Intelligent Operational Optimization Method for Salt Lake Chemical Industrial Processes Based on Data-Driven Modeling

2023 IEEE 18th Conference on Industrial Electronics and Applications (ICIEA) | 979-8-3503-1220-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICIEA58696.2023.10241917

Diju Liu
School of Automation
Central South University
Changsha 410083, Hunan, China
djliu@csu.edu.cn

Yalin Wang
School of Automation
Central South University
Changsha 410083, Hunan, China
ylwang@csu.edu.cn

Chenliang Liu
School of Automation
Central South University
Changsha 410083, Hunan, China
lcliang@csu.edu.cn

Kai Wang
School of Automation
Central South University
Changsha 410083, Hunan, China
kaiwang@csu.edu.cn

Xiaofeng Yuan
School of Automation
Central South University
Changsha 410083, Hunan, China
yuanxf@csu.edu.cn

***Abstract*—The process control performance of the washing process heavily influences the quality of products in the Salt Lake Chemical industrial process, which is mainly determined by two subtasks: process modeling and operation optimization. For these two tasks, this paper first demarcates the modeling and optimization objectives of the washing process by profoundly analyzing the process mechanism and operating rules. To fully consider the working condition type and sampling time, a novel global multi-condition-aware transformer (GMCA-Former) network is proposed for the process modeling task. After that, the relatively accurate GMCA-Former is used to simulate the process state. Next, a hybrid operational optimization strategy based on case matching and NSGA-II is applied to find the optimal operational parameter settings efficiently and stably. Notably, an event-triggered process model update mechanism is also designed to ensure the accuracy of the process model. Finally, the effectiveness of the proposed methods is verified in the actual industrial dataset.**

***Keywords*—Operational optimization, Salt Lake chemical industrial process, transformer, data-driven modeling**

I. Introduction

The intelligent improvement of process control in industry is one of the central tasks for modern industrial upgrading [1-3]. Generally, process control means using intelligent techniques to sense the process status and set the values of key operating variables to ensure the process output is within the target range [4]. Accordingly, there are two important sub-tasks in industrial process control: process modeling and operational optimization [5].

Process modeling is designed to sense the current output of the process for the following control action using hardware sensors or intelligent algorithms [6, 7]. If the output control quantity is easy to measure, such as temperature and pressure, then accurate real-time measurement of these variables can be achieved using hardware sensors. However, in actual process industries, some essential output control variables, such as the content of various ions, are often difficult to measure [8, 9]. These variables can only be obtained through offline laboratory analysis, which leads to huge time delays [10]. To solve this problem, soft sensor techniques are proposed and developed rapidly, which means building mathematical models between easy-to-measure physical quantities and difficult-to-measure control indicators to achieve an indirect estimate of control indicator values. Soft sensor modeling techniques for industrial processes can generally be categorized as mechanism-based or data-driven. However, with the increasing complexity of modern industries, mechanism-based modeling has become more challenging to implement effectively, and data-driven modeling has emerged as the predominant approach [11, 12]. Among them, deep learning algorithms represented by neural networks have developed most rapidly due to their excellent nonlinear feature extraction capability. For example, Yuan et al. [13] proposed a data-driven soft sensor modeling method named VW-SAE for predicting the key quality variable in the debutanizer column process. Jiang et al. [14] further discussed the problem of dimensional catastrophe in dealing with high-dimensional industrial data.

Although these successful applications all demonstrate the potential of deep learning algorithms for process sensing tasks, most of them neglect to consider the multi-condition properties of industrial processes [15]. That is, industrial processes often work at several different stable operating points, owing to fluctuations in the properties of the feedstock and changes in the operators. These will lead to similar evolution laws for the same operating conditions, which in turn guides process modeling. Inspired by the above idea, this paper innovatively proposes a novel global multi-condition-aware transformer (GMCA-Former) method for process modeling. GMCA-Former fully considers the similarities between the same working conditions and the influence of different working conditions. Besides, it also pays attention to the declining reference characteristics between samples with time. These advantages enable the

[*]This work was supported in part by the National Key Research and Development Program of China (2020YFB1713800), in part by National Natural Science Foundation of China (NSFC) (92267205), and in part by the science and technology innovation Program of Hunan Province in China (2021RC4054). *(Corresponding author: Chenliang Liu)*

proposed GMCA-Former to sense the industrial process changes timelier and more accurately. Moreover, to ensure effectiveness over time, this paper designs an event-triggered model update strategy to maintain its accuracy when there is a large degradation in performance.

Process modeling is only the basis of process control, while operational optimization is the key to achieving efficient and effective running of the whole process [16]. Operational optimization means optimizing the operation variables to ensure that the whole process runs in an optimal state [17]. Since most operational optimization problems are nonconvex, genetic algorithms are among the most used algorithms. For example, A. Alabdulkarem et al. [18] utilized genetic algorithms to minimize steam consumption in the condensate stabilization process. Bi et al. [19] determined the optimal process parameters for the hydrogen liquefaction process with dual hydrogen refrigeration using a genetic algorithm. However, the performance of traditional genetic algorithms is easily affected by the initial value when applied in the operational optimization process.

To solve this problem, a hybrid operational optimization strategy based on case matching and NSGA-II is developed in this paper. First, the selection criteria of excellent historical cases are determined through an in-depth analysis of the process mechanism. After that, the excellent cases that are most similar to the current case are matched among the excellent historical cases. Then the reference setting value of the current operation variables is obtained by weighting the similarities between the current case and excellent historical cases. Finally, using this set value as the initial value of the NSGA-II to improve the accuracy of the optimal search.

To validate the effectiveness of the proposed method, the mixed potassium washing process in Salt Lake Chemical is selected as validation in this paper. The proposed methods are applied to the process control problem of this process and compared with typical process models and classical operation optimization algorithms. In summary, the main contributions of this paper are as follows.

1) A global multi-condition-aware process modeling approach is proposed to fully perceive operational condition fluctuations;
2) A hybrid operational optimization strategy combining case matching and NSGA-II is designed to improve operational optimization performance;
3) After conducting an in-depth analysis of the Salt Lake Chemical industrial process, operational optimization problems and constraints are established for the mixed potassium washing process;
4) Extensive simulation experiments are constructed to demonstrate the superiority of the proposed methods.

The remainder of this paper is organized as follows. First, the whole problem is systematically described in Section II. Then the proposed GMCA-Former process model is introduced in Section III, and the hybrid operation optimization strategy used is shown in detail in Section IV. The practical application results are shown in Section V. Finally, the systematic summary of the paper is presented.

II. Process and Problem Description

The mixed potassium washing process is one of the most important production processes in Salt Lake Chemical's industrial process. Its main purpose is to adequately mix the crude potassium (mainly potassium chloride) and the soft potassium (mainly picromerite) obtained from the flotation process. The next step is to remove the sodium ions that disturb the subsequent reactions while retaining the valuable potassium ions by using the difference in solubility of the different substances. Fig. 1 presents a diagram of the mixed potassium washing process in Salt Lake Chemical. The feed material includes the crude and soft potassium and the freshwater and E2 mother liquor (the tail liquor of the crystallization process consists of water, potassium ions, sulfate ions, and chloride ions). Then, they are mixed in four tandem washing tanks in turn. The filtrate containing impurities of sodium ions is sent to a filtrate recovery tank for settling and then returned to the decomposition process for further recovery of the small number of valuable potassium ions. At the same time, the slurry containing large amounts of potassium ions is sent to a belt filter via a slurry distributor. Finally, the drier mixed potassium is sent to the crystallization process to produce potassium sulfate pellets.

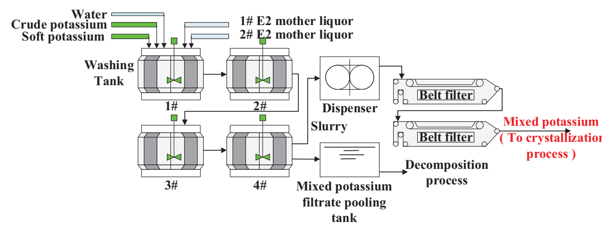

Fig. 1. Schematic diagram of the mixed potassium washing process in the Salt Lake Chemical.

The main goal of the mixed potassium washing process is to maximize the potassium ion content of the washed mixed potassium and minimize the content of impurity ions such as sodium ions. In addition, according to the five-phase system of water and salt, the subsequent crystallization process demands inlet mixed potassium with potassium to sulfate ratio as close as possible to 1.14 to decrease overall energy consumption. Hence, for the optimization of the entire production process, the potassium-to-sulfate ratio of mixed potassium is close to 1.14 is also one of the production requirements for the washing process. As mentioned before, the principle of washing is mainly because the solubility of the impurity sodium ions in the E2 mother liquor is much higher than that of the potassium and sulfate ions. Thus, the amount of 1# and 2# E2 mother liquor added are two extremely important operational variables in the washing process.

However, the solubility of sodium ions in the E2 mother liquor is not always sufficient to ensure that the sodium ions in the mixed potassium enter the solution, which requires an additional input of freshwater to increase the solubility of the sodium ions. In addition, freshwater addition increases the solubility of both potassium and sulfate ions resulting in lower potassium yields, so the amount of fresh water added should be kept to a minimum. Hence the problem of optimizing the washing process can be summarized by establishing the optimum values for three important operating variables: the amount of freshwater addition, the amount of 1# and 2# E2 mother liquor addition to ensure that the potassium content of the washed potash mix is as high as possible, the sodium content is as low as possible and the potassium to sulfate ratio is as close to 1.14 as possible.

Similar to many industrial processes, washing processes suffer from multiple working conditions, i.e., the washing process operates at several different stabilization points due to variations in the properties of the feedstock and the operating conditions of the preceding processes. The feed conditions of the process vary considerably from one working condition to another, and the corresponding rules for the selection of operating variables also differ. However, process industry data is often multidimensional and widely distributed, and the characteristics of the data conditions are difficult to capture. Therefore, accurately sensing the status of the washing process under different operating conditions is one of the challenges for efficient operation optimization.

III. Process Modeling based on Global Multi-Condition-Aware Transformer

To accurately sense the production state of multi-condition industrial processes, this paper designs a global multi-condition-aware transformer network (GMCA-Former) method to predict the ionic grade of washed mixed potassium. This section presents the core module of global multi-condition-aware attention and the overall framework.

A. Global multi-condition-aware attention

It is well known that the same conditions have strong similarities between them, while different conditions also influence each other. Besides, the closer the sampling time interval is, the stronger the interaction effect of the two samples. Based on these two facts, a global multi-condition-aware self-attention mechanism (GMCA-SA) is designed in this paper, as shown in Fig. 2.

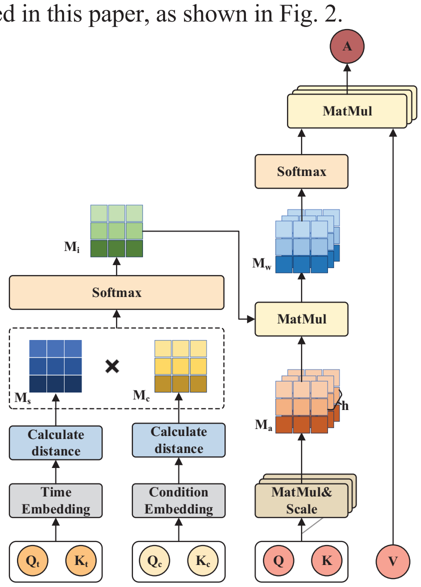

Fig. 2. Illustration of global multi-condition-aware self-attention mechanism.

Suppose the input key-value pairs and query vectors are $\{\mathbf{K},\mathbf{V}\} \in \mathbb{R}^{N\times d}, \mathbf{Q} \in \mathbb{R}^{N\times d}$, the corresponding working condition category centers are expressed as $\mathbf{K}_c \in \mathbb{R}^{N\times d}, \mathbf{Q}_c \in \mathbb{R}^{N\times d}$, where N is the samples number, d is the dimension. First, analogous to the positional encoding in transformer [20], the 1-dimensional time stamp is expanded to d-dimensions for enhancing the presentation capability, which denotes as $\mathbf{K}_s \in \mathbb{R}^{N\times d}, \mathbf{Q}_s \in \mathbb{R}^{N\times d}$. Then the sampling time interval and the working condition category center gap between queries and keys are calculated as

$$\mathbf{D}_{type} = \left(d_{ij} = \left\|\mathbf{Q}_{type}[i,:] - \mathbf{K}_{type}[j,:]\right\|\right) \in \mathbb{R}^{N\times N} \tag{1}$$

where $type = \{s,c\}$. Based on the fact that the closer the sampling time or the closer the working condition category is, the more influential the inter-sample effect is, this paper designs the following monotonically decreasing integrated weights to guide the attention calculation

$$\mathbf{M} = \mathrm{softmax}\left(\exp\left(-\mathbf{D}_s \cdot \mathbf{D}_c\right)\right) \in \mathbb{R}^{N\times N} \tag{2}$$

where $\mathbf{M}$ denotes the normalized integrated weights. By incorporating it into the original self-attention mechanism calculation process, it can assist the model to extract complex correlations between samples faster and more accurately. The formulaic description of the GMCA-SA calculation process is as

$$\mathbf{A} = \mathrm{softmax}\left(\mathbf{M}\frac{\mathbf{QK}^T}{\sqrt{d}}\right)\mathbf{V} \tag{3}$$

where $\mathbf{A} \in \mathbb{R}^{N\times d}$ denotes the extracted inter-sample correlations. Compared with the traditional attention mechanism, the GMCA-SA is superior in considering the time sequence relationship and working conditions more fully.

B. Global multi-condition-aware transformer

Utilizing the GMCA-SA as a fundamental module, this paper designs a global multi-condition-aware transformer framework (GMCA-Former), as illustrated in Fig. 3 . The framework is divided into three steps: KNN-based working condition segmentation, input data division, and feature extraction and prediction. The subsequent sections provide a detailed description of each of these steps.

(1) K-Means-based working condition segmentation

In GMCA-SA, the working condition labels require to be used, but it is very difficult to collect in practice. Thus, the classical K-Means clustering algorithm is used to unsupervised cluster all historical samples, and the obtained clustering center serves as the work condition center, which is expressed as

$$\mathbf{labels}, \mathbf{centers} = \text{K-Means}\left(\mathbf{X}_h\right) \tag{4}$$

where $\mathbf{X}_h \in \mathbb{R}^{N\times d_x}$ denotes the input historical data, d_x is the dimension, $\mathbf{labels} \in \mathbb{R}^{N\times 1}$ and $\mathbf{centers} \in \mathbb{R}^{N\times d_x}$ denote the working condition label and corresponding center, respectively.

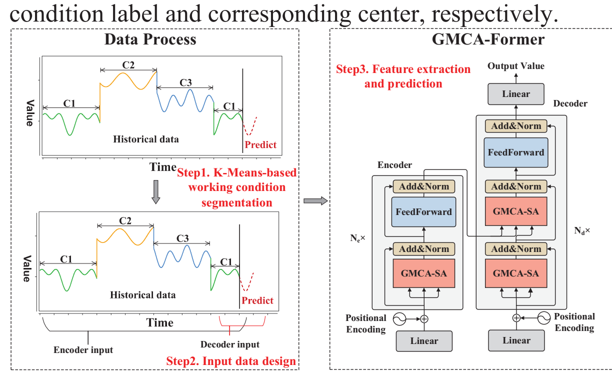

Fig. 3. Illustration of global multi-condition-aware transformer framework.

(2) Input data design

The encoder part of the transformer network serves to extract the complex correlations between the encoder input data. In contrast, the decoder part looks for the valuable parts of the features extracted from the encoder for the task. Therefore, in order to extract the global features, long-term range history samples are fed into the encoder. Whereas, since the temporal nearest neighbor samples are most similar to the future moments, a small portion of the most temporal nearest neighbor samples is used as decoder input to find task-relevant features among the massive history features. Therefore, the input data of GMCA-Former can be organized as

$$\mathbf{X}_e = \mathbf{X}_h\left[-L_e :\right] \tag{5}$$

$$\mathbf{X}_d = \mathbf{X}_h\left[-L_d :\right] \tag{6}$$

where $\mathbf{X}_e$ and $\mathbf{X}_d$ denote the encoder input data and decoder input data, L_e and $L_d\left(L_d \ll L_e\right)$ represent the length.

(3) Feature extraction and prediction

The overall architecture of GMCA-Former is similar to the original transformer network, except that the original multi-head attention mechanism is replaced by the multi-head GMCA-SA. Thus, the calculation process of each of the N_e stacked encoder modules is described as

$$\mathbf{X}_I = PE\left(Linear\left(\mathbf{X}_e\right)\right) \tag{7}$$

$$\mathbf{X}_{SA} = \text{LayerNorm}\left(\mathbf{X}_I + \text{GMCA-SA}\left(\mathbf{X}_I\right)\right) \tag{8}$$

$$\mathbf{X}_A = \text{LayerNorm}\left(\mathbf{X}_{SA} + FeedForward\left(\mathbf{X}_{SA}\right)\right) \tag{9}$$

where $\mathbf{X}_I$ and $\mathbf{X}_{SA}$ are hidden features, $\mathbf{X}_A$ is the final feature of this module. The N_d stacked decoder modules are similar to the encoder computation process. Differently, the key-value pairs in the second attention computation process in Eq. (8) are generated by the final features of the encoder, while the query matrix is generated by the decoder input. Due to the consideration of sample working conditions and sampling time in the feature extraction process, GMCA-Former is more suitable for modeling industrial processes with multiple working conditions than traditional transformer network.

IV. Hybrid Operational Optimization Strategy

In this study, a hybrid strategy for optimizing the operational variables of the Salt Lake washing process is proposed to ensure stable and efficient operation, as depicted in Fig.4. The overall optimization module is composed of case matching and multi-objective optimization NSGA-II algorithm, and the process model is constructed with the GMCA-Former mentioned in Section III. The entire optimization process is described in detail below.

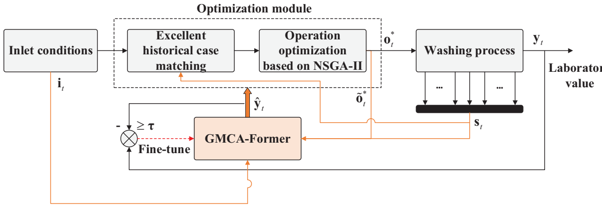

Fig. 4. Illustration of hybrid operation optimization.

The whole optimization process can be roughly divided into four steps: excellent historical case matching, GMCA-Former-based process sensing, NSGA-II-based operation variable seeking, and event-triggered-based process model updating.

(1) Excellent historical case matching

In the process of system running, many historical operation cases have been accumulated, among which there are some excellent cases with outstanding control, which can provide reference experience for current operations. Therefore, matching cases are first used to obtain similar historical excellent cases. The criteria for selecting excellent historical examples in the problem of this paper are as follows:

$$\begin{cases} \mathbf{y}_t\left(Na^+\right) \leq 1.4\% \\ \left|\mathbf{y}_t\left(K^+\right) / \mathbf{y}_t\left(SO_4^{2-}\right) - 1.14\right| \leq 1.14 \times 10\% \\ \mathbf{y}_t\left(K^+\right) \geq 26.25\% \end{cases} \tag{10}$$

where $\mathbf{y}_t\left(Na^+\right)$, $\mathbf{y}_t\left(K^+\right)$, and $\mathbf{y}_t\left(SO_4^{2-}\right)$ denote the content of sodium ions, potassium ions, and sulfate ions in the potassium mixture after washing, respectively. The specific constraint values are obtained based on the process mechanism and historical experience. Assuming that the excellent historical cases satisfying the criteria are $\mathbf{D}_h$ and the current entry condition and process state are $\mathbf{I}_t, \mathbf{S}_t$, respectively. Then, the matching process can be described as

$$\mathbf{W}_{sim} = Dis_{Euro}\left(\left[\mathbf{I}_t, \mathbf{S}_t\right], \mathbf{D}_h\left[\mathbf{I}, \mathbf{S}\right]\right) \tag{11}$$

$$\mathbf{D}_m = \mathbf{D}_h\left[top_k\left(\mathbf{W}_{sim}\right)\right] \tag{12}$$

$$\mathbf{O}_t^{init} = \text{softmax}\left(top_k\left(\mathbf{W}_{sim}\right)\right)\mathbf{D}_m\left[\mathbf{O}\right] \tag{13}$$

where $Dis_{Euro}(\bullet)$ denotes the European distance, $top_k(\bullet)$ denotes the selection of the smallest k, $\mathbf{D}_m$ and $\mathbf{O}_t^{init}$ denote the matched excellent historical cases and the obtained reference operation variables.

(2) GMCA-Former-based process sensing

Three GMCA-Former-based process aware models are represented as $g_{Na^+}(\bullet)$, $g_{K^+}(\bullet)$, and $g_{SO_4^{2-}}(\bullet)$. They are trained using all historical samples for predicting the sodium, potassium, and sulfate ion contents of the washed mixed potassium, respectively. The inputs of the model are set as the operational variables $\mathbf{O}_t$, the current inlet conditions $\mathbf{I}_t$, and the equipment state $\mathbf{S}_t$. If the constructed model is sufficiently accurate, the output of the model can be used to simulate the grade of the mixed potassium at the exit of the washing process when the operating variables are set. This approach is of great significance because the washed mixed potassium grade is obtained from the laboratory values with a huge time delay, which makes it difficult to apply in practice. The proposed GMCA-Former model demonstrated its efficacy in sensing the intricacies of the washing process due to its remarkable ability to capture and comprehend global working conditions.

(3) NSGA-II-based operation variable seeking

We assume that the grade of mixed potassium at the outlet of the washing process is simulated using the accurate GMCA-Former model. Then combined with the discussion of the whole process in Section II, the optimization problem can be formulated as

$$\min \begin{cases} g_{Na^+}\left(\mathbf{O}_t, \mathbf{I}_t, \mathbf{S}_t\right) \\ -g_{K^+}\left(\mathbf{O}_t, \mathbf{I}_t, \mathbf{S}_t\right) \\ \left|g_{K^+}\left(\mathbf{O}_t, \mathbf{I}_t, \mathbf{S}_t\right) / g_{SO_4^{2-}}\left(\mathbf{O}_t, \mathbf{I}_t, \mathbf{S}_t\right) - 1.14\right| \\ o_t^3 \end{cases} \tag{14}$$

where $\mathbf{O}_t = \left[o_t^1, o_t^2, o_t^3\right]$ denote three operational variables including 1# and 2# E2 mother liquor pump current and freshwater addition amount, its initial value is the $\mathbf{O}_t^{init}$ obtained by matching the excellent historical cases. It can be seen that this paper deals with a multi-objective optimization problem, so NSGA-II is chosen as the optimization-seeking algorithm. Moreover, due to the slow-varying characteristics of the process industry, the following constraints need to be added to the optimization search process

$$s.t.\begin{cases} o_t^1 - o_{t-1}^1 \le o_{t-1}^1 \times 2\% \\ o_t^2 - o_{t-1}^2 \le o_{t-1}^2 \times 2\% \\ o_t^3 - o_{t-1}^3 \le o_{t-1}^3 \times 30\% \end{cases} \quad (15)$$

where $o_t^1, o_t^2 \in [100, 850], o_t^3 \in [0, 80]$. In other words, it constrains the range of variation of the operating variables. The optimized operating variable is then used as the final operational variable setting for the actual process control.

(4) Event triggered-based process model updating

Due to the time-varying characteristics of the process, the established process model does not have permanent high accuracy, so the update of the process model also needs to be considered. In this paper, an event-triggered approach is used to update the process model. That is when the predicted value of the established model and the obtained true laboratory value disagree by more than $\boldsymbol{\tau}$, the newly collected samples are used to fine-tune the established model to maintain its high accuracy.

V. Experimental Results and Discussion

A. Performance validation of GMCA-Former

The proposed GMCA-Former is validated in this section using 4043 actual data collected from a Salt Lake Chemical industrial process. The first 3600 data are utilized for model training, and the last 443 for model testing. To illustrate the superiority of the proposed method, baseline methods such as PCR [21], VW-SAE [13], and LSTNet [22] are constructed under the same settings for comparison. All simulation experiments are performed on Python 3.7 with torch 1.8 and NVIDIA GeForce RTX 3060 Laptop GPU. To ensure the standardization of the comparison experiments, the root mean square error (RMSE) and the mean absolute error (MAE) are selected as evaluation metrics, which are calculated as

$$\text{RMSE} = \sqrt{\sum_{i=1}^{p} \left(\hat{y}_i - y_i\right)^2 \Big/ p} \quad (16)$$

$$\text{MAE} = \sum_{i=1}^{p} \left|\hat{y}_i - y_i\right| \Big/ p \quad (17)$$

where y_i and $\hat{y}_i$ represent the true value and prediction value, respectively. p is the total number of the test sample.

Since sodium ions, potassium ions, and sulfate ions in the exported mixed potassium of the washing process have strong correlations and their modeling processes are almost consistent, only the prediction of sulfate ions is adopted as an example in this paper to demonstrate the superiority of the proposed model. The optimal hyperparameters of all methods are obtained by trial-and-error technique. The experimental results obtained on the test dataset are shown in Table I. From the two evaluation metrics, it can be seen that the modeling performance of PCR is the worst. This is mainly because PCR is a linear model that is unable to model the complex nonlinear relationships in the washing process. While VW-SAE has improved performance compared to PCR benefiting from its nonlinear extraction capability, but still inferior. This is attributed to the fact that they are both static models that cannot capture the dynamic properties of the process. LSTNet adopts LSTM as the basic unit with stronger temporal feature extraction capability. Thus the performance of LSTNet is better than both the previous two. The results of these three also show that the dynamic properties of the washing process are more important compared to the nonlinear ones. The proposed GMCA-Former not only considers the long-time dependent properties of the process but also takes into account the multi-conditions properties of the process. Therefore, its performance is further improved compared to other methods.

TABLE I
COMPARISON RESULTS OF FOUR METHODS FOR PREDICTING THE SULFATE ION CONTENT IN THE MIXED POTASSIUM WASHING PROCESS

Method	RMSE	MAE
GMCA-Former	**0.2816**	**0.2286**
LSTNet	0.3664	0.2862
VW-SAE	0.3763	0.2966
PCR	0.3786	0.2954

To more visually illustrate the effectiveness of the proposed method, Fig. 5 further presents the prediction curves of all methods on the test dataset. It is clear from the figures that the proposed GMCA-Former has better performance in the peaks and valleys of the curve. This is because the peaks and valleys of the curve are mostly when the process conditions change, and GMCA-Former can precisely capture such changes well, making it perform much better than other methods in these areas.

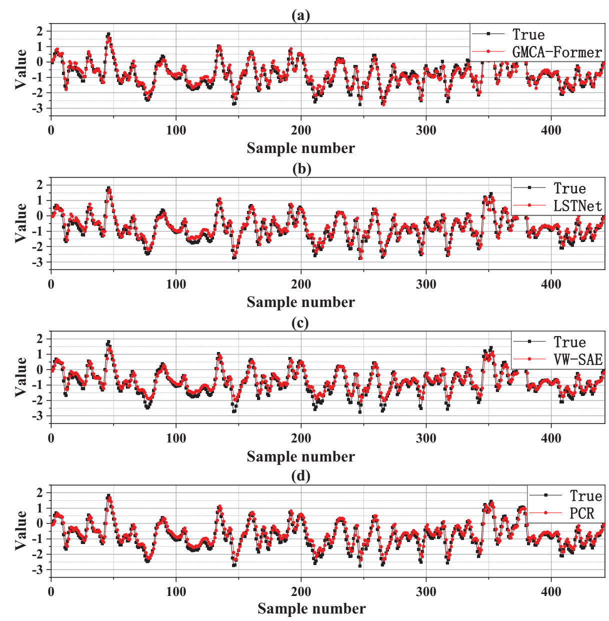

Fig. 5. Modeling performance of four methods for predicting the sulfate ion content in the washing process. (a) GMCA-Former. (b) LSTNet. (c) VW-SAE. (d) PCR.

From the above discussion, it can be concluded that the proposed GMCA-Former model has excellent modeling capability to accurately predict the grade of mixed potassium at the outlet of the washing process. Therefore, the GMCA-Former can be utilized to simulate the state of the washing process and to guide the subsequent operation optimization.

B. Performance validation of hybrid operation optimization strategy

In this paper, the proposed hybrid operation optimization strategy is further structured on the same dataset based on the established GMCA-Former. It is compared with manual adjustment, historical case matching, and NSGA-II-only-based operation optimization strategies. The corresponding evaluation metrics are the average value of four optimization objectives for the task in the test dataset. The importance of four optimization objectives according to the process requirements is sodium ion content, potassium ion to sodium ion content ratio, potassium content, and freshwater addition. In addition, since several algorithm-based methods could not obtain the mixed potassium grade obtained after the operating variable values are sent down, the performance of all methods except the manual adjustment method is replaced by the predicted values of the constructed GMCA-Former model. The results obtained after several experiments on the test dataset are shown in Table II.

From the results in the table, it can be seen that the proposed hybrid operation optimization strategy achieves optimal values for most of the objectives. This also proves the superiority of the proposed method. In contrast, the case where only NSGA-II or case-matching is used has difficulty coordinating multiple objectives to reach the optimum. Even 10% of the samples cannot be solved, which further demonstrates the efficiency of the proposed strategy.

TABLE II
COMPARISON RESULTS OF FOUR METHODS FOR OPERATION OPTIMIZATION OF THE MIXED POTASSIUM WASHING PROCESS

Method	$Na^+ \downarrow$	K^+/SO_4^{2-}	$K^+ \uparrow$	$O_3 \downarrow$
Proposed	0.9981	1.0251	**27.5781**	**9.3245**
NSGA-II (44 samples failed)	1.0028	1.0131	27.4836	12.5053
Case matching	1.0321	1.0022	27.3156	13.4028
Manual	**0.9905**	**1.0828**	26.9621	14.8149

VI. CONCLUSION

This paper proposes a novel GMCA-Former network to consider the interactions between conditions and the decreasing impact properties with time intervals for multiple conditions in the industrial processes. The experimental results provide evidence that the proposed method exhibits superior modeling performance, particularly in instances where the modeling index curve reaches its peak or trough. Furthermore, based on GMCA-Former to build a relatively accurate process model, a hybrid operational optimization strategy is proposed in this paper. Specifically, the weighted values of operational variables obtained from the excellent case matching are utilized as the initial values of NSGA-II, which helps the algorithm to find better values. Experimental results show that the proposed optimization method performs better in optimizing operating variables than existing methods.

Even with the advancements achieved by the proposed GMCA-Former method, it remains vulnerable to high computational complexity in the ultra-long time range feature extraction, indicating the need for future research to improve its computational efficiency.

REFERENCES

[1] Y. S. Perera, D. A. A. C. Ratnaweera, C. H. Dasanayaka, *et al.*, "The role of artificial intelligence-driven soft sensors in advanced sustainable process industries: A critical review," *Engineering Applications of Artificial Intelligence*, vol. 121, p. 105988, 2023.

[2] F. Qian, "The Future of Smart Process Manufacturing," *Engineering*, vol. 22, pp. 20-22, 2023.

[3] Y. Fang, Z. Jiang, D. Pan, *et al.*, "Soft Sensors Based on Adaptive Stacked Polymorphic Model for Silicon Content Prediction in Ironmaking Process," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, p. 2503412, 2021.

[4] D. Liu, Y. Wang, C. Liu, *et al.*, "Data Mode Related Interpretable Transformer Network for Predictive Modeling and Key Sample Analysis in Industrial Processes," *IEEE Transactions on Industrial Informatics*, early access, 2022, doi: 10.1109/TII.2022.3227731.

[5] P. Petsagkourakis, I. O. Sandoval, E. Bradford, *et al.*, "Chance constrained policy optimization for process control and optimization," *Journal of Process Control*, vol. 111, pp. 35-45, 2022.

[6] C. Liu, K. Wang, Y. Wang, *et al.*, "Learning deep multimanifold structure feature representation for quality prediction with an industrial application," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 9, pp. 5849-5858, 2022.

[7] Y. Luo, X. Zhang, M. Kano, *et al.*, "Data-driven soft sensors in blast furnace ironmaking: a survey," *Frontiers of Information Technology & Electronic Engineering*, vol. 24, no. 3, pp. 327-354, 2023.

[8] Y. Wang, J. Luo, C. Liu, *et al.*, "Layer-wise residual-guided feature learning with deep learning networks for industrial quality prediction," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, p. 2520011, 2022.

[9] C. Liu, Y. Wang, K. Wang, *et al.*, "Deep learning with nonlocal and local structure preserving stacked autoencoder for soft sensor in industrial processes," *Engineering Applications of Artificial Intelligence*, vol. 104, p. 104341, 2021.

[10] M. R. Parvez, W. Hu, and T. Chen, "Real-time pattern matching and ranking for early prediction of industrial alarm floods," *Control Engineering Practice*, vol. 120, p. 105004, 2022.

[11] C. Liu, K. Wang, Y. Wang, *et al.*, "Deep Nonlinear Dynamic Feature Extraction for Quality Prediction Based on Spatiotemporal Neighborhood Preserving SAE," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, p. 2516310, 2021.

[12] Y. Liu, C. Yang, K. Huang, et al., "A Systematic Procurement Supply Chain Optimization Technique Based on Industrial Internet of Things and Application," *IEEE Internet of Things Journal*, vol. 10, no. 8, pp. 7272-7292, 2023.

[13] X. Yuan, B. Huang, Y. Wang, *et al.*, "Deep Learning-Based Feature Representation and Its Application for Soft Sensor Modeling With Variable-Wise Weighted SAE," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3235-3243, 2018.

[14] X. Jiang, X. Kong, and Z. Ge, "Augmented Industrial Data-Driven Modeling Under the Curse of Dimensionality," *IEEE/CAA Journal of Automatica Sinica*, vol. 10, no. 6, pp. 1-17, 2023.

[15] C. Liu, K. Wang, L. Ye, *et al.*, "Deep learning with neighborhood preserving embedding regularization and its application for soft sensor in an industrial hydrocracking process," *Information Sciences*, vol. 567, pp. 42-57, 2021.

[16] J. Ding, C. Yang, and T. Chai, "Recent Progress on Data-Based Optimization for Mineral Processing Plants," *Engineering*, vol. 3, no. 2, pp. 183-187, 2017.

[17] Q. Chen, J. Ding, T. Chai, *et al.*, "Evolutionary Optimization Under Uncertainty: The Strategies to Handle Varied Constraints for Fluid Catalytic Cracking Operation," *IEEE Transactions on Cybernetics*, vol. 52, no. 4, pp. 2249-2262, 2022.

[18] A. Alabdulkarem and N. Rahmanian, "Steam consumption minimization using genetic algorithm optimization method: an industrial case study," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, pp. 1-15, 2020.

[19] Y. Bi, L. Yin, T. He, *et al.*, "Optimization and analysis of a novel hydrogen liquefaction process for circulating hydrogen refrigeration," *International Journal of Hydrogen Energy*, vol. 47, no. 1, pp. 348-364, 2022.

[20] A. Vaswani, N. Shazeer, N. Parmar, *et al.*, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, no. 2017.

[21] X. Yuan, Z. Ge, B. Huang, *et al.*, "Semisupervised JITL Framework for Nonlinear Industrial Soft Sensing Based on Locally Semisupervised Weighted PCR," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 532-541, 2017.

[22] G. Lai, W.-C. Chang, Y. Yang, *et al.*, "Modeling Long- and Short-Term Temporal Patterns with Deep Neural Networks," in *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, Ann Arbor, MI, USA, 2018.