



OPEN

## Digital twin key technology on rare earth process

Hui Yang<sup>1,2,5</sup>, Zhiqin Kuang<sup>1,2,5</sup>, Jianyong Zhu<sup>1,2</sup>✉, Fangping Xu<sup>1,2,5</sup>, Feiping Nie<sup>3,5</sup> & Shuchen Sun<sup>4,5</sup>

Digital twin can be defined as a digital equivalent of an object of which it can mirror its behavior and status or virtual replicas of real physical entities in Cyberspace. To an extent, it also can simulate and predict the states of equipment or systems through smart algorithms and massive data. Hence, the digital twin is emerging used in intelligent manufacturing Systems in real-time and predicting system failure and also has introduced into a variety of traditional industries such as construction, Agriculture. Rare earth production is a typical process industry, and its Extraction Process enjoys the top priority in the industry. However, the extraction process is usually characterized by nonlinear behavior, large time delays, and strong coupling of various process variables. In case of failures happened in the process, the whole line would be shut down. Therefore, the digital twin is introduced into the design of process simulation to promote the efficiency and intelligent level of the Extraction Process. This paper proposes the techniques to build the rare earth digital twin such as soft measurement of component content, component content process simulation, control optimization strategy, and virtual workshop, etc. At the end, the validity of the model is verified, and a case study is conducted to verify the feasibility of the whole Digital twin framework.

Currently, rare studies deal with the simulation and prediction of the complete manufacturing process. Even little study has been done on process optimization based on process simulation data. Rare Earths (RE), sometimes known as industrial vitamins, are widely applied in a variety of areas, including metallurgy, high-speed trains, and the national defense military industry. The Cascade Extraction Process, which is a cascade of several extraction tanks, is frequently used to increase the purity of rare earth in Chinese RE facilities. Furthermore, the Rare Earth Extraction Process (REEP), as a typical complex industrial process, is characterized by nonlinear behavior, long-time delays, and significant coupling of numerous process variables. In this study, we propose a DT system for REEP that incorporates inspection, virtual inspection, control, and process simulation modules. Cascade extraction is separated into thousands of extraction sub-processes, which is the principal reason of the lengthy process and over-reliance on human operation for failure judgment relating to the factory's normal operation.

With the advancement of information technology, such as the industrial internet, the internet of things, and artificial intelligence, the concept of smart manufacturing is increasingly becoming such a reality. Germany, a powerful manufacturing country, suggested the 'industry 4.0' concept<sup>1</sup>. One of the key challenges for smart manufacturing is to connect the physical and virtual spaces. It is necessary to set up Cyber-physical systems<sup>2</sup>, even Cyber-physical production systems<sup>3</sup>. Currently, The importance of digital twin (DT) is increasingly recognized by both academia and industry. many DT applications have been successfully implemented in equipment maintenance, production health management, process simulation, auxiliary design, fault diagnostics, and so on<sup>4</sup>.

The DT may be traced back to a product life-cycle management program lectured by Dr. Michael Grieves, who proposed an unique notion called "Information Mirroring Model", which was then refined to "Digital Twin"<sup>5</sup>. Yang covered essential DT technologies such as data capture, transmission, and processing, data-driven and model fusion collaborative control, virtual interaction and collaboration, and its applications in smart manufacturing, product development, and smart cities, among others<sup>6</sup>.

This paper proposes four key methods to build a rare earth digital twin system to achieve the functions of real-time sensing, rapid prediction, control optimization and efficient inspection of the digital twin system: a soft measurement method of solution component based on the color characteristics of rare earth solution is proposed to obtain the solution component content in the rare earth extraction process in real time, a component content prediction method based on mechanism compensation is proposed to simulate the component content change

<sup>1</sup>School of Electrical and Electronic Engineering, East China Jiaotong University, Nanchang 330013, China. <sup>2</sup>Key Laboratory of Advanced Control and Optimization of Jiangxi Province, Nanchang 330013, China. <sup>3</sup>School of Computer, Northwestern Polytechnical University, Xi'an 710072, China. <sup>4</sup>School of Materials & Metallurgy, Northeastern University, Shenyang 110004, China. <sup>5</sup>These authors contributed equally: Hui Yang, Zhiqin Kuang, Fangping Xu, Feiping Nie and Shuchen Sun. ✉email: zhujiemail@163.com

of the extraction solution. In order to improve the control effect, a control optimization method based on case search is proposed, and a virtual plant using virtual reality technology is proposed to facilitate quick understanding of the working conditions and equipment status.

The paper's overall structure is as follows: A literature review on digital twin shop intelligent production lines technology and Digital twin technology for the process industry are provided in “[Related works](#)” section “[rare earth production DT framework](#)” section describes the architecture of a rare earth digital twin system. “[Key technologies](#)” section introduces and discusses four key enabling technologies, including soft measurement of rare earth component content, as well as the modeling approach of process simulation and the principle of parameter optimization “[Model validation and case study](#)” section presents illustrative applications and discussions to verify the proposed model. Finally, conclusions are presented in “[Conclusions](#)” section.

## Related works

Previously, More applications for digital twins can be found in the fields of equipment maintenance and flaw diagnosis. The Airframe Digital Twin (ADT), which is envisioned to be an ultra-realistic, cradle-to-grave computer model of an aircraft structure, was built by Gockel to analyze the aircraft's capacity to satisfy mission criteria<sup>7</sup>. Bielefeldt then ingeniously constructed a DT of an aircraft wing subjected to flight loading and replicates the behavior of these localized particles while simultaneously lowering computing time<sup>8</sup>. Wang introduced a data-driven DT of the rotor system for accurate diagnosis and adaptive deterioration analysis of rotating equipment since this rotor system emulates an imbalance fault and its advancement, unbalanced quantification<sup>9</sup>. Jiang presented a single description and standardization of DT, which is used in a 110 kV substation's designed prognosis and health management (PHM) system<sup>10</sup>. Li employed the notion of a dynamic Bayesian network to construct an adaptable probabilistic model for diagnosis and prognosis, and it showed the Feasible technique using an aircraft wing fatigue crack growth example<sup>11</sup>.

With advancements in virtualization, sensor technologies, and computing capacity, the notion of DT has grown to design, create, and operate complex systems by providing a secure virtual area for testing and validation<sup>12</sup>. Virtual reality (VR) technology captures human movements in a virtual environment, which are subsequently replicated by a real robot<sup>13</sup>. Bilderberg and Malik discuss an object-oriented and event-driven simulation of a flexible assembly cell coordinated with a robot to execute assembly duties alongside humans<sup>14</sup>. Corral-Acero created a visual DT for cardiology diagnosis and uses computer models and artificial intelligence to properly forecast cardiology. They give remedies that are personalized to each patient and optimize the healthcare system's efficacy and efficiency<sup>15</sup>. Kang used DT technology to create a multimedia knowledge-based bridge health monitoring system that evaluates aberrant bridge circumstances and suggests the best time for maintenance<sup>16</sup>.

Scholars mainly work on the application of the digital twin in the Manufacturing industry. The implement of DT in manufacturing can optimize production processes and reduce energy consumption and costs. Tao conducted a framework to construct a digital twin system that included a real workshop, a virtual workshop, a database, and a service system<sup>17</sup>. Stark focused on the theoretical underpinning for the DT design framework and offered an 8-dimension Digital Twin model. The DT design components and their effect factors on the DT 8-dimension model were used to the smart manufacturing cell's development and ramp-up operations<sup>18</sup>. Jones summarized important vocabulary, and associated procedures, and lists 13 major features such as physical Entity, virtual Entity, and so on<sup>19</sup>. The use of simulation techniques brings digital twins to life and allows them to be tested. Schluse pioneered the notion of EDTs (experimental digital twins) as a new structuring element for simulation-based systems engineering processes and their interdisciplinary and cross-domain simulation, which enables comprehensive simulations on the system level<sup>20</sup>. A DT case for a welding manufacturing line is created using an application framework and a virtual model<sup>21</sup>. Wu introduced four important technologies for achieving real-time monitoring based on DT, including data modeling and transmission, event-driven virtual and real mapping, workshop logic modeling, and information visualization<sup>22</sup>. Considering the influence of carbon emissions from manufacturing processes, Zhao presented a method for dynamically optimizing machining process parameters to decrease carbon emissions based on real-time observation of the machining circumstances<sup>23</sup>. Zhang presented a special framework for simulation optimization utilizing a virtual workshop in response to several issues encountered during the design stage of a dual-manipulator cooperative unit<sup>24</sup>.

The current state of digital twin technology is mostly employed in the discrete industry. Due to the complexities of process modeling, the process industry, is still missing in DT trials. Among the six essential technologies advocated by Li for process industry DT are data intelligent perception, multisource heterogeneous data integration, data-efficient transmission, digital twin creation, enhanced interaction, and transformation application<sup>25</sup>. Zhou streamlined the raw materials delivery schedule at an ironmaking factory with five sintering machines and seven blast furnaces utilizing cloud computing and DT technology to cut production expenses to reduce the mean coke ratio of an ironmaking plant<sup>26</sup>. Soares showed how a digital twin was successfully implemented to simulate a four-stage multi-effect evaporation train from an industrial sugar-cane processing facility<sup>27</sup>. There is presently no reference model especially established for risk control and prevention in the oil and gas industries. Bevilacqua created a Digital Twin reference model to offer conceptual criteria for DT deployment for risk prediction and prevention<sup>28</sup>.

## Rare earth production DT framework

Figure 1 depicts the REEP (rare earth extraction process), which includes a dissolve circuit, extraction circuit, disposal circuit, and dehydration circuit. In an acid solution, the feed liquid and raw material powder are dissolved. In the dissolution process, the raw material powder is mixed with acid and water, and then neutralized with alkali to a certain pH value, and then precipitated and filtered to obtain the next stage of the raw material. In the extraction process, the solution configured in the previous stage is separated from a variety of separation

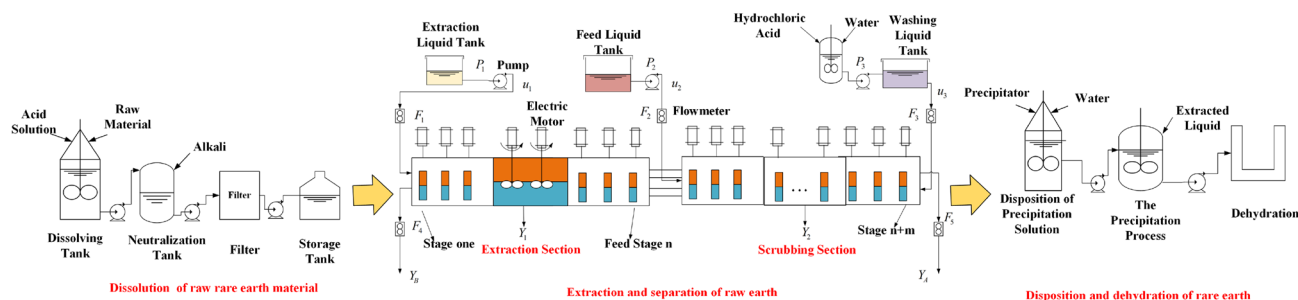


Figure 1. The Main procedure of rare earth production.

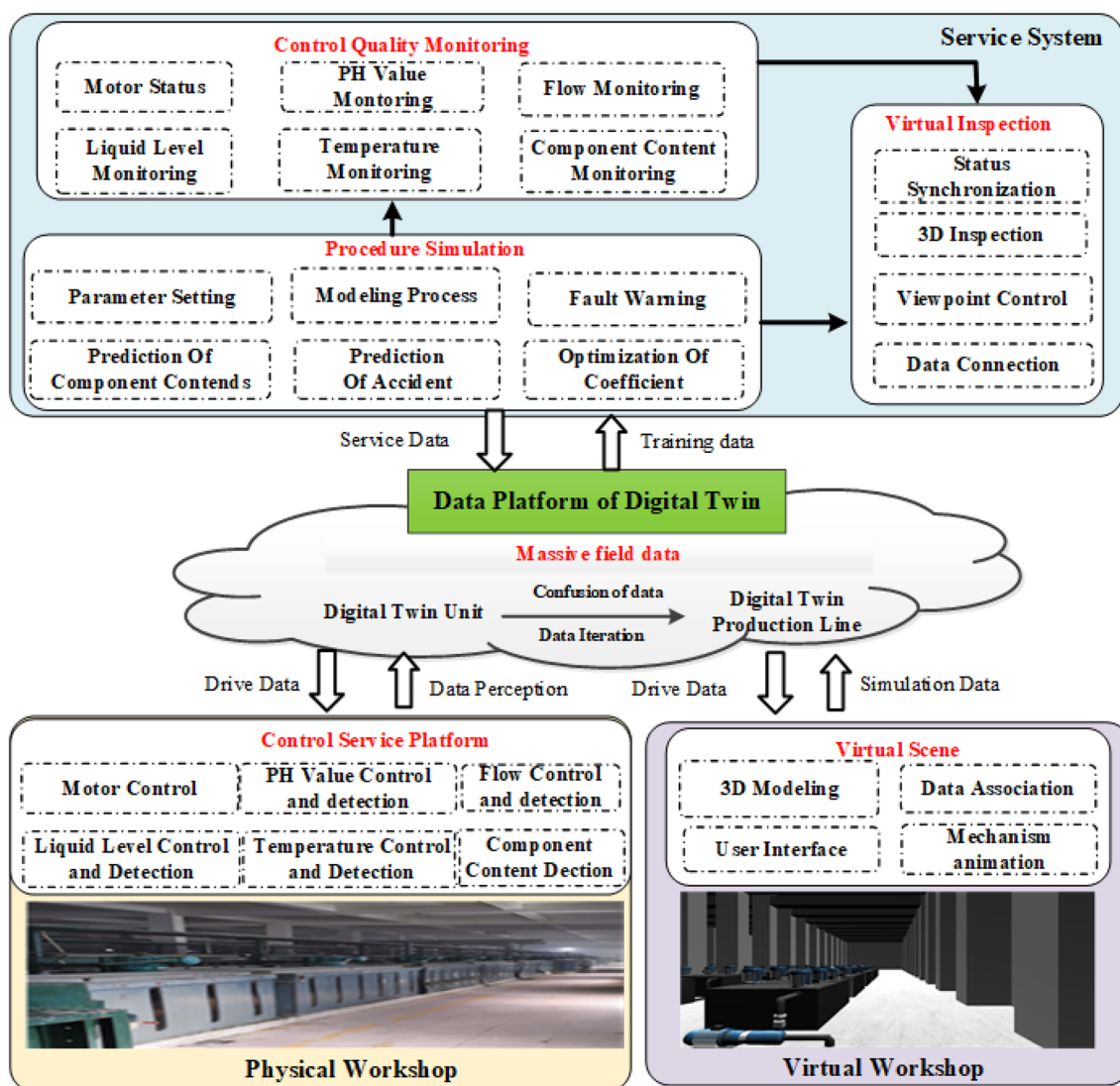
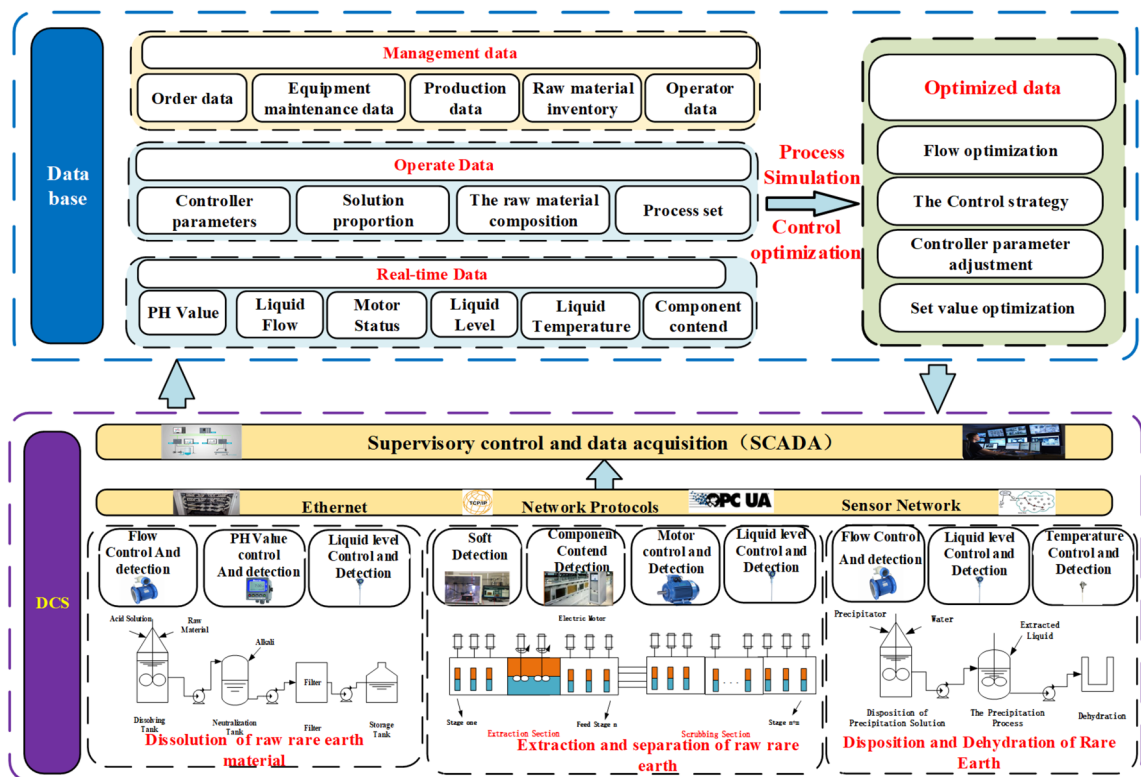


Figure 2. Basic framework of rare earth DT.

products through the action of multi-stage extraction tank. These separated products can be obtained by adding precipitating agent in the precipitation process and stirring to obtain various desired products.

Figure 2 depicts the four key components of the DT architecture, which comprise a physical workshop, a virtual workshop in digital space, DT service systems, and dynamic data bases. Firstly, the physical workshop transmits the process index control data such as motor, PH, temperature, level, temperature and the collected component content data to the database for easy use by the service system and the virtual system. Then the virtual equipment in the virtual shop will synchronize the collected data in the database and realize the overrun warning and the corresponding process animation.



**Figure 3.** Framework for real-time collection, composition, management data of workshop.

Finally, according to the real-time data, the service system can optimize the control index to facilitate the setting of control parameters and simulate the component content of the extraction process at all levels, while the user can quickly inspect the virtual workshop through the virtual inspection module. Taking the rare earth extraction process as an example, the digital twin system data flow is described as follows.

(1) The physical workshop realizes the control of motor, dosing pump and solenoid valve, and gets the real-time data such as solution component content of detection level and extraction tank level, temperature, inlet and outlet flow through the soft measurement system of solution component content, which is transmitted to the digital twin data platform. (2) The digital twin data platform will update the collected data into the database of the virtual workshop and the database of the service system. (3) The virtual workshop synchronizes the process simulation data from the data platform to the corresponding extraction tank with real-time data. (4) The service system includes control optimization, virtual inspection module, and process simulation module. The process simulation module can read real-time data from the DT data platform and then simulate the component content of each extraction level based on the algorithm. The control optimization module will calculate the optimal control strategy based on the real-time control data and the case-based control optimization strategy.

Figure 3 presents the Framework for real-time collection, composition, management data of workshop. The framework is divided into two parts: DCS and database, where the DCS is responsible for flow, level, temperature, motor control and process index detection in each production step. The control and inspection data are transferred to the SCADA system<sup>29</sup> via Ethernet and sensor networks and then uploaded to the database. The database contains real-time data, management data, operation data, and optimization data. The real-time data mainly includes the actual site control and testing data, and the operation information includes the solution proportioning, controller parameters, raw material component content and other artificially set data. Management data includes order, inventory, operator, equipment, and raw material data. Optimization data is used to store the parameter settings of the control optimization strategy such as controller settings, flow rate settings, and to send them to the SCADA system to guide the production.

### Key technologies

**Soft measurement of rare earth component content.** In actual production, process engineers mainly use the color change of the extracted rare earth solution to determine the component content and abnormal working conditions. Inspired by this, Yang proposed a soft measurement model of component content using the color characteristics of the extracted solution<sup>30</sup>. Zhu further improved the accuracy of the soft measurement model by considering the interference generated by light changes on the solution image acquisition<sup>31</sup>. The soft measurement method proposed in this paper can effectively decrease the influence of illumination on the detection, and the detection accuracy is better.

The soft measurement method is used to detect the color features of the liquid to be tested and then estimate component amount. However, detection errors will arise in the field inspection due to the fading of the light source and the hostile environment. To solve this issue, the initial picture must be compensated before modeling.

it is ought to set an objective function based on the Grey Edge algorithm, establish the parameters to be improved, and then apply the Grey Edge algorithm based on the Genetic Algorithm (GA) to optimize the illumination compensation model. To use Extraction of RGB features and HSI traits from light-compensated acquisition photos as the soft measurement model's input variables. The component content is gently measured using the weighted least squares support vector machine (WLSSVM). To guarantee that the recovered color features are accurate, the lighting conditions at the time of data collecting must be consistent, hence color constancy is required for color feature extraction. The color constancy computation is a two-step method that begins with estimating the picture's light color and ends with using the diagonal model to adjust the image to the standard light. The Grey Edge technique is a well-known unsupervised lighting estimating algorithm. The color constancy approach derived from the gray world hypothesis and the gray edge method is as Eq. (1).

$$\left( \int \left| \frac{\partial^n f^\sigma(X)}{\partial X^n} \right|^p dX \right) = ke^{n,p,\sigma} \tag{1}$$

$f^\sigma = f \otimes G^\sigma$  show the convolution of the image  $f$  and the Gauss filter  $G^\sigma$  and  $\partial^n(.)/\partial X^n$  refers the  $n$  order derivation. The color under two diverse illumination conditions can be translated by a diagonal matrix. For an instance, the color of the image under the light  $e^u = [R^u, G^u, B^u]$  can convert to the color  $f^c$  under the light  $e^c = [R^c, G^c, B^c]$ . only needs to be multiplied by a diagonal matrix, which is as Eq. (2).

$$f^c = \begin{bmatrix} R^c/R^u & 0 & 0 \\ 0 & G^c/G^u & 0 \\ 0 & 0 & B^c/B^u \end{bmatrix} \tag{2}$$

According to the production requirements of the actual site, only the middle solution image is valid for the acquired image, so we only need to calculate the error for the area of the solution region. The color of the central region of the image to be estimated is  $(R_a, G_a, B_a)$ , the average RGB value of the image region after the algorithm is  $(R_c, G_c, B_c)$ . Angle error is calculated by Eq. (3):

$$E_a = \cos^{-1} \left( \frac{(R_a, G_a, B_a) \bullet (R_c, G_c, B_c)}{\sqrt{R_a^2 + G_a^2 + B_a^2} \times \sqrt{R_c^2 + G_c^2 + B_c^2}} \right) \times \frac{180^\circ}{\pi} \tag{3}$$

Pre-correction color  $f^u = [R^u, G^u, B^u]$  is known, only estimate the unknown light source in the image to be corrected  $e^u = [R^u, G^u, B^u]$  and standard light sources in standard images  $e^c = [R^c, G^c, B^c]$ . The corrected color can be obtained  $f^c = (R_c, G_c, B_c)$ . Thus, the image correction problem based on illumination estimation is transformed into an optimization problem containing three parameters with the following optimization formulas as Eqs. (4)–(6).

$$\min E_a = \cos^{-1} \left( \frac{(R_a, G_a, B_a) \bullet (R_c, G_c, B_c)}{\sqrt{R_a^2 + G_a^2 + B_a^2} \times \sqrt{R_c^2 + G_c^2 + B_c^2}} \right) \times \frac{180^\circ}{\pi} \tag{4}$$

$$s.t. \begin{pmatrix} R_c \\ G_c \\ B_c \end{pmatrix} = \begin{bmatrix} R^c/R^u & 0 & 0 \\ 0 & G^c/G^u & 0 \\ 0 & 0 & B^c/B^u \end{bmatrix} \begin{pmatrix} R_u \\ G_u \\ B_u \end{pmatrix} \tag{5}$$

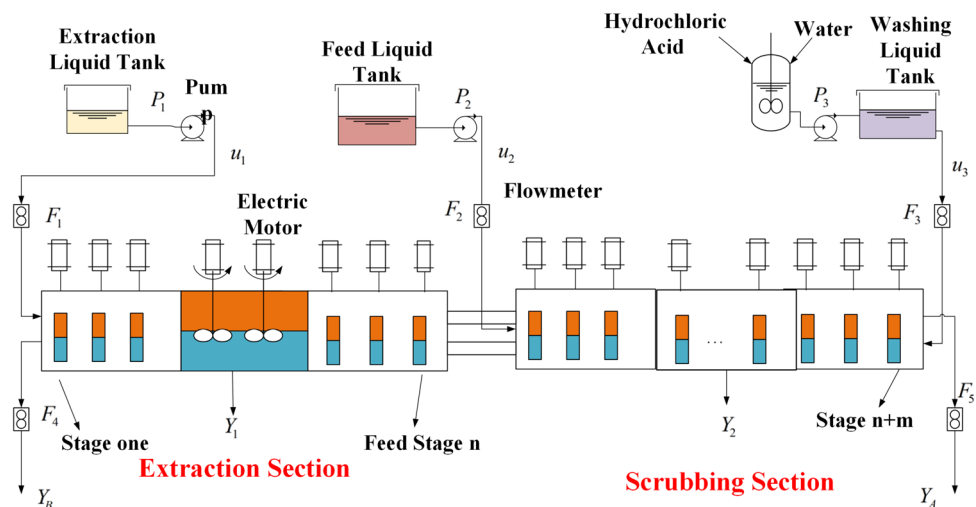
$$ke^{n,p,\sigma} = \left( \int \left| \frac{\partial^n f^\sigma(X)}{\partial X^n} \right|^p dX \right)^{1/p} \tag{6}$$

These three parameters  $n, p, \sigma$  can be used as a set of independent variables, and the best estimate of the actual light color can be obtained by parameter optimization through genetic algorithms (GA), particle swarm optimization (PSO), differential evolution and other optimization algorithms. According to the optimal combination  $n, p, \sigma$ . The estimated light color can be obtained using above equation. After substituting the above results into the transformation matrix The final corrected color is obtained  $f'_c = (R'_c, G'_c, B'_c)$ .

The H, S and I color characteristic components are extracted from the images of rare earth solutions acquired at industrial sites under the HSI color space. Since each value of HSI has different degrees of influence on the component content, each color component is weighted. Similarly, RGB features can be added and the relationship between the solution image color and component content can be expressed as Eq. (7).

$$y = f(w_h C_h, w_s C_s, w_i C_i, w_R C_R, w_G C_G, w_B C_B) \tag{7}$$

$w_h, w_s, w_i, w_R, w_G, w_B$  denote the weights of the H, S and I components of the HSI color space, respectively,  $C_h, C_s, C_i, C_R, C_G, C_B$  The H, S and I components of the HSI color space and R G B components of the RGB color space are represented respectively. soft measurements of rare earth component content are mainly established, with a nonlinear relationship between the solution color HSI characteristics and the rare earth component content. The nonlinear relationship can be modeled using nonlinear regression, neural network methods. Due to the limitations of the field process, it is difficult to collect a large amount of continuous data as the training set of the neural network, so the model of soft measurement uses WLSSVM. this method features fast learning capability, simple calculation and is suitable for small sample objects.



**Figure 4.** Structure of rare earth extraction process.

**Process simulation of the extraction process.** Rare earth, a vital raw material, has been widely employed in the production of military projects, electronic and communication equipment. The process of rare earth production can be classified a series of sub-processes. What matters most is Extraction Process which characterized by nonlinear behavior, large time delays, and strong coupling of various process variables. The extraction process has the features of long process flow, complex mechanism and many influence factors. Figure 4 shows the flow chart of rare earth extraction process. The process consists of an extraction stage and a washing stage. The mixed liquid of raw material liquid is poured into middle stage, the extraction solvent P507 and washing liquid HCL are separately added into first stage and last stage. Hard extracted product  $Y_B$  in aqueous phase can be exported, and the easy extracted product  $Y_A$  in organic phase is originated from the last stage. The organic phase flows in the opposite direction of the aqueous phase.

An accurate Mathematical model is critical in anticipating the states of component content in each mixer-setter in order to simulate the component content. To simulate the extraction process, Wu developed a primary model that relies on mass and element balance<sup>32</sup>. Most academics deny the accuracy, yet it sheds insight on a suitable strategy for rare earth extraction simulation. JIA developed a soft-sensor model technique by incorporating the subtraction clustering algorithm<sup>33</sup>. Yang proposed a model that was based on an enhanced principal model with optimized parameters<sup>34</sup>. The previous process simulation method mainly uses a static mechanism method, which cannot adjust the model according to the actual field data, resulting in large errors in the predicted component content. In this paper, the approaches are utilized to develop a dynamically optimized model of process simulation which could further improve the accuracy of process simulation Training with historical data Compared to previous studies.

When the extraction system reached stability, the organic phase component was set to  $Y$  and the aqueous phase component to  $X$ ;  $A$  denotes easy extracted component content,  $B$  denotes hard extracted component content,  $N$  denotes the number of components,  $I$  denotes the number of components, and  $F$  denotes the raw material composition. There are  $n$  extraction steps and  $m$  scrubbing stages. Since its material liquid remains invariant after extraction, the following formulas for the separation coefficient of various components may be obtained as Eqs. (8)–(10).

$$\beta_{(i+1)/i} = \frac{Y_{i+1} \times X_i}{Y_i \times X_{i+1}} \quad (8)$$

$N$  refers the last component,  $i$  refers component  $i$ .

$$\beta_{(1)/i} = \frac{Y_i \times X_1}{Y_1 \times X_i} \quad (9)$$

$$\beta_{(i)/N} = \frac{Y_N \times X_i}{Y_i \times X_N} \quad (10)$$

To describe the relationship of single component between aqueous phase and organic phase, there are some formulas about the transformation with separation factor as Eqs. (11)–(12).

$$Y_i = \left( X_i \times \prod_{i=1}^N \beta_{1/i} \right) / \sum_{i=1}^N \left( X_i \times \prod_{i=1}^N \beta_{1/i} \right) \quad (11)$$

$$X_i = \left( Y_i \times \prod_{i=1}^N \beta_{1/N} \right) / \sum_{i=1}^N \left( Y_i \times \prod_{i=1}^N \beta_{1/N} \right) \quad (12)$$

According to the insufficient extraction, a compensation coefficient  $K$  have to be added in extraction coefficient to improve accuracy of model. the expression like Eqs. (13)–(14).

$$Y_i = \left( X_i \times \prod_{i=1}^N K * \beta_{1/i} \right) / \sum_{i=1}^N \left( X_i \times \prod_{i=1}^N \beta_{1/i} \right) \quad (13)$$

$$X_i = \left( Y_i \times \prod_{i=1}^N K * \beta_{1/N} \right) / \sum_{i=1}^N \left( Y_i \times \prod_{i=1}^N \beta_{1/N} \right) \quad (14)$$

To increase the accuracy of this model, an improved PSO (Particle Swarm Optimization) was employed in the optimization of compensation coefficient. It is necessary to set fitness function of optimization like following depiction. The fitness function of extraction stage as Eqs. (15)–(16).

$$\min J = \sum_{i=1}^L \left( Y_i - \left( X_i \times \prod_{i=1}^N K * \beta_{1/i} / \sum_{i=1}^N \left( X_i \times \prod_{i=1}^N K * \beta_{1/i} \right) \right) \right)^2 \quad (15)$$

$$\min G = \sum_{i=1}^{N-L} \left( X_i - \left( Y_i \times \prod_{i=1}^N K * \beta_{1/N} / \sum_{i=1}^N \left( Y_i \times \prod_{i=1}^N K * \beta_{1/N} \right) \right) \right)^2 \quad (16)$$

An appropriate compensation coefficient requires the constant iteration through the optimization algorithm and massive date. Here presents the main process: In the initial stage of building mathematical model, the principal model with compensation coefficient was built before Application, the coefficient will be iterated by update parameters.

Compared the common PSO model, improved PSO model based on functional inertia weight and constant constriction factor to optimize compensation coefficient which is more excellent in effect of optimization. This paper is enlightened by the asynchronously improved PSO model applied in text feature selection and utilize it in compensation coefficient optimization<sup>35</sup>.

Common PSO model creates a large number of particles, which move in whole solution space with a fixed law. Each particle own a couple of messages about their position and fitness which also influence other particles to adjust their position and velocity of moving. The velocity of particle  $i$  expresses as  $V_i = (V_{i1}, V_{i2}, \dots, V_{iD})$ , the location of particle  $i$  express as  $(X_{i1}, X_{i2}, \dots, X_{iD})$ , the optimal location of particle  $i$  expresses as  $P_i = (p_{i1}, p_{i2}, \dots, p_{iD})$  it is also regarded as  $p_{best}$ . The best optimum location is  $P_g = (p_{g1}, p_{g2}, \dots, p_{gD})$  or express as  $g_{best}$ . Each particle own a single fitness calculated from fitness function. This kind of improved PSO concentrate on the promotion of update formulae of the dimension  $d$  in Eqs. (17) and (18):

$$v_{id} = w \times v_{id} + c_1 \times \text{Rand}() \times (P_{id} - x_{id}) + c_2 \times \text{Rand}() \times (P_{gd} - x_{id}) \quad (17)$$

$$x_{id} = x_{id} + v_{id} \quad (18)$$

In standard PSO,  $Q$  refers particle quantity,  $w$  refers inertia weight  $c_1$  and  $c_2$  both refers acceleration constant,  $v_{max}$  is the maximum velocity,  $G_{max}$  refers the maximum number of iterations  $\text{rand}()$  or  $\text{Rand}()$  both refers the random functions with values in  $[0,1]$ .

Due to massive data conducted in production process, common PSO model fail to rapidly search the global optimum through the fixed generation. the constriction factor  $K$  is conducted into Eq. (19) show the improved velocity formula.

$$v_{id} = K[v_{id} + c_1 \times \text{Rand}() \times (P_{id} - x_{id}) + c_2 \times \text{Rand}() \times (P_{gd} - x_{id})] \quad (19)$$

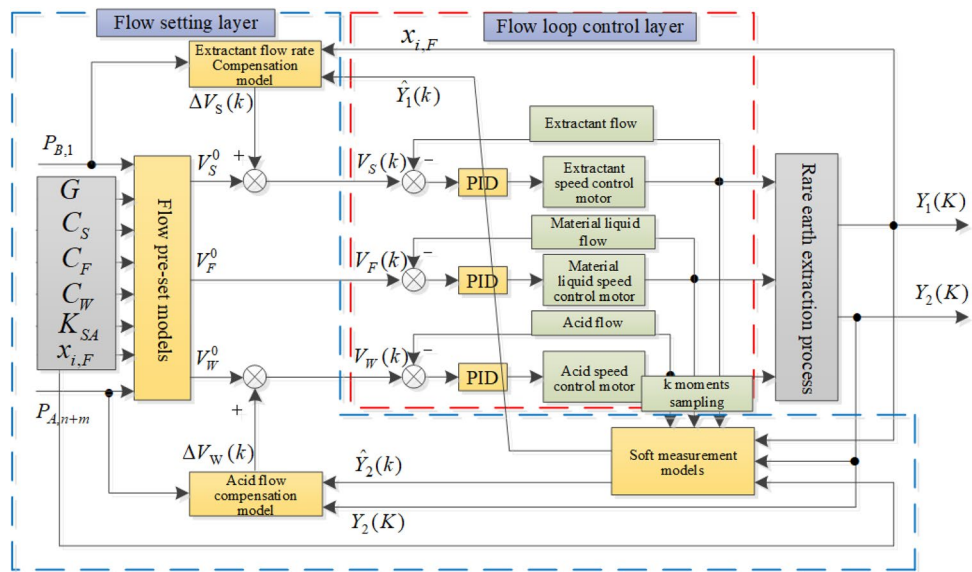
In the early generation, the particle needs to move in a wide range to search the optimum location. it needs to develop the precision within a small range to determine the best single point. Hence, the constriction factor keep a large value in early iteration, and adjust its values with the generation change. the functional constriction factor is presented as Eq. (20). refers the iterations.

$$K = \frac{\cos(\Pi/G_{max}) \times T + 2.5}{4} \quad (20)$$

To Improve asynchronously inertia weight and constriction factor, the principal that original speed influenced by inertia weight, the convergence of PSO is affected by constriction can to be described as Eq. (21).

$$\begin{cases} v_{id} = w \times v_{id} + c_1 \times \text{Rand}() \times (P_{id} - x_{id}) + c_2 \times \text{Rand}() \times (P_{gd} - x_{id}) & \text{if } T < \frac{G_{max}}{2} \\ v_{id} = K[\alpha \times v_{id} + c_1 \times \text{Rand}() \times (P_{id} - x_{id}) + c_2 \times \text{Rand}() \times (P_{gd} - x_{id})] & \text{if } T \geq \frac{G_{max}}{2} \end{cases} \quad (21)$$

$\alpha$  is the default original velocity of particle. The change way of appears in Eq. (22).



**Figure 5.** Control strategy for extraction process twin.

$$\begin{cases} w = w_{end} + (w_{start} - w_{end})(1 - (T/G_{max})) \text{ if } P_{gd} \neq x_{id} \\ w = w_{end} \text{ if } P_{gd} = x_{id} \end{cases} \quad (22)$$

$w_{end}$  and  $w_{start}$  both are set by researchers, which determine the range of inertia weight. To predict the future behavior of plant-wide process, internal parameters of a principle model must be obtained and updated. The actual production process can infer a real model through idealization Approximation. Subsequently, a series of formulae are abstracted from this hypothetical model and constitute a principle model. The correlation between the principle model and actual collected data needs to be ensured by at least one optimal parameter. The optimization model is perfect by iterative optimization using improved PSO.

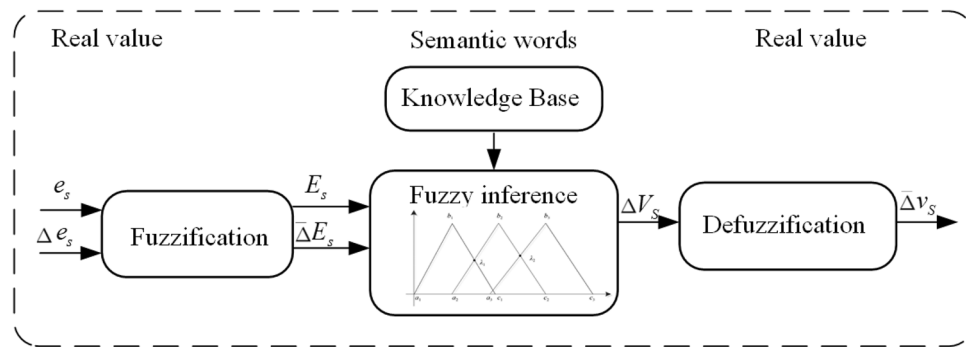
**Process control strategy.** The current production process has a low level of automation and production optimization relying on manual experience. In order to guarantee the quality of the extracted products, it is necessary to design a dosing control strategy with optimization function. Yang proposed the component content distribution profile control which can be automatically regulated by dynamically compensating the related extract or scrubbing liquid flow-rate<sup>36</sup>. Lu proposed a model based on generalized prediction of the content of rare earth extracted components to meet the demand of component content floating in a specific interval<sup>37</sup>. A paper applied Static Setting and dynamic compensation based optimal control strategy for the Flow Rate of the Reagent in Ce Pr/Nd Extraction Process<sup>38</sup>. It can be depicted as Fig. 5. The control strategy in this thesis can quickly generate optimized control parameters based on real-time measured component content and historical cases, which is more practical than previous studies.

Staff usually rely on operational experience to pre-set the extractant and detergent flow rate values based on the raw material processing volume, rare earth feed liquid element distribution, saponification degree, extractant concentration, feed liquid concentration, acid concentration, and other entry conditions during the actual production process of the entire rare earth extraction. The control system obtains the detergent and acid settings from the case library that best match the real process based on the production parameters of the actual production process. CBR(Case-based reasoning) is an artificial intelligence reasoning strategy that uses human thinking to gather expert knowledge from instances in order to solve new issues. The case inference approach is utilized to predetermine the extractant/detergent for the rare earth extraction procedure based on this. Using case-based reasoning, the starting value of the extractant and detergent for the rare earth extraction method was derived. To decide which class a new instance belongs to, the classifier based on CBR approach is primarily focused on the known classes in of a known class. The data demonstrates that the first production parameters, Process Indicators, both have a substantial influence on the initial setup of extractant and acid by studying the process mechanism and process circumstances of the rare earth extraction manufacturing process. Each case consists of a description of the working conditions and a case solution stated by Eq. (23):

$$Case(X_k, V_k) = Case((x_{1,k}, \dots, x_{7,k}), (v_{1,k}, v_{2,k})) \quad (23)$$

$X_k = (x_{1,k}, \dots, x_{7,k})$  denotes the kth condition description.  $x_{1,k}, \dots, x_{7,k}$  presents all initial production parameters, Process Indicators.  $Y_A$  refers the rare earth material distribution,  $C_F$  refers raw material Liquid concentration,  $C_S$  refers Extractant concentration,  $C_W$  refers Detergent concentration,  $G$  refers Material and liquid handling capacity,  $P_{B1}$  and  $P_{A,n+m}$  both are Product purity indexes.  $v_k = (v_{1,k}, v_{2,k})$  enotes the solution corresponding to the kth condition description;  $v_{1,k}, v_{2,k}$  respectively denote the extraction volume preset  $V_s$  and the washing





**Figure 6.** Compensated control based on fuzzy inferences.

volume preset  $V_w$  respectively. The case retrieval strategy can employ the nearest neighbor search strategy. case retrieval strategy can employ the nearest neighbor retrieval strategy, by calculating the distance between the target case and the source case. The smaller the distance, the higher the similarity between the two cases. Multiple search strategies are feasible such as KNN (K-nearest neighbor) search strategy. Intelligent compensation model for initial flow setting using a preset model based on case-based reasoning. Then a relatively reasonable preset value of extractant and detergent flow rate can be obtained. but the model does not take into account of the working conditions caused by real-time external disturbances. To ensure the requirements of the production process, the staff needs to periodically adjust the settings of the extractant and detergent according to the color of the solution.

Intelligent compensation model for initial flow setting is beneficial for the optimism of initial flow setting. The model mainly consists of a soft measurement model based on the color of rare earth solution and an intelligent compensation model based on fuzzy inference. The collected color images of rare earth solutions were extracted and converted to tree components. The H, S and I color components are extracted from the HSI color space. The color characteristics do not have a one-to-one correspondence with the component content, but the Least Squares Support Vector Machine can be applied in the Description of this correspondence since its excellent performance for modeling with small samples. To simulate the regulation of the initial flow by expert experience, a fuzzy control method can be utilized to build an intelligent compensation model. The process is presented as Fig. 6.

The component content  $Y_B$  and  $Y_A$  obtained from soft measurement model are compared with target component content  $P_{B,1}$  and  $P_{A,n+m}$ . The error can be presented as formulas  $e_s = Y_A(K-1) - P_{A,n+m}$ ,  $e_w = Y_B(K-1) - P_{B,1}$ . The Variation of error show as  $\Delta e_s = Y_A(K-1) - Y_A(K-2)$ ,  $\Delta e_w = Y_B(K-1) - Y_B(K-2)$  The variation is set as the input of the fuzzy control model and the variation of extractant and acid is available as output.

## Model validation and case study

**Soft measurement model validation of solution component content.** In this paper, the solutions in 100 extraction tanks are collected as experimental samples, and after collecting the solution images, light compensation is performed with soft measurement modeling.

After extracting the HSI and RGB features of the compensated images, they are substituted into the WLSSVM model to obtain the predicted output of the component content. To Evaluate the accuracy of the soft-measurement model by comparing the actual assay values with the soft-measurement component content values. Relative error percentage as the Critical metrics for evaluating models as shown in Fig. 7. In order to verify the proposed method, the experiments are conducted in four cases, WCC means without color compensation, and CC means with color compensation. The experiments show that CC-WLSSVM has the lowest relative error compared to the actual value and the soft measurement effect is the last.

**Validation of process simulation methods.** The process parameters set in this paper are as follows: The Ce component content of the inlet is 0.445, the Pr component content of inlet is 0.1842, and the Nd component content of inlet is 0.3713. The separation coefficient between Ce and Pr is 2.03 and The separation coefficient between Pr and Nd is 1.55. The number of extraction stages  $n$  is 26 and the number of washing stages  $m$  is 34. The extraction amount is 0.9995 and the export index is 0.9995. The parameters of the improved PSO are set as: acceleration constants  $c_1$  is 1.3,  $c_2$  is 1.7,  $w_{end}$  is set as 0.86  $w_{start}$  is set as 1,  $\alpha$  is 0.7. The predictive values of the component content of rare earth components at each stage are shown in Fig. 8.

To further evaluate the prediction effect of the model, In this paper, we use the mean relative error and the maximum relative error to evaluate the deviation of the model from the actual value, and the root mean square error to reflect the degree of deviation as Eqs. (24)–(26).  $z$  is the model output value,  $z^*$  is the actual component content. Relative error of each component after calibration as shown in Table 1.

$$MEANRE = \frac{1}{n+m} \sum_{i=1}^{n+m} \frac{|z - z^*|}{z^*} \times 100 \quad (24)$$

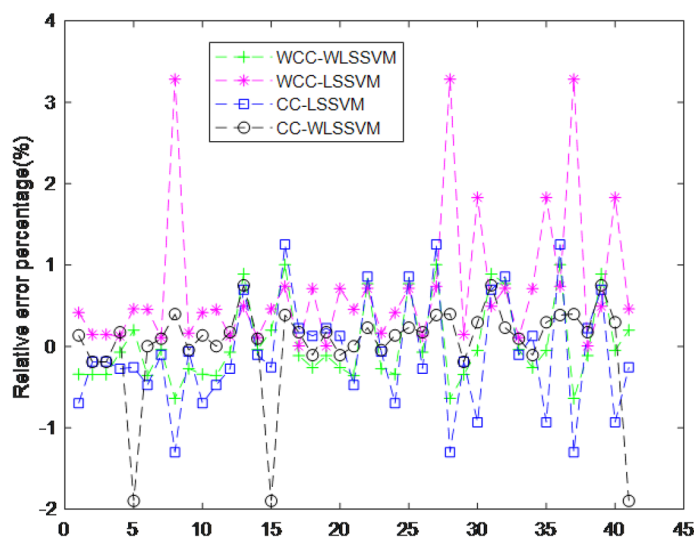


Figure 7. Relative error percentage of WLSSVM.

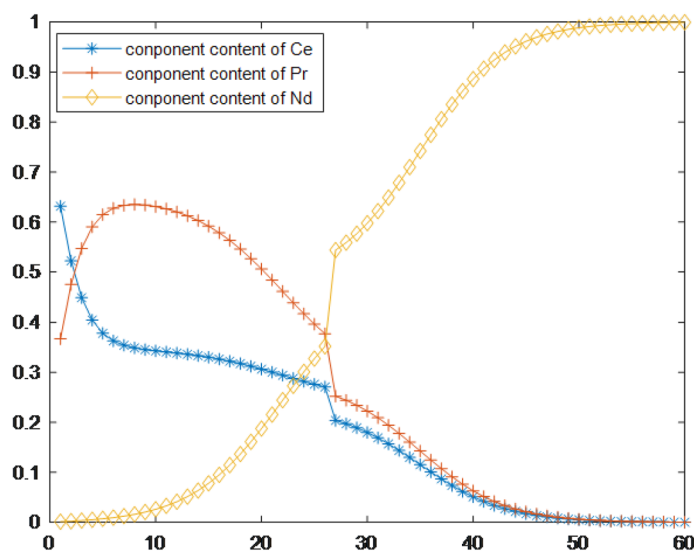


Figure 8. Predicted values of component content for each stage.

Component content	MEANRE(%)	MAXRE(%)	RMSE(%)
Ce	0.425422	4.434351	$1.382 \times 10^{-4}$
Pr	0.417555	3.656577	$2.246 \times 10^{-4}$
Nd	0.175327	3.269205	$1.439 \times 10^{-4}$

Table 1. Relative error of each component after calibration

$$MAXRE = MAX \left( \frac{|z - z^*|}{z^*} \times 100\% \right) \tag{25}$$

$$RMSE = \sqrt{\frac{1}{n+m} \sum_{i=1}^{n+m} (z - z^*)^2} \tag{26}$$

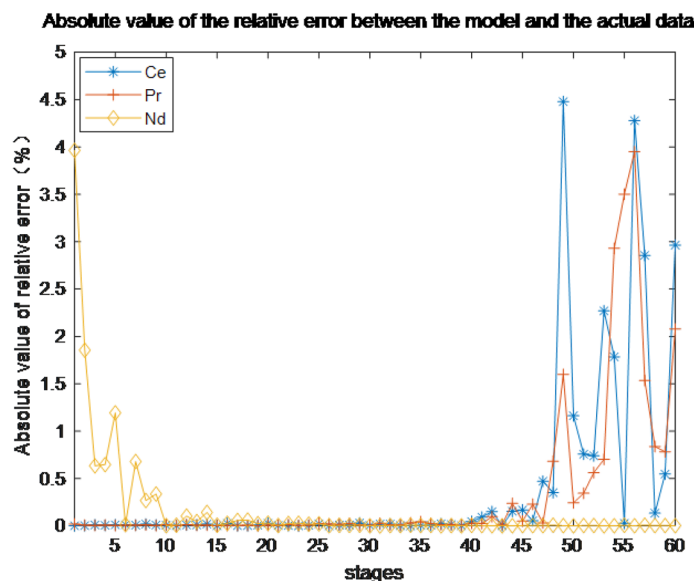


Figure 9. Absolute value of the relative error between the model and actual data twin.

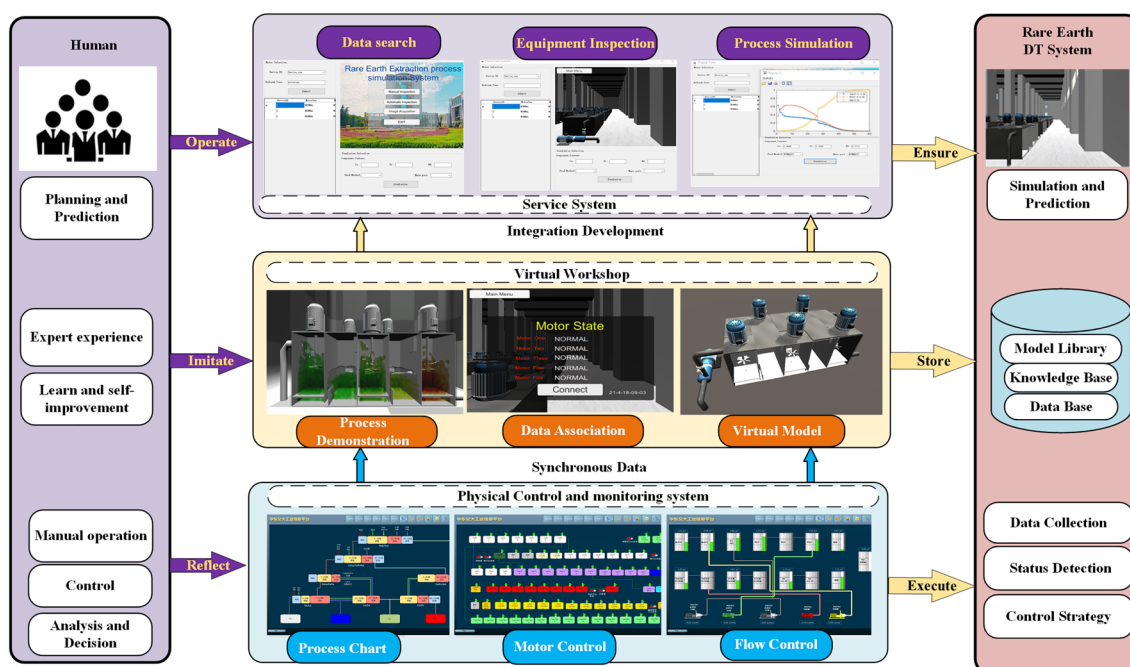
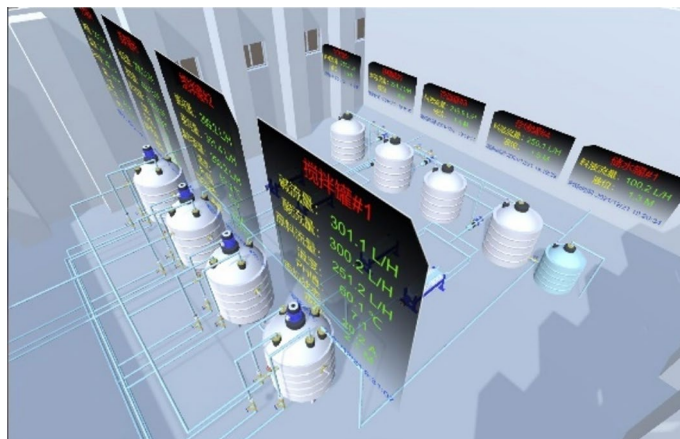


Figure 10. Test of rare earth DT system.

The absolute value of the relative error between the prediction results and the actual data is shown in Fig. 9. If the maximum relative error of the component content of each level does not exceed 5%, the error of the model is deemed to lie within a reasonable error range, and it can be seen from the figure that the error of the established model is within the normal range.

**System development and case study.** The specific applications include process simulation, virtual inspection, and data search. Finally, the contributions and limits of this research are examined. To further illustrate the efficacy of the aforementioned technique and framework, this study will perform experiments in the following areas: physical control and monitoring system, virtual workshop, service system, and human-DT connection. Figure 10 depicts the total framework diagram.

This paper is based on an existing information management system for rare earths production that has been enhanced to enable the sensing, control, and data gathering tasks of the digital twin plant. Configuration software such as WINCC can be used to implement the module. The Physical control and monitoring system is



**Figure 11.** virtual workshop for rare earth process.

made up of three parts: the flow chart, the motor control module, and the flow control module. Here, the flow chart displays all procedures for monitoring each advice, such as the mixing motor and the extraction tank, and allows users to easily access the control and monitoring interface for the related process. The flow control and motor control modules generate reference control parameters such as material flow, extractant flow, and detergent flow. To develop control optimization strategies, these two modules can use the control optimization technique outlined above. The real-time data is synced to the database every two hours with the changeover of communication equipment for further analysis and utilization.

Virtual workshop is mainly convenient for users to quickly inspect the status of equipment in each workshop, and real-time working conditions which is integrated into the DT service system. The rare earth production line's virtual workshop module is created by merging a 3D model with real-time data. Virtual-real interaction, data synchronization, and consistent behavior with Physical Equipment are the major goals of the virtual workshop. As a result, the virtual workshop may be broken down into three sections: process demonstration, data association, and virtual model. This section mostly relies on 3Dmaxs and Unity software to fulfill its goals. Demonstrating the change and distribution of materials liquid in the manufacturing process, which is accomplished by Fluid Animation in Unity, is a process demonstration. Data Association is realized by C# scripts about connection of data base with virtual model. The virtual system enables free inspection of each equipment and fast warning of abnormal working conditions, which greatly reduces production risks and unnecessary waste. To the end, A virtual workshop that allows data synchronization and free inspection is shown as Fig. 11.

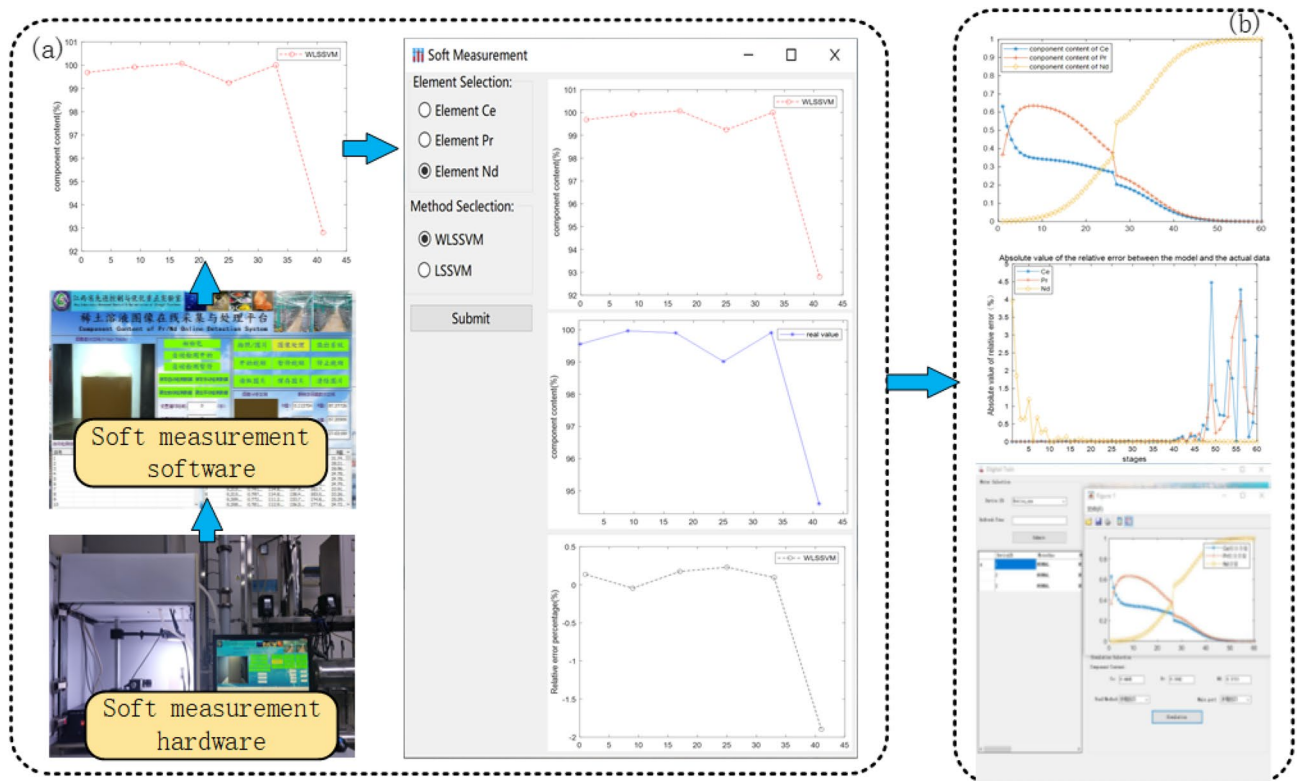
The digital twin service system mainly realizes process simulation of rare earth production, virtual inspection, and soft measurement data retrieval of component content. The virtual workshop module, algorithms, and C# scripts, on the other hand, are enclosed inside subroutines for DT service system calls. A state searching application that can quickly display the states of motors retrieved from a database. To avoid possible failures in the process, the system may scan motor states using the time stamp and device ID and provide the predictive outcome of the component malfunction. The purpose of the virtual inspection is to guarantee that each piece of equipment receives a free and intuitive check, as well as a defect alarm. The process simulation is divided into two major subsystems: component content prediction for the extraction process and soft component content measurement. The proposed model with compensating coefficient is used to forecast component content.

Process simulation is the core function of DT system, to realize this function, it is necessary to solve the problem of soft measurement of component content. A subsystem was built based on the previously described soft measurement method for component content and Existing Soft measuring equipment and software.

After extracting the HSI and RGB features of 6 compensated images collected from Detection level, those are substituted into the WLSSVM model to obtain the predicted output of the component content. To Evaluate the accuracy of the soft-measurement model by comparing the actual assay values with the soft-measurement component content values. The results are shown in the Fig. 12a.

From Relative error percentage of soft measurement, it can be seen that the model is able to accurately predict the component content based on the RGB and HSI characteristics of the solution graph. In this paper, 200 groups of component contents were detected as the training set of the process simulation model. A mathematical model with parameter optimization for component content prediction can be achieved using the above method. In order to verify the applicability of the model, some parameters are preset, a large amount of actual data is provided to train the model, and finally the prediction effect of the model is evaluated using the data from the validation set. The predictive values of the component content of rare earth components at each stage are shown in Fig. 12b. It can be seen from the error of the established that model is within the normal range.

The developed rare earth digital twin service system has greatly reduced the inspection time and detection time, the inspection time from the original 30 min to 5 min, and the component content detection time from the original 1 h to 10 min.



**Figure 12.** subsystem of DT service system (a) soft measurement of component content (b) process simulation of component content.

## Conclusions

Rare earth production process relies heavily on manual labor for solution component content detection, process control and equipment inspection, and the production process has a large time lag and is prone to abnormal working conditions caused by untimely troubleshooting, which eventually affects the output quality of products. This paper proposes a soft measurement method of solution component based on the color characteristics of rare earth solution and a component process simulation method based on mechanism compensation for the problem of untimely detection of solution components, and adopts a case-based control optimization strategy for the problem of difficult control optimization, and builds a virtual scene to facilitate users to quickly understand the process and equipment conditions in order to reduce manual inspection. The soft measurement method of solution components is easily affected by the light and transparency of the container. The following points will be improved in future studies: Soft measurement method for solution fraction is slow to calculate and accuracy needs to be enhanced. The process simulation algorithm will be further improved to enhance the simulation accuracy. The control effect of the control strategy needs to be further improved, and the integration degree of the current DT system. The usability of the virtual workshop need to be enhanced. The interface will be presented in the form of a web page, and the server will complete the calculation and prediction.

## Data availability

The datasets generated and analysed during the current study are not publicly available due to sensitive data involving production units but are available from the corresponding author on reasonable request.

Received: 19 April 2022; Accepted: 24 August 2022

Published online: 30 August 2022

## References

- Duan, L. & Da Xu, L. Data analytics in industry 4.0: A survey. *Inf. Syst. Front.* <https://doi.org/10.1007/s10796-021-10190-0> (2021).
- Konstantinov, S., Ahmad, M., Ananthanarayan, K. & Harrison, R. The cyber-physical e-machine manufacturing system: Virtual engineering for complete lifecycle support. *Procedia CIRP* **63**, 119–124. <https://doi.org/10.1016/j.procir.2017.02.035> (2017).
- Uhlemann, T. H. J., Lehmann, C. & Steinhilper, R. The digital twin: Realizing the cyber-physical production system for industry 4.0. *Procedia Cirp* **61**, 335–340. <https://doi.org/10.1016/j.procir.2016.11.152> (2017).
- Tao, F., Zhang, H., Liu, A. & Nee, A. Y. Digital twin in industry: State-of-the-art. *IEEE Trans. Ind. Inf.* **15**, 2405–2415. <https://doi.org/10.1109/TII.2018.2873186> (2018).
- Grievies, M. Digital twin: Manufacturing excellence through virtual factory replication. *White Pap.* **1**, 1–7 (2014).
- Yang, L., Chen, S., Wang, X., Zhang, J. & Wang, C. Digital twins and parallel systems: State of the art, comparisons and prospect. *Acta Autom. Sin.* **45**, 2001–2031. <https://doi.org/10.16383/j.aas.2019.y000002> (2019).

7. Gockel, B., Tudor, A., Brandyberry, M., Penmetsa, R. & Tuegel, E. Challenges with structural life forecasting using realistic mission profiles. In: *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, 1813 (2012).
8. Bielefeldt, B., Hochhalter, J. & Hartl, D. Computationally efficient analysis of sma sensory particles embedded in complex aerostuctures using a substructure approach. In: *Smart Materials, Adaptive Structures and Intelligent Systems*, vol. 57298, V001T02A007, <https://doi.org/10.2514/6.2012-1813> (2015).
9. Wang, J., Ye, L., Gao, R. X., Li, C. & Zhang, L. Digital twin for rotating machinery fault diagnosis in smart manufacturing. *Int. J. Prod. Res.* **57**, 3920–3934. <https://doi.org/10.1080/00207543.2018.1552032> (2019).
10. Jiang, Z., Lv, H., Li, Y. & Guo, Y. A novel application architecture of digital twin in smart grid. *J. Ambient Intell. Human. Comput.* <https://doi.org/10.1080/00207543.2018.1552032> (2021).
11. Li, C., Mahadevan, S., Ling, Y., Choze, S. & Wang, L. Dynamic Bayesian network for aircraft wing health monitoring digital twin. *Aiaa J.* **55**, 930–941. <https://doi.org/10.2514/1.J055201> (2017).
12. Oyekan, J. O. *et al.* The effectiveness of virtual environments in developing collaborative strategies between industrial robots and humans. *Robot. Comput. Integr. Manuf.* **55**, 41–54. <https://doi.org/10.1016/j.rcim.2018.07.006> (2019).
13. Burghardt, A. *et al.* Programming of industrial robots using virtual reality and digital twins. *Appl. Sci.* **10**, 486. <https://doi.org/10.3390/app10020486> (2020).
14. Bilberg, A. & Malik, A. A. Digital twin driven human-robot collaborative assembly. *CIRP Ann.* **68**, 499–502. <https://doi.org/10.1016/j.cirp.2019.04.011> (2019).
15. Corral-Acero, J. *et al.* The 'digital twin' to enable the vision of precision cardiology. *Eur. Heart J.* **41**, 4556–4564. <https://doi.org/10.1093/eurheartj/ehaa159> (2020).
16. Kang, J.-S., Chung, K. & Hong, E. J. Multimedia knowledge-based bridge health monitoring using digital twin. *Multimedia Tools Appl.* <https://doi.org/10.1007/s11042-021-10649-x> (2021).
17. Tao, F. *et al.* Five-dimension digital twin model and its ten applications. *Comput. Integr. Manuf. Syst.* **25**, 1–18 (2019).
18. Stark, R., Fresemann, C. & Lindow, K. Development and operation of digital twins for technical systems and services. *CIRP Ann.* **68**, 129–132. <https://doi.org/10.1016/j.cirp.2019.04.024> (2019).
19. Jones, D., Snider, C., Nassehi, A., Yon, J. & Hicks, B. Characterising the digital twin: A systematic literature review. *CIRP J. Manuf. Sci. Technol.* **29**, 36–52. <https://doi.org/10.1016/j.cirpj.2020.02.002> (2020).
20. Schluse, M., Priggemeyer, M., Atorf, L. & Rossmann, J. Experimentable digital twins-streamlining simulation-based systems engineering for industry 4.0. *IEEE Trans. Ind. Inf.* **14**, 1722–1731. <https://doi.org/10.1109/TII.2018.2804917> (2018).
21. Zheng, Y., Yang, S. & Cheng, H. An application framework of digital twin and its case study. *J. Ambient Intell. Human. Comput.* **10**, 1141–1153. <https://doi.org/10.1109/10.1007/s12652-018-0911> (2019).
22. Wu, P. *et al.* Visual real-time monitoring method for discrete manufacturing workshop based on digital twin. *Com. Int. Man. Sys. China* **27**, 1605–1616 (2021).
23. Zhao, L., Fang, Y., Lou, P., Yan, J. & Xiao, A. Cutting parameter optimization for reducing carbon emissions using digital twin. *Int. J. Precis. Eng. Manuf.* **22**, 933–949. <https://doi.org/10.1007/s12541-021-00486-1> (2021).
24. Zhang, Z. *et al.* Digital twin system design for dual-manipulator cooperation unit. In: *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, vol. 1, 1431–1434. <https://doi.org/10.1109/ITNEC48623.2020.9084652> (2020).
25. Li, Y., Yang, C., ZHANG, H. & LI, J. Discussion on key technologies of digital twin in process industry. *Acta Autom. Sin.* **47**, 501–514. <https://doi.org/10.16383/j.aas.c200147> (2021).
26. Zhou, H., Yang, C. & Sun, Y. A collaborative optimization strategy for energy reduction in ironmaking digital twin. *IEEE Access* **8**, 177570–177579. <https://doi.org/10.1109/ACCESS.2020.3027544> (2020).
27. Soares, R. M., Câmara, M. M., Feital, T. & Pinto, J. C. Digital twin for monitoring of industrial multi-effect evaporation. *Processes* **7**, 537. <https://doi.org/10.3390/pr7080537> (2019).
28. Bevilacqua, M. *et al.* Digital twin reference model development to prevent operators' risk in process plants. *Sustainability* **12**, 1088. <https://doi.org/10.3390/su12031088> (2020).
29. Uhlemann, T.H.-J., Schock, C., Lehmann, C., Freiburger, S. & Steinhilper, R. The digital twin: Demonstrating the potential of real time data acquisition in production systems. *Procedia Manuf.* **9**, 113–120. <https://doi.org/10.1016/j.promfg.2017.04.043> (2017).
30. Lu, R. & Yang, H. Soft measurement for component content based on adaptive model of pr/nd color features. *Chin. J. Chem. Eng.* **23**, 1981–1986. <https://doi.org/10.1016/j.cjche.2015.10.007> (2015).
31. Yang, H. *et al.* Color constancy of rare-earth solution image with adaptive parameter. *IFAC-PapersOnLine* **51**, 295–300. <https://doi.org/10.1016/j.ifacol.2018.09.435> (2018).
32. Sheng, W. Static design for multiple components and multiple outlets rare earth countercurrent extraction (i): Algorithm of static design. *J. Chin. Rare Earth Soc. Chin. Ed.* **22**, 17–21 (2004).
33. Chai, T. & Jia, W. Soft-sensor of element component content based on multiple models for the rare earth cascade extraction process. *Control Theory Appl.* (2007).
34. Yang, H., DAI, W., Lu, R. & ZHU, J. Simulation of rare earth extraction process based on separation coefficient correction. *CIESC J.* **71**, 3180–3190 (2020).
35. Lu, Y., Liang, M., Ye, Z. & Cao, L. Improved particle swarm optimization algorithm and its application in text feature selection. *Appl. Comput.* **35**, 629–636. <https://doi.org/10.1016/j.asoc.2015.07.005> (2015).
36. Yang, H., Xu, F., Lu, R. & Ding, Y. Component content distribution profile control in rare earth countercurrent extraction process. *Chin. J. Chem. Eng.* **23**, 192–198. <https://doi.org/10.1016/j.cjche.2014.09.046> (2015).
37. Yang, H., He, L., Zhang, Z., Lu, R. & Tan, C. Multiple-model predictive control for component content of cepr/nd countercurrent extraction process. *Inf. Sci.* **360**, 244–255. <https://doi.org/10.1016/j.ins.2016.04.031> (2016).
38. ZHU, J., YANG, H., LU, R., XU, F. & YU, Y. Static setting and dynamic compensation based optimal control for the flow rate of the reagent in cepr/nd extraction process. *Acta Autom. Sin.* **45**, 1186–1197. <https://doi.org/10.16383/j.aas.c170666> (2019).

## Acknowledgements

This work was supported by the National key research and Development Program of China No.2020YFC1909005;the National Natural Science Foundation of China Nos. 61733005, 61963015.

## Author contributions

H.Y. Conceptualization, Supervision, Interpretation. Z.K. Editing, Interpretation, Software. J.Z. Methodology, Writing-review, Conceptualization, Supervision, Validation. F.X. Software, Supervision. F.N. Conceptualization, Supervision.S.S. Conceptualization, Supervision.

## Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to J.Z.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022