail: daijiayang@csu.edu.cn*)
***** *School of Automation, Central South University, Changsha, 410083 China* (*E-mail: ychh@csu.edu.cn*)
******Hunan Shanshan Energy Technology Co. , LTD, Changsha, 410205,.China* (*E-mail: li.bo@shanshanenergy.com*)

Abstract: Soft-sensor technique is often used to estimate key variables in industrial manufacturing, of which the commonly used approaches as the first-principle modeling and data-driven modeling both have their limitations. To take full advantage of the modeling methods and overcome the problems of unmodeled dynamics in industrial manufacturing, a grey-box modeling method combining the first-principle analysis with dynamic data-driven model is developed in this paper. In the framework of the presented grey-box model, the unmodeled dynamics in the first-principle model are obtained by the dynamic probabilistic latent variable model. On this basis, availability of models can be improved. Finally, the actual industrial data in the roller kiln of ternary cathode material manufacturing is used for simulation to verify the validity of the model. The results have practical guiding significance.

Keywords: Grey-Box Model, Dynamic Probabilistic Latent Variable Model, Roller Kiln

1. INTRODUCTION

Soft-sensor techniques have been widely adopted in the process industry through the past decades. Key variables can be estimated by establishing models with available variables. To date, soft-sensor techniques often include three forms, first-principle modeling, data-driven modeling and grey-box modeling. The first-principle modeling method, also called as the white-box modeling, is to establish a mathematical model by analyzing the mechanism of related physical, chemical as well as biological reactions in industrial processes (Ferrer *et al*. (2019)). However, as to systems with complex manufacturing processes, the first-principle model often cannot perform well. Under this circumstance, there are growing appeals for data-driven modeling, as known as the black-box modeling. This method does not require a deep understanding of the mechanism and only rely on the data to build the model. Commonly used methods for data-driven modeling include principal component regression (PCA), partial least squares regression (PLSR) and artificial neural network (Kadlec *et al.* (2010)). However, with the development of technology, the dimensionality and variety of data are increasing, the data-driven model faces the challenge of selecting and applying data.

To handle the imperfection of first-principle model and data-driven model, the grey-box model addressing the concerns. (Ahmad *et at*. (2020)) In order to improve the accuracy and reliability of the model in industrial systems, the grey-box modeling which combines multiple models has received much attention (Yang et al. (2018), Sun et al(2013), Xie et al.(2015), Chen et al.(2017)). Although these models achieve better performance than standalone models, the dynamic characteristics in the first-principle analysis are hardly described. To this end, grey-box model can be adopted in compensating the deficiencies of standalone models. To gain the dynamic characteristics, a dynamic form is introduced into the probabilistic principal component analysis model (Nyamundanda *et al*. (2014)). However, this method is an unsupervised algorithm, thus the established models have weak correlation between input and output, leading to unsatisfactory results sometimes. On this basis, a dynamic probabilistic latent variable model is developed (Ge *et al*. (2019)). This method not only takes dynamic characteristics into consideration but develops a supervised form algorithm.

Therefore, to combine advantages of dynamic form and grey-box model, this paper develops a grey-box model combining first-principle and dynamic probabilistic latent variable model. Furthermore, in order to ensure the same trend of the predicted and the actual value, a trend index is adopted to improve the credibility of the model in parameter identification step.

The remaining parts of this paper are organized as follows. Section 2 provides the framework of the grey-box model and the descriptions of dynamic data-driven model. Section 3 introduces first-principle model sintering roller hearth kiln in the ternary cathode material manufacturing. Then, the first-principle model and the dynamic model are combined to obtain a developed grey-box model. Section 4 presents the results of simulation to verify the feasibility and superiority of the model. Finally, conclusion is delivered in the Section 5.

2. FRAMEWORK OF GREY-BOX MODEL

Although many grey-box models have been proposed, there are situations that have not been taken into consideration such as the case of unmodeled dynamic variables (Ahmad *et at*. (2014)). To this end, a grey-box model combining dynamic data-driven model and first-principle model is developed in the paper. In this grey-box model, the dynamic characteristics of parameters difficult determined in the first-principle model are obtained. In this way, not only the mechanism in the processes are fully used, but also the unmodeled dynamic variable can be obtained. The schematic diagram of grey-box is as shown in Figure 1.

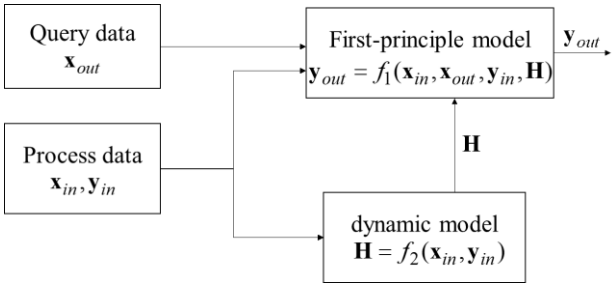

Figure 1 Framework of grey-box model

In order to describe the dynamic relationships among the process variables, a dynamic probabilistic latent variable model is constructed based on the framework of the linear dynamic systems (Ge *et al.* (2019)). Assuming the input and output variables are generated by latent variables $\mathbf{h}_t$. The two major assumptions about the latent variables can be applied in this model. First, the latent variables of all samples are independent and equally distributed. Second, all observations with latent variables are independent of each other. Give the assumption that latent variable $\mathbf{h} \in R^H$ and input variables $\mathbf{x} \in R^V$ can be generated using linearly transformed by latent variables, as shown in Eq. (1).

$$\begin{aligned} \mathbf{h}_t &= \mathbf{A}\mathbf{h}_{t-1} + \boldsymbol{\eta}_t \\ \mathbf{x}_t &= \mathbf{P}\mathbf{h}_t + \mathbf{e}_t \end{aligned} \tag{1}$$

where $\mathbf{A} \in R^{H\times H}$ is the load matrix of latent variable, $\mathbf{P} \in R^{V\times H}$ is the load matrix of the input, $\boldsymbol{\eta}_t \in R^{H\times 1}$ and $\mathbf{e}_t \in R^{V\times 1}$ are noises terms, which obey $\boldsymbol{\eta}_t \sim N(0, \boldsymbol{\Sigma}_\eta)$ and $\mathbf{e}_t \sim N(0, \boldsymbol{\Sigma}_e)$, respectively. H and V are the numbers of latent variables and input variables.

In the similar way, assuming the output variable $\mathbf{y} \in R^L$ can be generated as shown in Eq. (2)

$$\mathbf{y}_t = \mathbf{C}\mathbf{h}_t + \mathbf{f}_t \tag{2}$$

where $\mathbf{C} \in R^{L\times H}$ is the load matrix of the output, L is the number of output, and $\mathbf{f}_t \in R^{L\times 1}$ is the noise term of the output, which obey $\mathbf{f}_t \sim N(0, \boldsymbol{\Sigma}_f)$.

From the above description, although the predictive distribution of the observed variables $\mathbf{y}_t$ only relies on its corresponding latent variable $\mathbf{h}_t$, the latent variable $\mathbf{h}_t$ is assumed to summarize all useful information from all previous results.

According to the description of the dynamic model, the parameter set needs to be determined is $\boldsymbol{\Theta} = \{\boldsymbol{\mu}_\pi, \boldsymbol{\Sigma}_\pi, \mathbf{A}, \boldsymbol{\Sigma}_\eta, \mathbf{P}, \mathbf{C}, \boldsymbol{\Sigma}_e, \boldsymbol{\Sigma}_f\}$, where $\boldsymbol{\mu}_\pi, \boldsymbol{\Sigma}_\pi$ are mean and variance of latent variable $\mathbf{h}_1$. By maximizing the log-likelihood function, this parameter set can be obtained. The log-likelihood function for dynamic model is as Eq. (3)

$$\begin{aligned} L(\theta) &= \ln p(\mathbf{x}_{1:N}, \mathbf{y}_{1:N}, \mathbf{h}_{1:N}) \\ &= \ln \left[p(\mathbf{h}_1) p(\mathbf{x}_1 | \mathbf{h}_1) p(\mathbf{y}_1 | \mathbf{h}_1) \prod_{t=2}^{N} p(\mathbf{h}_t | \mathbf{h}_{t-1}) p(\mathbf{x}_t | \mathbf{h}_t) p(\mathbf{y}_t | \mathbf{h}_t) \right] \\ &= \ln p(\mathbf{h}_1) + \sum_{t=2}^{N} \ln p(\mathbf{h}_t | \mathbf{h}_{t-1}) + \sum_{t=1}^{N} \ln p(\mathbf{x}_t | \mathbf{h}_t) \\ &\quad + \sum_{t=1}^{N} \ln p(\mathbf{y}_t | \mathbf{h}_t) \end{aligned} \tag{3}$$

The Expectation Maximization algorithm (EM algorithm) can be adopted for iteration to obtain the parameters of the model (Dempster *et al.* (1977), Xie *et al.* (2010)).

In the E-step, the three statistics $E(\mathbf{h}_t), E(\mathbf{h}_t \mathbf{h}_t^T), E(\mathbf{h}_t \mathbf{h}_{t+1}^T)$ can be determined by forward and backward algorithm which includes forward filtering step and backward correction step. In the forward filtering step, posterior distribution of the latent variable at time t can be obtained, whose mean and variance are as shown in Eq. (4)

$$\begin{aligned} \mathbf{l}_t &= \boldsymbol{\mu}_h + (\mathbf{P}^T \boldsymbol{\Sigma}_e^{-1} \mathbf{P} + \mathbf{C}^T \boldsymbol{\Sigma}_f^{-1} \mathbf{C} + \boldsymbol{\Sigma}_{hh}^{-1})^{-1} \\ &\quad \times \left[(\boldsymbol{\Sigma}_e^{-1} \mathbf{P})^T (\mathbf{x}_t - \boldsymbol{\mu}_x) + (\boldsymbol{\Sigma}_f^{-1} \mathbf{C})^T (\mathbf{y}_t - \boldsymbol{\mu}_y) \right] \\ \mathbf{L}_t &= (\mathbf{P}^T \boldsymbol{\Sigma}_e^{-1} \mathbf{P} + \mathbf{C}^T \boldsymbol{\Sigma}_f^{-1} \mathbf{C} + \boldsymbol{\Sigma}_{hh}^{-1})^{-1} \end{aligned} \tag{4}$$

The initial posterior distribution can be obtained by Eq. (5)

$$\begin{aligned} \boldsymbol{\mu}_x &= \mathbf{P}\boldsymbol{\mu}_\pi, \boldsymbol{\mu}_y = \mathbf{C}\boldsymbol{\mu}_\pi \\ \mathbf{l}_1 &= \boldsymbol{\mu}_\pi + (\mathbf{P}^T \boldsymbol{\Sigma}_e{}^{-1} \mathbf{P} + \mathbf{C}^T \boldsymbol{\Sigma}_f{}^{-1} \mathbf{C} + \boldsymbol{\Sigma}_\pi{}^{-1})^{-1} \\ &\quad \times \left[\left(\boldsymbol{\Sigma}_e{}^{-1} \mathbf{P} \right)^T (\mathbf{x}_1 - \boldsymbol{\mu}_x) + \left(\boldsymbol{\Sigma}_f{}^{-1} \mathbf{C} \right)^T (\mathbf{y}_t - \boldsymbol{\mu}_y) \right] \\ \mathbf{L}_1 &= (\mathbf{P}^T \boldsymbol{\Sigma}_e{}^{-1} \mathbf{P} + \mathbf{C}^T \boldsymbol{\Sigma}_f{}^{-1} \mathbf{C} + \boldsymbol{\Sigma}_\pi{}^{-1})^{-1} \end{aligned} \tag{5}$$

In the backward correction step, the initial latent variable distribution can be obtained from the forward step $N(\mathbf{m}_N, \mathbf{M}_N) = N(\mathbf{l}_N, \mathbf{L}_N)$. The posterior distribution at time t can be calculated as Eq. (6)

$$\begin{aligned}
\mathbf{m}_t &= \overleftarrow{\mathbf{A}}_t \mathbf{m}_{t+1} + \overleftarrow{\boldsymbol{\mu}}_t \\
\mathbf{M}_t &= \overleftarrow{\mathbf{A}}_t \mathbf{M}_{t+1} \overleftarrow{\mathbf{A}}_t^T + \overleftarrow{\boldsymbol{\Sigma}}_t \\
\overleftarrow{\mathbf{A}}_t &= \mathbf{L}_t \mathbf{A}^T \left(\mathbf{A} \mathbf{L}_t \mathbf{A}^T + \boldsymbol{\Sigma}_\eta \right)^{-1} \\
\overleftarrow{\boldsymbol{\mu}}_t &= \mathbf{l}_t - \mathbf{L}_t \mathbf{A}^T \left(\mathbf{A} \mathbf{L}_t \mathbf{A}^T + \boldsymbol{\Sigma}_\eta \right)^{-1} \mathbf{A} \mathbf{l}_t \\
\overleftarrow{\boldsymbol{\Sigma}}_t &= \mathbf{L}_t - \mathbf{L}_t \mathbf{A}^T \left(\mathbf{A} \mathbf{L}_t \mathbf{A}^T + \boldsymbol{\Sigma}_\eta \right)^{-1} \mathbf{A} \mathbf{L}_t
\end{aligned} \tag{6}$$

Then, $E(\mathbf{h}_t), E(\mathbf{h}_t \mathbf{h}_t^T), E(\mathbf{h}_t \mathbf{h}_{t+1}^T)$ can be calculated as Eq. (7)

$$\begin{aligned}
E\left(\mathbf{h}_t\right) &= \mathbf{m}_t \\
E\left(\mathbf{h}_t \mathbf{h}_t{}^T\right) &= \mathbf{M}_t + \mathbf{m}_t \mathbf{m}_t{}^T \\
E\left(\mathbf{h}_{t+1} \mathbf{h}_t^T\right) &= \mathbf{M}_{t+1} \left[\mathbf{L}_t \mathbf{A}^T \left(\mathbf{A} \mathbf{L}_t \mathbf{A}^T + \boldsymbol{\Sigma}_\eta \right)^{-1} \right]^T + \mathbf{m}_{t+1} \mathbf{m}_t^T
\end{aligned} \tag{7}$$

The parameter update equations can be obtained as

$$\begin{aligned}
\boldsymbol{\mu}_\pi^{new} &= E(\mathbf{h}_1) \\
\boldsymbol{\Sigma}_\pi^{new} &= E(\mathbf{h}_1 \mathbf{h}_1^T) - E(\mathbf{h}_1) E(\mathbf{h}_1^T) \\
\mathbf{A}^{new} &= \sum_{t=2}^{N} E(\mathbf{h}_t \mathbf{h}_{t-1}^T) \left[\sum_{t=2}^{N} E(\mathbf{h}_{t-1}) E(\mathbf{h}_{t-1}^T) \right]^{-1} \\
\mathbf{P}^{new} &= \sum\nolimits_{t=1}^{n} \mathbf{x}_t E\left(\mathbf{h}_t{}^T\right) \left[\sum\nolimits_{t=1}^{n} E\left(\mathbf{h}_t \mathbf{h}_t{}^T\right) \right]^{-1} \\
\mathbf{C}^{new} &= \sum\nolimits_{t=1}^{n} \mathbf{y}_t E\left(\mathbf{h}_t{}^T\right) \left[\sum\nolimits_{t=1}^{n} E\left(\mathbf{h}_t \mathbf{h}_t{}^T\right) \right]^{-1} \\
\boldsymbol{\Sigma}_\eta^{new} &= \frac{1}{N-1} \sum_{t=2}^{N} \begin{bmatrix} E\left(\mathbf{h}_t \mathbf{h}_t{}^T\right) - \mathbf{A}^{new} E\left(\mathbf{h}_{t-1} \mathbf{h}_t^T\right) \\ -E\left(\mathbf{h}_t \mathbf{h}_{t-1}^T\right) (\mathbf{A}^{new})^T \\ +\mathbf{A}^{new} E\left(\mathbf{h}_t \mathbf{h}_t^T\right) (\mathbf{A}^{new})^T \end{bmatrix} \\
\boldsymbol{\Sigma}_e^{new} &= \frac{1}{N} \sum_{t=1}^{N} \begin{bmatrix} \mathbf{x}_t \mathbf{x}_t^T - \mathbf{P}^{new} E(\mathbf{h}_t) \mathbf{x}_t^T - \mathbf{x}_t E(\mathbf{h}_t^T) (\mathbf{P}^{new})^T \\ +\mathbf{P}^{new} E(\mathbf{h}_t \mathbf{h}_t^T) (\mathbf{P}^{new})^T \end{bmatrix} \\
\boldsymbol{\Sigma}_f^{new} &= \frac{1}{N} \sum_{t=1}^{N} \begin{bmatrix} \mathbf{y}_t \mathbf{y}_t^T - \mathbf{C}^{new} E(\mathbf{h}_t) \mathbf{y}_t^T - \mathbf{y}_t E(\mathbf{h}_t^T) (\mathbf{C}^{new})^T \\ +\mathbf{C}^{new} E(\mathbf{h}_t \mathbf{h}_t^T) (\mathbf{C}^{new})^T \end{bmatrix}
\end{aligned} \tag{8}$$

After obtaining the parameters of the model, the output of a series of query samples can be calculated. Assuming the input of query samples are $\mathbf{X}^q = \left[\mathbf{x}_1^q, \mathbf{x}_2^q, \cdots, \mathbf{x}_{k-1}^q, \mathbf{x}_k^q \right]$. The posterior distribution of the first latent variable obeys $N(\mathbf{l}_1, \mathbf{L}_1)$,which can be calculated as Eq. (9)

$$\begin{aligned}
\boldsymbol{\mu}_x &= \mathbf{P} \boldsymbol{\mu}_\pi \\
\mathbf{l}_1 &= \boldsymbol{\mu}_\pi + (\mathbf{P}^T \boldsymbol{\Sigma}_e^{-1} \mathbf{P} + \boldsymbol{\Sigma}_\pi^{-1})^{-1} (\boldsymbol{\Sigma}_e^{-1} \mathbf{P})^{-1} \times (\mathbf{x}_1^{new} - \boldsymbol{\mu}_x) \\
\mathbf{L}_1 &= (\mathbf{P}^T \boldsymbol{\Sigma}_e^{-1} \mathbf{P} + \boldsymbol{\Sigma}_\pi^{-1})^{-1}
\end{aligned} \tag{9}$$

Then, the subsequent posterior distribution of the latent variables at time t obey $N(\mathbf{l}_t, \mathbf{L}_t)$ can be calculated as Eq. (10)

$$\begin{aligned}
\boldsymbol{\mu}_h &= \mathbf{A} \mathbf{l}_{t-1}, \mu_x = \mathbf{P} \boldsymbol{\mu}_h \\
\boldsymbol{\Sigma}_{hh} &= \mathbf{A} \mathbf{L}_{t-1} \mathbf{A}^T + \boldsymbol{\Sigma}_\eta, \boldsymbol{\Sigma}_{xx} = \mathbf{P} \boldsymbol{\Sigma}_{hh} \mathbf{P}^T + \boldsymbol{\Sigma}_e \\
\boldsymbol{\Sigma}_{xh} &= \mathbf{P} \boldsymbol{\Sigma}_{hh} \\
\mathbf{l}_t &= \boldsymbol{\mu}_h + \boldsymbol{\Sigma}_{xh}^T \boldsymbol{\Sigma}_{xx}^{-1} (\mathbf{x}_t^{new} - \boldsymbol{\mu}_x) \\
\mathbf{K} &= \boldsymbol{\Sigma}_{hh} \mathbf{P}^T \boldsymbol{\Sigma}_{xx}^{-1} \\
\mathbf{L}_t &= (\mathbf{I} - \mathbf{K}\mathbf{P}) \boldsymbol{\Sigma}_{hh} (\mathbf{I} - \mathbf{K}\mathbf{P})^T + \mathbf{K} \boldsymbol{\Sigma}_e \mathbf{K}^T
\end{aligned} \tag{10}$$

Finally, the output of the query samples at time t can be calculated as Eq. (11)

$$\mathbf{y}_t^q = \mathbf{C} \mathbf{h}_t^q, t = 1, 2, \cdots, k \tag{11}$$

3. A GREY MODEL FOR SINTERING PROCESS OF TERNARY CATHODE MATRIAL

In this section, the first-principle model of roller kiln will be established. On this basis, the grey-box model will be built.

3.1 Description of roller kiln and the first-principle model

In the manufacturing of lithium batteries, the material that affects products' performance and price most is the cathode material, which determines not only the safety performance but also the prospects of lithium batteries (Li et al.(2014), Liu et al.(2017)). By analyzing the manufacturing of the ternary cathode material, it can be found that the sintering temperature has a great influence on the product quality. In order to ensure product quality, it is necessary to monitor the sintering temperature, which is of great significance for improving product quality and reaction efficiency.

Roller kiln is the main equipment of controlling sintering temperatures. The overall diagram of the roller kiln is shown in Figure 2.

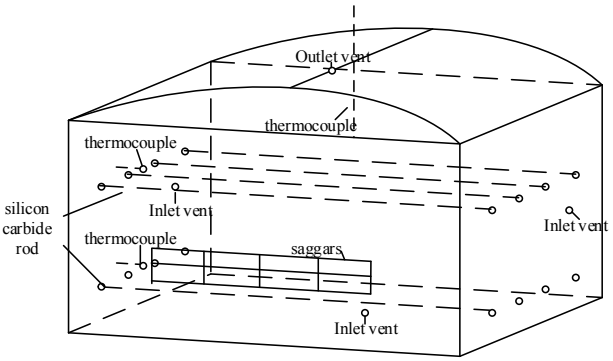

Figure 2 The overall diagram of the roller kiln

It can be found in the Figure 1 that roll kiln is mainly composed of four systems: inlet/outlet air system, transmission system, temperature control system and monitoring system. The roller kiln has a total length of 40 meters and is divided into 21 temperature zones. The temperature zones of 1-9 is the heating section, the temperature zones of 10-18 is the constant temperature section, and the temperature zones of 19-21 is the cooling section. Moreover, the 21 temperature zones are divided into upper temperature zones and the lower temperature zones.

The roller kiln is a typical system in which the relevant variables in the first-principle model cannot track the dynamic characteristics. In the roller kiln temperature prediction model, the predicted temperature is related to

many factors, such as the temperatures of the adjacent temperature zones, the temperature of the kiln wall, the heating current, and the temperature at the inlet and outlet of the atmosphere. Among them, the atmosphere outlet temperature as a related variable is difficult to obtain by using the first-principle, and the measurement interval of the data is large, so that some of the output data is missing. The data characteristics of the atmosphere outlet temperature are just suitable for the industrial characteristics in this paper. Therefore, the grey-box model of the first-principle analysis and data-driven model of the roller kiln will be established with the outlet temperature of the atmosphere as the output of the data-driven model. A first-principle model of *i*-th sub-temperature zone can be established according to the thermodynamic principle (Dai et al. (2019))

$$\begin{cases} \dfrac{dT_{i1}(t)}{dt} = a_{i1}[\int_{t-\Delta t}^{t} I_{i1}^2(t)R_{i1}dt + \omega_{i1}^a(T_{(i+1),1}(t) - T_{i1}(t)) \\ \quad +V_{i1}^a S_{i1}^a(T_{i2}(t) - T_{i1}(t)) + C_{i1}m_{i1}^s(T_{(i-1),1}(t) - T_{i1}(t)) \\ \quad -2490m_{i1}^w - V_{i1}^b S_{i1}^b(T_{i1}(t) - T_{q1}) - \omega_{i1}^b(T_{i1}(t) - T_{(i-1),1}(t)) \\ \quad -\omega_{i1}^c(T_{i1}(t) - T_{i2}(t)) - \omega_{i1}^d(T_w - T_f)] \\ \dfrac{dT_{i2}(t)}{dt} = a_{i2}[\int_{t-\Delta t}^{t} I_{i2}^2(t)R_{i2}dt + \omega_{i2}^a(T_{(i+1),2}(t) - T_{i2}(t)) \\ \quad +V_{i2}^a S_{i2}^a(T_{q0} - T_{i2}(t)) + C_{i2}m_{i2}^s(T_{(i-1),2}(t) - T_{i2}(t)) \\ \quad -2490m_{i2}^w - V_{i1}^a S_{i1}^a(T_{i2}(t) - T_{i1}(t)) - \omega_{i2}^b(T_{i2}(t) \\ \quad -T_{(i-1),2}(t)) - \omega_{i2}^c(T_{i1}(t) - T_{i2}(t)) - \omega_{i2}^d(T_w - T_f)] \end{cases} \tag{12}$$

where a_{i1} and a_{i2} are the heat conversion coefficient, i_1, i_2 are the upper and lower sub-temperature zone, respectively, I_i is the heating current of the *i*-th sub-temperature zone, R_i is the heating resistance of the *i*-th sub-temperature zone, ω_i^a, ω_i^b, ω_i^c, ω_i^d are the conversion coefficients of heat and temperature of *i*-th the sub-temperature zone, C_i is the heat capacities of saggars and material for the *i*-th sub-temperature zone, V_i^a, V_i^b are the specific convective heat transfer coefficients of the *i*-th sub-temperature zone, S_i^a, S_i^b specific heat convection areas of the *i*-th sub-temperature zone, T_i is the temperature of *i*-th sub-temperature zone, m_i^s is the mass of the inlet charging saggars of the sub-temperature zone, m_i^w is the mass of water evaporated in the sub-temperature zone.

By some mathematical operation, the discrete form of first-principle model for the *i*-th temperature zone can be written as Eq. (13)

$$\begin{cases} x_{i1}(k+1) = a'_{i1}u_{i1}(k) + b'_{i1}x_{i1}(k) + c'_{i1}x_{i2}(k) \\ \quad + d'_{i1}x_{(i-1),1}(k) + e'_{i1}x_{(i+1),1}(k) \\ \quad + f'_{i1}T_{q1}(k) + g'_{i1} \\ x_{i2}(k+1) = a'_{i2}u_{i2}(k) + b'_{i2}x_{i1}(k) + c'_{i2}x_{i2}(k) \\ \quad + d'_{i2}x_{(i-1),2}(k) + e'_{i2}x_{(i+1),2}(k) \\ \quad + f'_{i2}u'_i(k) + g'_{i2} \end{cases} \tag{13}$$

where a'_{i1}, a'_{i2}, b'_{i1}, b'_{i2}, c'_{i1}, c'_{i2}, d'_{i1}, d'_{i2}, e'_{i1}, e'_{i2} f'_{i1}, f'_{i2}, g'_{i1} g'_{i2} are the influence coefficient. x_{i1}, x_{i2} is the temperature of temperature zone i_1 and i_2, $x_{(i-1),1}$, $x_{(i-1),2}$, $x_{(i+1),1}$, $x_{(i+1),2}$ is the temperature of previous and subsequent temperature zone of the *i*-th upper and lower temperature zone, u_{i1} ,u_{i2} is the square of the heating current in the temperature zones i_1 and i_2, u'_i is inlet flow of atmosphere in the *i*-th temperature zone. T_{q1} is the temperature of atmospheric outlet which cannot be measured in time and can be predicted by data-driven model.

3.2 Model integration and parameter optimization

By analysing the relevant factors affecting the outlet temperature of the atmosphere, the input variables of the atmosphere outlet temperature data-driven model are selected as the upper and lower temperature zones currents, the adjacent and present temperature zones. Based on the dynamic data-driven model of the atmosphere outlet temperature, the predicted output is taken as the dynamic variable in the first-principle model. The grey-box model can be written as:

$$\begin{cases} x_{i1}(k+1) = a'_{i1}u_{i1}(k) + b'_{i1}x_{i1}(k) + c'_{i1}x_{i2}(k) \\ \quad + d'_{i1}x_{(i-1),1}(k) + e'_{i1}x_{(i+1),1}(k) \\ \quad + f'_{i1}H_i(\mathbf{x}_k) + g'_{i1} \\ x_{i2}(k+1) = a'_{i2}u_{i2}(k) + b'_{i2}x_{i1}(k) + c'_{i2}x_{i2}(k) \\ \quad + d'_{i2}x_{(i-1),2}(k) + e'_{i2}x_{(i+1),2}(k) \\ \quad + f'_{i2}u'_i(k) + g'_{i2} \\ H_i(\mathbf{x}_k) = T_{q1} = f_i(\mathbf{x}_k) \end{cases} \tag{14}$$

where $\mathbf{x}_k$ is the input samples. Then identify the parameters in the grey-box model, define the parameter vector to be identified as $\varsigma = [a'_{i1}, a'_{i2}, b'_{i1}, b'_{i2}, c'_{i1}, c'_{i2}, d'_{i1}, d'_{i2}, e'_{i1}, e'_{i2}, f'_{i1}, f'_{i2}, g'_{i1}, g'_{i2}]$. The traditional parameters identify methods computationally time-consuming and the range of parameters are limited. Therefore, they are often not optimal parameters, nor can they optimize mean square error, mean absolute value error or maximum absolute error. In this paper, Particle swarm optimization is used to identify model parameters.

In addition, the real industrial manufacturing process can be reflected when the trend of the model predicts the same trend as the actual value. Therefore, in order to gain overall optimal parameters and ensure the same trend of the predicted and actual value, the trend consistency index is adopted to the optimization target. Then, converting the multi-objective optimization problem into a single-objective optimization problem, the fitness function can be obtained as shown in Eq. (15).

$$\begin{aligned} \min J(\xi) = \frac{1}{2}\sum_{j=1}^{N}(e_{j1}^2 + e_{j2}^2) \\ -\alpha\sum_{j=2}^{N}\mathrm{sgn}\left(\frac{y_{j,1} - y_{j-1,1}}{\bar{x}_j - \bar{x}_{j-1}}\right)\mathrm{sgn}\left(\frac{\hat{y}_{j,1} - \hat{y}_{j-1,1}}{\bar{x}_j - \bar{x}_{j-1}}\right) \\ -\beta\sum_{j=2}^{N}\mathrm{sgn}\left(\frac{y_{j,2} - y_{j-1,2}}{\bar{x}_j - \bar{x}_{j-1}}\right)\mathrm{sgn}\left(\frac{\hat{y}_{j,2} - \hat{y}_{j-1,2}}{\bar{x}_j - \bar{x}_{j-1}}\right) \end{aligned} \tag{15}$$

where $e_{j1} = y_{j1} - \hat{y}_{j1}$, $e_{j2} = y_{j2} - \hat{y}_{j2}$. y_{j1}, y_{j2} are the actual temperature of *j*-th sample upper and lower

temperature zones, $\hat{y}_{j1}$, $\hat{y}_{j2}$ are the predicted temperature of j-th sample upper and lower temperature zones, $y_{j-1,1}$, $y_{j-1,2}$ are the actual temperature of (j-1)-th sample upper and lower temperature zones, $\hat{y}_{j-1,1}$, $\hat{y}_{j-1,2}$ are the predicted temperature of (j-1)-th sample upper and lower temperature zones, $\bar{x}_j$, $\bar{x}_{j-1}$ are the mean of the input sample of j-th and (j-1)-th sample, respectively, α and β are constant coefficient, N is the total number of training samples. The parameter identification steps are as follows:

Step1: Initialization: set the data-driven model parameters, the particle group parameter range, learning factor and so on;

Step2: Establish a data-driven model of the atmosphere outlet temperature and bring it into the established first-principle model;

Step3: Individual evaluation: taking the initial position of each particle as the individual extreme value, calculating the initial fitness J of each particle in the group according to Eq. (15), and finding the optimal position of the population;

Step4: Update the velocity and position of the particle, generate a new population. Calculate the current fitness value of the particle, and its own historical optimal value p_{Ns}. If it is better than p_{Ns}, replace p_{Ns} with the current fitness value and update the particle position; the optimal population value is *BestS*, if it is better than *BestS*, *BestS* replaces with the current fitness value and updates the global optimal value of the population;

Step5: Check the end condition, if it is satisfied, end the optimization; otherwise go to Step3. Repeat until the optimization reaches the maximum number of iteration steps or the evaluation value is less than the given precision.

4.VERIFICATION OF THE GREY-BOX MODEL

Taking a ternary cathode material manufacture plant in china as an example, the data of the industrial manufacturing roller kiln in October 2016 is collected to verify the developed model. The standard errors RMSE and R^2 are used as performance indices when evaluating the predictive performance of the model.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{16}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{17}$$

The third temperature zone in the heating section is regarded as the main objective in the simulation. After using the Pauta criteria to eliminate the abnormal data, 1458 sets of data are obtained. For verifying the feasibility of the established model, the 1458 sets of data are denoted as data set W1, 30% of which are taken as training samples in data-driven model, 30% as parameters identification training samples, and the remaining 40% are randomly selected 150 samples as the test samples.

4.1 Performance comparison of different parameter

The related parameter in the model need to be determined first. In this paper, the latent variable dimension H need to be determined and varied from 2 to 6, the prediction results are shown as Table 1.

Table 1 the prediction error of different zone

H	2	3	4	5	6
upper	3.9878	3.0412	2.6995	3.1109	2.9358
lower	4.0218	2.5430	2.3368	2.0099	3.0711

It can be seen from Table 1, the results show that when the latent variable dimension H is 4, better prediction results can be obtained. Therefore, the latent variable dimensions H are set to 4 in the subsequent simulation.

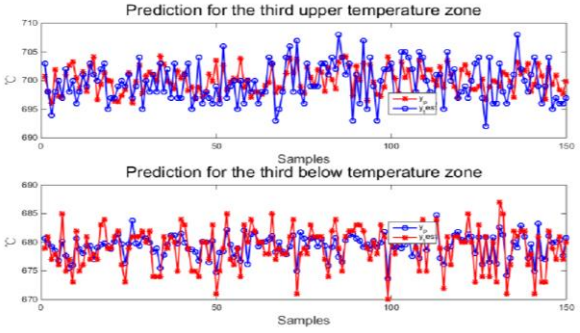

Figure 3 The predicted results with H=4

4.2 Performance comparison of different models

To verify the superiority of the developed model, the mechanism model and the improved hybrid model are compared, which are denoted as Model 1 and Model 2, respectively. In addition, the models are established using the same datasets. The predicted results are shown in Table 2.

Table 2 the prediction error of different zone

Temperature zone		upper	lower
Model 1	RMSE	5.3008	2.8503
	R^2	-0.4910	0.2538
Model 2	RMSE	2.6995	2.3368
	R^2	0.2159	0.3059

It can be seen from Table 2, the grey-box model developed in this paper has a good effect in both the upper temperature zone and the lower temperature zone. This result shows that the grey-box model is able to gain the dynamic characteristic of variable in the mechanism model, thus have better prediction results than mechanism model. The predicted results are shown in Figure 4.

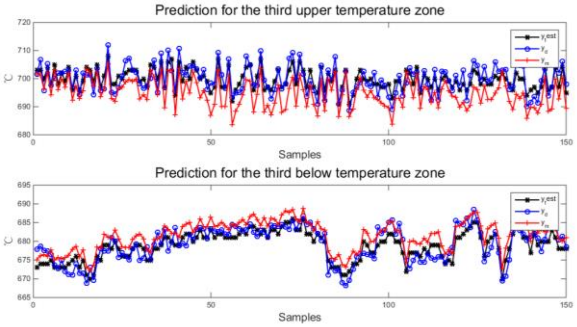

Figure 4 The predicted results with grey-box model

5. CONCLUSIONS

In this paper, a grey-box modelling method combining first-principle analysis and dynamic data-driven model is developed. Variables in the first-principle model whose dynamic characteristics hardly obtained were measured

using a data-driven model. In the data-driven model, a dynamic probabilistic latent variable was adopted to match the characteristics of data in industrial manufacturing. Finally, the temperature in roller kiln was estimated by the grey-box model. The results showed that the developed model has good performance.

REFERENCES

Ahmad, I., Ayub, A., Kano, M., & Cheema, I. I. (2020). Gray-box Soft Sensors in Process Industry: Current Practice, and Future Prospects in Era of Big Data. *Processes*, 8(2), 243.

Ahmad, I., Kano, M., Hasebe, S., Kitada, H., & Murata, N. (2014). Gray-box modeling for prediction and control of molten steel temperature in tundish. J*ournal of Process Control*, 24(4), 375-382.

Chen, N., Luo, L., Gui, W., & Guo, Y. (2017). Integrated modeling for roller kiln temperature prediction. In *2017 Chinese Automation Congress (CAC)* (pp. 3064-3069). IEEE.

Dai, J., Chen, N., Yuan, X., Gui, W., & Luo, L. (2019). Temperature prediction for roller kiln based on hybrid first-principle model and data-driven MW-DLWKPCR model. *ISA transactions*.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1-22.

Ferrer, S., Mezquita, A., Aguilella, V. M., & Monfort, E. (2019). Beyond the energy balance: Exergy analysis of an industrial roller kiln firing porcelain tiles. *Applied Thermal Engineering*, *150*, 1002-1015.

Ge, Z., & Chen, X. (2017). Dynamic probabilistic latent variable model for process data modeling and regression application. *IEEE Transactions on Control Systems Technology*, 27(1), 323-331.

Kadlec, P., Grbić, R., & Gabrys, B. (2011). Review of adaptation mechanisms for data-driven soft sensors. *Computers & chemical engineering*, *35*(1), 1-24.

Li, X. Q., Xiong, X. H., Wang, Z. X., & Chen, Q. Y. (2014). Effect of sintering temperature on cycling performance and rate performance of LiNi0. 8Co0. 1Mn0. 1O2. *Transactions of Nonferrous Metals Society of China*, *24*(12), 4023-4029.

Liu L. (2017) LiNi0.5 Co0.2 Mn0.3 O2 ternary anode material surface residual lithium and pH control research. *Journal of Functional Materials*. 48(9):09175-09179.

Nyamundanda, G., Gormley, I. C., & Brennan, L. (2014). A dynamic probabilistic principal components model for the analysis of longitudinal metabolomics data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 63(5), 763-782.

Sun, B., Gui, W. H., Wu, T. B., Wang, Y. L., & Yang, C. H. (2013). An integrated prediction model of cobalt ion concentration based on oxidation–reduction potential. *Hydrometallurgy*, 140, 102-110.

Xie, F. C., Lin, J. G., & Wei, B. C. (2010). Influence analysis of additive mixed-effects nonlinear regression models via EM algorithm. *Journal of Statistical Computation and Simulation*, *80*(10), 1115-1129.

Xie, Y., Xie, S., Chen, X., Gui, W., Yang, C., & Caccetta, L. (2015). An integrated predictive model with an on-line updating strategy for iron precipitation in zinc hydrometallurgy. *Hydrometallurgy*, *151*, 62-72.

Yang, J., Chai, T., Luo, C., & Yu, W. (2018). Intelligent Demand Forecasting of Smelting Process Using Data-Driven and Mechanism Model. *IEEE Transactions on Industrial Electronics*, 66(12), 9745-9755.