BIG DATA MINING AND ANALYTICS
ISSN 2096-0654
Volume 5, Number
DOI: 10.26599/BDMA.2023.9020002

Artificial Intelligence Methods Applied to Catalytic Cracking Processes

Fan Yang, Mao Xu, Wenqiang Lei, and Jiancheng Lv*

Abstract: Fluidic Catalytic Cracking (FCC) is a complex petrochemical process affected by many highly non-linear and interrelated factors. Product yield analysis, flue gas desulfurization prediction, and abnormal condition warning are several key research directions in FCC. This paper will sort out the relevant research results of the existing Artificial Intelligence (AI) algorithms applied to the analysis and optimization of catalytic cracking processes, with a view to providing help for the follow-up research. Compared with the traditional mathematical mechanism method, the AI method can effectively solve the difficulties in FCC process modeling, such as high-dimensional, nonlinear, strong correlation, and large delay. AI methods applied in product yield analysis build models based on massive data. By fitting the functional relationship between operating variables and products, the excessive simplification of mechanism model can be avoided, resulting in high model accuracy. AI methods applied in flue gas desulfurization can be usually divided into two stages: modeling and optimization. In the modeling stage, data-driven methods are often used to build the system model or rule base; In the optimization stage, heuristic search or reinforcement learning methods can be applied to find the optimal operating parameters based on the constructed model or rule base. AI methods, including data-driven and knowledge-driven algorithms, are widely used in the abnormal condition warning. Knowledge-driven methods have advantages in interpretability and generalization, but disadvantages in construction difficulty and prediction recall. While the data-driven methods are just the opposite. Thus, some studies combine these two methods to obtain better results.

Key words: intelligent optimization algorithm; neural networks; catalytic cracking; lumped kinetics

1 Introduction

With the aggravating trend of heavier and deteriorated of the crude oil, Fluid Catalytic Cracking (FCC), one of the key processes for processing heavy oil, has attracted more and more attention. FCC is a series of chemical reactions dominated by cracking reactions of heavy oil that occur at about 500 ℃ and 1×105∼3×105 Pa in the presence of acid catalysts. The reactions mainly produce light oil, gas and coke. In China, diesel and gasoline produced by Fluid Catalytic Cracking Units (FCCU) account for about 30% and 70% of the total diesel and gasoline product, respectively. FCC has become one of the most important methods for heavy oil processing[1–4].

FCC is one of the most complicated processes in process industry. At present, the methods for process analysis of FCCU can be roughly divided into mathematical model based methods and artificial intelligence based methods[5]. Molecular scale kinetic mode[6] and lumped kinetic model[7] are typical mathematical model based methods. They have been widely studied in the field of FCC process analysis.

• Fan Yang is with the College of Computer Science, Sichuan University, Chengdu 610041, China, and also with the Algorithm and Big Data Center, New Hope Liuhe Co Ltd, Sichuan 610000, China. E-mail: yang.fan@live.com.
• Mao Xu is with the Data Intelligence Lab, New Hope Liuhe Co Ltd, Sichuan 610000, China. E-mail: xumao3021@163.com.
• Wenqiang Lei and Jiancheng Lv are with the College of Computer Science, Sichuan University, Chengdu 610041, China. E-mail: {wenqianglei, lvjiancheng}@scu.edu.cn.
* To whom correspondence should be addressed.
Manuscript received: 2022-10-12; revised: 2022-12-30; accepted: 2023-03-09

© The author(s) 2023. The articles published in this open access journal are distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).

On the basis of the description of the process principle and its physical and chemical processes, these methods calculate the change trend of the target to be analyzed by constructing a differential equation model. Restricted by the complexity of modeling, this kind of methods can only use a small number of process parameters, leading to low accuracy, so that there are large errors in practical application.

With the increasing automation control level of the industrial production process and the continuous improvement of the process control system, the recorded data and operating condition parameters of the process equipment can be obtained in real time from the database platform of the unit[8]. These data record the characteristics and performance of the FCCU, reflecting the essence of the production system, which provide a foundation for the application of Artificial Intelligence (AI) in FCC[9]. In recent years, AI-based methods have shown strong advantages. In addition to the Internet[10,11], they have also played an important role in education[12], transportation[13], medical care[14] and smart cities[15]. Also, more and more AI methods have been applied to the analysis of FCC process[5,16]. The FCC analysis method based on AI will become a research focus.

High efficiency, environmental sustainability, and safety are the three core objectives in the process industry[17]. In the FCC, product analysis and optimization, Flue Gas Desulfurization (FGD) analysis and optimization, and early warning and diagnosis of abnormal condition are the research emphases and hotspots in the direction of high efficiency, environmental protection, and safety. In this paper, we provide a review of the application of AI method in the FCC analysis from three aspects: product analysis and optimization, FGD analysis and optimization, and early warning and diagnosis of abnormal condition, hoping to provide possible help for future research.

2 FCC Product Analysis and Optimization

FCC is a complex process influenced by many highly nonlinear and strongly interrelated factors. In this process, many factors, including the nature of feed oil, the nature of reaction regeneration catalyst, and the conditions of reaction, will affect the quality and yield of product. The mathematical modeling analysis of the process oriented to the product quality or yield of the catalytic cracking unit has always been a focus and challenge in the field of petroleum processing[18,19].

The traditional methods of FCC analysis are mainly based on mathematical model[20,21]. That is, the process principle and its physical and chemical processes are described by mathematical formulas, which can effectively reflect the transfer and reaction laws of the process. Mathematical model based methods have the characteristics of clear engineering background, strong interpretability, and strong traceability, mainly including analysis methods, such as correlation mode[22], lumped kinetic model[23], and molecular scale dynamics model[24]. Since the raw materials and product of the FCC are complex mixtures composed of a large number of hydrocarbons and non-hydrocarbons, the process involves a large number of complex reaction systems. Considering the complexity of the model establishment and its practicality in industry, the lumped kinetic model is the most commonly used method for mechanism analysis.

Mathematical model based methods often need to make idealized and simplified assumptions, ignoring some secondary factors selectively, which will lead to the loss of model accuracy. And the accumulation of the simplification of a single reaction unit is easy to cause the error to be amplified step by step, which makes it difficult to guarantee the convergence and stability of the whole system model. In addition, due to the poor timeliness and long cycle of lumped modeling, it is impossible to update and analyze the process status in real time[25].

With the development of advanced sensor and database technology, a large number of production process data can be collected into the real-time database. These data recording the characteristics, performance, and changes of the chemical process, are a comprehensive and detailed description of the chemical process, which could provide favorable conditions for the data-driven methods. Data-driven methods represented by AI methods directly build models based on massive data. With the development of AI and the proposal of “Industry 4.0”, more and more AI technologies are introduced into the modern industry chain to improve production efficiency, reduce operation cost, improve operation safety, and realize risk avoidance. Meanwhile, Deep Learning (DL), as an important technology of AI, has made amazing progress in theoretical and applied research in the industry, which vigorously promotes the development of informatization, digitization, and intelligence of the modern industry[26]. By fitting the functional relationship between operating variables and product, these methods

can analyze the reaction process and its influence mechanism from multiple angles in an all-round way, avoiding the over simplification of a large number of factors in the mechanical modeling and finding the relationship between the input and output of system process when the process mechanism is unknown or too complex. Appropriate machine learning methods can effectively solve the problems of high dimension in chemical process, strong correlation of influence factors, nonlinearity, time-varying, lagging, and uncertainty, thereby obtaining high-precision prediction results, which have shown great advantages in chemical process modeling[27–29].

2.1 Machine learning approaches

Statistical learning based methods are called traditional machine learning methods. Multiple linear regression is one of the simplest traditional machine learning methods. It uses multiple linear functions as hypothetical functions and solves the coefficients of hypothetical functions by least squares. Multiple regression is widely used in modeling simple chemical processes. Wei et al.[30] used multiple linear regression to accurately predict the quality of chemical product. Lv et al.[31] used multiple linear regression to predict chemical product. Gmeinbauer et al.[32] used multiple linear regression to predict the product distribution of gasoline and liquefied gas in FCCU.

Compared with the multiple linear regression, Bayesian regression can improve the accuracy and generalization of chemical process models by introducing appropriate prior distribution[33]. Support Vector Regression (SVR) finds a hyperplane that minimizes the maximum distance to all samples under the constraint that the error between the regression prediction value and the sample annotation value is small enough. Sun et al.[34] used SVR to model catalytic cracking product, and found a series of optimized conditions that can improve the yield. Roy et al.[35] used multiple regression and SVR separately to predict methane content in natural gas.

FCC is a complex chemical process with highly nonlinear characteristics. Therefore, processing FCC directly with a linear model always gets poor results, while nonlinear models tend to get better results. On the one hand, a linear model can be transformed into a nonlinear model by introducing a kernel function. Roy et al.[35] introduced a polynomial kernel into SVR, which makes the prediction accuracy increase from 40% to 52%. And after using Gaussian kernel, the accuracy reaches 98%. The use of kernel functions based on dimension reduction, such as kernel PCA[35] and kernel Partial Least Squares (PLS)[36], can bring significant performance improvement in chemical process modeling. On the other hand, nonlinear machine learning models can be directly used to fit the relationship between various influencing factors and production results, such as polynomial regression[36], C4.5 decision tree[37], random forest[38], and Gradient Boosted Decision Tree (GBDT)[39], which have been applied in the chemical process modeling.

The process parameters of FCCU are usually strongly correlated with each other, so feature engineering is necessary. The common feature selection methods include filter, wrapper, and embedding. The filter-based feature selection method aims to calculate the importance of features to rank them. The top-ranked feature variables are usually high-importance features, while the bottom-ranked feature variables are irrelevant or less importance features. The filter method selects features whose importance is greater than a specified threshold as input variables or selects the top-k features with the greatest importance (k is a manually set super parameter). The wrapper-based feature selection method takes the performance of machine learning model as the criterion for evaluating the feature subset, and obtains the target feature subset through continuous iteration of the search algorithm. Unlike filter and wrapper, the embedding method integrates feature selection into the model training of the machine learning algorithm, that is, the model automatically selects features during the model training process of the machine learning algorithm. Correlation analysis is a popular method to analyze the importance of features in filtering feature selection. Zhao[40] combined Pearson correlation coefficient analysis in SPSS software and process production experience to filter all variables of raw oil. In FCCU, regenerant and reaction regeneration systems are based on the data obtained from data preprocessing, reducing the complexity of machine learning models. Wang et al.[41] combined the filter method with the wrapper method to select features. It is a data-driven spontaneous feature variable selection method that does not rely on FCC prior knowledge in the process of selecting input variables. This method used the production data of the FCCU to select the

input variables for the prediction model of dry gas and coke yield, and proposed a model with high prediction accuracy and moderate number of input variables.

Embedded selecting of features is included in machine learning methods, such as GBDT and random forest. Wang et al.[42] grouped the features combined with the expert experience and knowledge of FCC on the basis of GBDT, and proposed an adaptive feature selecting method, which effectively improved the generalization ability of FCCU gasoline yield prediction.

The machine learning based method has been widely used in the analysis of FCC products, which can effectively model the FCC process, thereby realizing the analysis of FCC product yield.

2.2 Neural networks approaches

Compared with machine learning models, neural networks are more widely used in the modeling of FCC process because of their powerful fitting ability and flexible structure. Lv et al.[43] built a neural network model to predict the product distribution of FCC and optimized the reaction-regeneration system to obtain conditions with better product yields. Wang[44] proposed a prediction model for hydrocracking product yield by using Back Propagation (BP) neural networks, and under the guidance of the model prediction results, the process parameters were optimized and adjusted according to the actual working conditions. The results show that the total yield of high-value product has been greatly improved and greater economic benefits have been obtained after the optimization and adjustment of the process parameters.

Jiang et al.[45] established a neural network prediction model with 17 input variables based on the Generalized Regression Neural Networks (GRNNs) and the adaptive enhancement algorithm. The results show that the mean squared error between the predicted gasoline production and actual production was at a low level. Shang et al.[46] used the deep neural network to predict the 95% cut-off point of heavy diesel oil in atmospheric and vacuum distillation unit. The results show that the prediction results are closer to the actual values compared with other data-driven methods.

FCC process analysis is a time series problem, which is suitable for solving with neural network models dealing with time series. Long Short-Term Memory (LSTM) is a recurrent neural network architecture to deal with time series problems. It introduces gating mechanisms to better learn the dependencies in time series data[47].

Ke et al.[48] proposed an LSTM-based deep neural network to deal with the strong nonlinearity and dynamics of chemical processes. The validity of the model is verified by the benchmark test of the sulfur recovery unit. Finally, the method was applied to a practical soft sensing system, and the results show that the method is particularly suitable for the modeling of dynamic soft sensing system for coal gasification.

Xu et al.[49] extracted the time feature of feed oil properties, catalyst properties, and process parameters of catalytic cracking unit by using multi-layer Bidirectional LSTM (Bi-LSTM), which predicted the production effectively. And an automatic encoding-decoding method based on the cyclic structure was proposed to extract valuable features from the data and reduce the input dimension, and at the same time can effectively reduce the noise of sensor data acquisition[50].

Neural networks for complex chemical processes are often high-dimensional nonlinear models. The nonlinear activation function is introduced to enhance the fitting ability of the neural networks, but it also makes the loss function non-convex. For convex optimization problems, the optimization algorithm can always find the optimal solution. However, for convex functions in high-dimensional space, there is little guarantee that the function will converge to the global optimum as the gradient close to 0. Even if the loss function only converges to a local optimal value, the obtained neural network error is often small. But in high-dimensional space, it tends to fall into a saddle point when the gradient of loss function is close to 0, resulting in a large error[51].

Saddle-Free Newton method (SFN)[51] is proposed to fits the loss function in the Hessian matrix. An eigenvalue less than 0 is found, and the saddle point can be separated along the direction of the corresponding eigenvector to continue to optimize the loss function. Because the calculation of Hessian matrix is very large, SFN is rarely used to optimize the neural networks of chemical process. Some scholars have used the Levenberg-Marquardt (LM) training algorithm with good results, because the LM uses a Jacobian matrix which is simpler than Hessian, and the calculation is simpler. Dasila et al.[52] used the artificial neural networks, took the conventional properties, such as density, distillation temperature, Conradson Carbon Residue (CCR), sulfur, and total nitrogen content, as inputs, and adopted the LM algorithm to study several

BP neural networks with different neuron numbers. By using the ten-lumped kinetic model of the FCCU, the performance of several different feedstocks and prediction of the detailed composition of the FCC feedstock have been successfully simulated.

Using stochastic gradient descent to find the minimum value of the loss function is very sensitive to the initial value of the parameters, and inappropriate initial value will often cause the optimization result to fall into a saddle point[53]. Applying stochastic optimization algorithms, such as Genetic Algorithm (GA), Particle Swarm Optimization algorithm (PSO), and simulated annealing algorithm to the initial value setting of neural networks parameters, can efficiently avoid saddle points.

In the process of using the GA to optimize the BP neural networks (GA-BP), the neural network is the first to be determined, the network parameters are initialized, and the initial values of the neural network parameters are encoded by the GA. Second, the optimal values of the network parameters are found through operations, such as selection, crossover, and mutation, until the constraints are satisfied. Finally, the neural network parameters are updated according to the optimal values of the network parameters found by the GA. Figure 1 shows the algorithmic flow diagram of the GA-BP[54].

The team of professor Ouyang from East China University of Science and Technology has carried out a lot of work in analyzing FCC process by using BP neural networks combined with GA[55–57]. Fang[58] selected 19 variables, including feed oil properties, catalyst properties, and conditions, as the neural network inputs, and the yields of liquefied gas, gasoline, diesel oil, and coke as the neural network outputs. They built a GA-BP to obtain the conditions of the reaction-regeneration system under the optimal gasoline yield. Zhao[40] proposed a GA-BP for gasoline yield on all the data of the MIP device and the data after clustering, and finally optimized the gasoline yield model. Su et al.[59] compared the prediction results of BP neural networks and GA-BP with industrial data separately, showing that the GA-optimized prediction model has better results in terms of accuracy and stability. In addition, the accuracy of the GA-BP was further proved by investigating the influence of single key parameters, such as raw material carbon residue and reaction temperature on coke yield. Wang et al.[60] proposed a method referred FNN-GA that combines the Fuzzy Neural Network (FNN) with GA to correlate input values, i.e, raw material components and operating variables, and output values, i.e., the production of upgraded gasoline and olefin components therein. And GA was used to optimize the input of operating variables to maximize the olefin limited gasoline of different feedstocks. The experimental results are in good agreement with the predicted results. The optimized operating conditions significantly improved the gasoline yield. It can be seen that the GA directly operates the structural objects through operations, such as selection, crossover, and mutation, adaptively adjusting the search direction of parameters, resulting in good global convergence. When GA is deeply combined with neural networks, more reasonable parameter values can be searched for the initialization of neural network parameters, and then the learning ability of the model can be optimized, thereby greatly improving the stability and accuracy of the neural network.

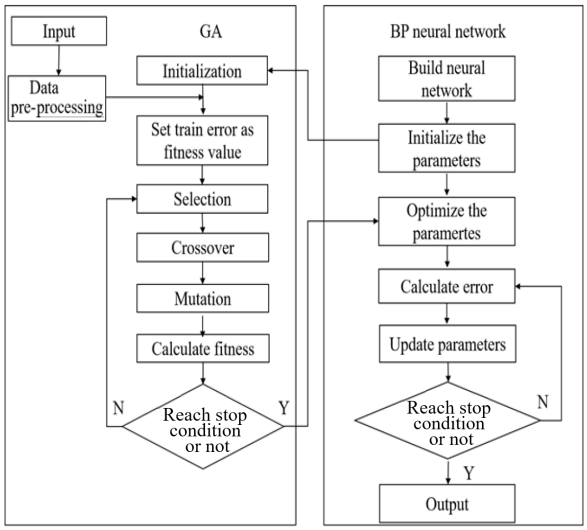


Fig. 1 Abstract illustration of GA-BP.

Neural networks combined with PSO also show good applicability. Gao et al.[61] constructed a BP neural network for FCC regeneration, and used PSO to optimize the initial weight and threshold of the neural network. Compared with the model without optimization, the prediction accuracy of PSO-BP has been greatly improved. Shang et al.[62] applied the Particle Swarm Optimization with Pre-Crossover (PSOPC) algorithm to the soft sensing modeling of C3 content in dry gas of FCCU and found that the PSOPC model has higher accuracy and better generalization. Wang et al.[63] constructed the Principal Component Analysis (PCA) based neural networks and PSO-BP model, and analyzed and compared their simulation results, which showed that the performance of the PSO-BP is better than that

of the PCA-based neural network. The PCA-based soft sensing model can better reflect the change trend and value of the total hydrocarbon content above C3 in dry gas, meeting the requirements of the actual production process.

In addition, some researchers have explored the combination of neural networks and simulated annealing algorithms. Tang[64] proposed a gasoline yield prediction model of MIP unit based on the generalized regression neural networks and AdaBoost algorithm. And they optimized the gasoline yield prediction model by using the simulated annealing algorithm in the individual behavior based optimization algorithm and the GA in the group behavior based optimization algorithm, founding that both algorithms can obtain the optimal gasoline yield. However, the simulated annealing algorithm fell into local optimization and did not get the optimal value at each optimization, indicating that the algorithm has poor stability. They finally chose the modified GA as the optimization algorithm to optimize gasoline yield.

Compared with the machine learning based analysis methods of FCC, neural networks based methods are more suitable for chemical engineering modeling. First, neural networks have a complex multi-layer structure, which can contain richer information and has strong feature extraction abilities. Second, neural networks, as a latent variable model, can help describe highly correlated process variables. At the same time, the massive data continuously collected in the chemical production process just meet the data requirements of the neural network model, providing data support for the construction of a complex network structure.

2.3 Data-driven approaches with mathematical mechanistic models

The data-driven-based methods obtain the hidden information and laws by analyzing a large number of historical data in FCC process, and then determine the mapping relationship among the factors, such as raw material properties, catalyst properties, operating process conditions, and production results. However, such methods cannot clearly describe the transfer and reaction process, resulting in poor interpretability. Moreover, these methods are completely data-driven, so the calculation results often depend heavily on the number and quality of data. It is easy to overfit the environmental noise, resulting in poor prediction generalization, which makes it difficult to analyze and explain the process mechanism at a deep level.

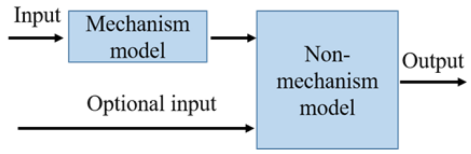
At the same time, there is a "survivor bias" when using the historical data of the actual FCC production process. All operating condition values are only distributed within the range controlled by the experience of the production line workers in the past, and the production results are also guided by the profit goal of the production line, Therefore, the sample space is too small to indicate a complete data distribution law.

According to the characteristics of highly nonlinear and strong correlation, as well as influencing factors during the FCC process, the mathematical mechanism models represented by lumped dynamics is combined with the data-driven-based AI methods to construct a mechanism-data hybrid driven analysis model, which can make full use of the existing prior knowledge, mine the effective information in the data, and improve the efficiency and accuracy of modeling. And then the FCC process is optimized to further improve the model's ability to predict product distribution.

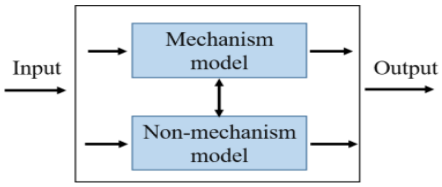
According to the connection mechanism of models, hybrid modeling can be divided into series, parallel, and hybrid. Series means that the input variables first enter the mechanism model for operation, the output of the mechanism model is then used as the input of the non-mechanism model, and the output of the non-mechanism model is used as the final output. Parallel means that the mechanism model and the non-mechanism model are used to perform parallel calculation on the input data. When the error of the output data obtained by using the mechanism model alone is large, the appropriate non-mechanism model can be used for error learning, and the total output is the sum of the results of the mechanism model and the non-mechanism model, so that the obtained output error will be reduced. Hybrid modeling can integrate several mechanism models and several non-mechanism models at the same time, which has a large degree of freedom and is very suitable for complex petrochemical production process modeling. The schematic diagram of series and parallel modeling is shown in Fig. 2[65].

Yang et al.[5] used the calculation results of the lumped kinetic mechanism model of FCC as the input of neural networks, and the final performance was significantly higher than that of using only the feedstock properties as the input features of the neural network.

Bollas et al.[66,67] proposed a hybrid model combining FCC mechanism and neural networks, and found that the hybrid model can improve the prediction accuracy better than the simple mechanism model and the simple neural



(a) Parallel modeling



(b) Series modeling

Fig. 2 Schematic of the hybrid model construction method.

networks model by comparing with the industrial data from Greek refineries. Liu[68] took the MIP process as the research object, constructed an eight-lumped reaction network and calculated the product distribution. The input layer of the BP neural network of 14 variables, including the main raw material properties, catalyst properties, and conditions, are selected, and 5 variables, including the error between the predicted value and the industrial actual value of the yield of diesel, gasoline, liquefied gas, dry gas, and coke calculated by the lumped kinetic model are used as the output of the BP neural network. They constructed a 14-7-5 BP neural network hybrid model, in which a total of 14 variables, including main raw material properties, catalyst properties, and operating conditions, were selected as the BP neural network input, and a total of 5 variables, i.e., errors between predicted and industrial yields of diesel, gasoline, liquefied gas, dry gas, and coke yields calculated by lumped kinetic models, were as the BP neural network output. And they combined the error values obtained by the hybrid model with the prediction value of the product distribution by the lumped model to obtain the final product prediction value, which is closer to the industrial measured value. It can be seen from the above studies that the hybrid models combining the mechanism-driven lumped kinetic model and the data-driven neural networks take into account the advantages of the two methods, having a good effect on further improving the accuracy and precision of the FCC product prediction.

3 FGD Prediction

Sulfur dioxide is one of the main air pollutions caused by industrial production, mainly from the processing and use of the fossil fuels[69]. Acid rain formed by sulfur dioxide can cause direct harm to the local environment, damaging crops, forests, and human health. In China, the sulfur oxides emitted by FCCU account for about 5% of the total emissions. FCCU also contains a large amount of nitrogen oxides, dust particles, etc., and the treatment of its emissions is receiving increasing attention[70].

Wet Flue Gas Desulfurization (WFGD), using the alkaline absorbent solution to remove SO_2 from flue gas, is currently the most important method for the Flue Gas Desulfurization (FGD) with low cost and high efficiency[69]. The waste gas enters from the inlet at the bottom of the WFGD unit. The alkaline absorbent solution is introduced through the nozzles of three to five spray stages at the upper part of the tower. After mass transfer and chemical reaction, SO_2 is absorbed into the liquid phase. The more alkaline absorbents are used, the more SO_2 is absorbed and the greater the power consumption is.

The AI method for modeling and optimization of WFGD desulfurization has received more and more research to take into account both the environmental protection requirements and economic benefits. This kind of research usually divides modeling and optimization into two independent stages. In the first stage, the model- or rule- based of the system is established according to the process mechanism or data. The second stage is to find the optimal operating parameters based on the built model or rule base.

3.1 Modeling of WFGD

The main task of WFGD modeling is to find out the relationship between desulfurization efficiency and various operating parameters. Given the input SO_2 concentration and WFGD process parameters φ, the residual SO_2 concentration C_{out} after desulfurization can be predicted,

$$c_{\text{out}} = f_{\text{WFGD}}\left(c_{\text{in}}, \varphi\right) \tag{1}$$

WFGD modeling methods can be divided into two categories: mathematical model based methods and data driven methods. Mathematical model based WFGD modeling methods can be further divided into micro mass transfer theory based methods[71], reaction kinetics based methods[72], etc. The calculation results are in good agreement with experiments and industrial practice.

On the one hand, the input concentration of SO_2 can be directly detected. On the other hand, SO_2 can be predicted as a product of FCCU using a product analysis method similar to that mentioned in Section 2. Wu et

al.[73] combined Convolutional Neural Network (CNN) and LSTM to accurately predict the exhaust gas emission of FCCU, where the global time cumulative influencing factors can be captured by LSTM and the correlation among the influencing factors within a time window is captured by CNN.

Similar to the product analysis of FCC, mathematical model based WFGD modeling methods are often not accurate enough due to their simplified assumptions. The data-driven methods can fit more complex details and have more accurate prediction results. Reference [74] proposed a data clustering method to obtain the continuous optimal operation mode, where the guidance of operation is provided by constraining the parameter vector to continuously approach the optimal operation mode. Some studies[75,76] used neural networks to predict the concentration of SO_2 at the outlet under different operating parameters, which achieved good results.

Because data-driven methods usually have the defects of interpretability and generalization, some scholars combine data-driven methods with mathematical model based methods to improve performance. According to Ref. [77], the economic performance mathematical models of SO_2 removal efficiency and alkaline absorbent consumption, power consumption, and process water consumption were derived, and then the undetermined coefficients were obtained by data-driven regression. Reference [75] proposed that the residual error of the prediction results based on the mathematical model should be compensated by the artificial neural network to improve the prediction accuracy. The mathematical model for calculating residual error of the prediction results f_{MAT} is shown in the following:

$$f_{\mathrm{MAT}} = c_{\mathrm{in}} \cdot \exp\left(-K \cdot pH^{a_1} \cdot \left(N_{\mathrm{pump}}/Load\right)^{a_2}\right) \quad (2)$$

where c_{in} is the concentration of input SO_2, pH^{a_1} is the PH value of alkaline absorbent, N_{pump} is the number of circulating pumps, $Load$ is the load of circulating pumps, and K, a_1, and a_2 are undetermined coefficients, which are calculated by data-driven regression method.

On the basis of the above mathematical model, a fully connected neural network is added to compensate for the residual error of the mathematical model prediction results. There are 9 input units in its input layer, including 9 process parameters: gas flow, PM concentration, temperature, alkaline absorbent level, alkaline absorbent density, input SO_2 concentration, PH value of alkaline absorbent, number and load of circulating pumps. Compared with using mathematical model alone or fully connected neural network, the combination of mathematical model and data-driven method has better performance, which indicates that the comprehensive input variables enable the data-driven model to compensate for neglected factors and eliminate the errors of the mathematical model more effectively.

3.2 Optimization of FGD

The main task in the optimization stage of operating parameters is to study the operating parameters with the best economic performance (lowest cost) on the premise of meeting the FGD requirements (SO_2 residue does not exceed the environmental protection standard). In industrial modeling, if a significant analytical mathematical model is built, mathematical programming can be used to find the optimal operating parameters, such as linear programming[78], nonlinear programming[79], and dynamic programming[80]. Reference [77] used multi-objective programming to solve the optimal solutions of slurry PH value, calcium sulfur molar ratio, and liquid gas molar ratio after obtaining the model of the relationship between output SO_2 concentration and alkaline solvent consumption, power consumption, and water consumption, which takes into account the economic benefits while protecting the environment.

Data-driven models that often do not have an explicit analytical form typically use metaheuristic algorithms to find optimal operating parameters. Reference [77] used particle swarm optimization algorithm based on its hybrid method. Reference [75] used GA to find the optimal operation method to reduce the operation cost. Reference [81] determined the flow of alkaline absorbent and the number of circulating pumps (or the current corresponding to the number of circulating pumps) through various meta-heuristic methods, such as GA, simulated annealing, and particle swarm optimization algorithm, based on the constructed neural networks for predicting SO_2 concentration.

In recent years, some Reinforcement Learning (RL) methods have been applied to the optimization of operation parameters in process industry. Reference [82] proposed a model-based method whose performance is directly affected by the modeling accuracy, using an on-policy RL method to plan the operating parameters of wastewater treatment pumps after building a model with gradient boosting trees and multi-linear regression. For WFGD unit, Ref. [83] proposed a model-free off-

policy RL method, the core idea of which is to record the state-action value through double Q value neural networks with dueling architectures, and then obtain each step optimal operation parameters through the greedy strategy.

The observed state at the K-th step is determined by the power load prediction value $L_{k+1}$and sulfur content S_{k+1} at the next $(K+1)$-th step. The observed state at the K-th step could be written as

$$s_k = \left[L_{k+1}, S_{k+1}, M_k^{(1)}, \ldots, M_k^{(n)}, T_k^{(1)}, \ldots, T_k^{(n)}\right] \tag{3}$$

where $M_k^{(i)}$ and $T_k^{(i)}$ are the state of the i-th circulating pump at time k and the duration of this state, respectively. The return at moment k is defined as

$$r_k = w \cdot f_s\left(T_{k+1}\right) + (1-w) \cdot f_{ee}\left(C_{os,\,k+1}, P_k\right) \tag{4}$$

where f_s and f_{ee} are switching reward and economic emission reward defined by business rules, respectively. w represents the weight coefficient, $C_{os,\,k+1}$ and P_k are the value of outlet SO_2 concentration and total power consumption, respectively. The calculation of $f_s\left(T_{k+1}\right)$ and $f_{ee}\left(C_{os,\,k+1}, P_k\right)$ are as follows:

$$f_s\left(T_{k+1}\right) = \sum_{i=1}^{n} \frac{T_{i,\,k+1}}{4 \times I} \tag{5}$$

$$f_{ee}\left(C_{os,\,k+1}, P_{k+1}\right) = \begin{cases} 0.5 \times \mathrm{e}^{-4 \times C_{os,\,k+1}}, \\ \quad \text{if } C_{os,\,k+1} > C_{os,limit}; \\ x \frac{0.5}{1+\mathrm{e}^{3 \times (P_{k+1}-7)}} + 0.5, \\ \quad \text{else} \end{cases} \tag{6}$$

where I is positive inertia coefficient, and $C_{os,limit}$ is the emission standard of outlet SO_2 concentration.

The Q value neural network is composed of two fully connected neural networks, value network $V\left(s\theta_v\right)$ with parameter θ_v and advanced network $A\left(s, a\theta_a\right)$, with parameter θ_a. The definition of Q network is as follows:

$$Q\left(s, a\theta_{\theta_v,\theta_a}\right) = V\left(s\theta_v\right) + A\left(s, a\theta_a\right) - \frac{1}{2^n}\sum_{a'} A\left(s, a'\theta_a\right) \tag{7}$$

where $Q\left(s, a\theta_{\theta_v,\theta_a}\right)$ is the Q value neural network with parameters $\theta_{\theta_v,\theta_a}$, and inputs s and a are the discount factors.

After using the experience pool data processed from the DCS database of WFGD and learning the parameters of Q network through Q-learning, we define the greedy strategy as

$$a_{greedy} = \arg\max_a Q(s, a|\theta) \tag{8}$$

The operation parameters that are currently being executed is selected. Experiments show that the method can obtain the rewards close to the theoretical optimal solution in the WFGD steady state. At the same time, the strategy keeps its performance through concept drift adaptation without manual intervention.

4 Early Warning and Diagnosis of Abnormal Conditions

The FCCU is very large in appearance. Under the harsh conditions of high temperature and high pressure, a large number of toxic, harmful, flammable, and explosive hazardous chemicals may be produced in the production process. If there is a production problem, it may cause huge safety and environmental protection accidents, and cause major life and property losses[84]. In Sinopec's oil refining units, more than 37% of the alarms come from FCC, far exceeding other units. Among the unplanned shutdown events, 53% occurred in the FCCU, and the shutdown duration accounted for more than 72% of the total unplanned shutdown duration of all unit[85]. It is an important guarantee means for safety production to monitor the real-time operation status of production devices in real time, find and predict abnormal conditions and give early warning timely, find out the causes of problems through intelligent diagnosis, and guide staff to intervene in advance.

Early warning of abnormal conditions of FCC refers to the timely detection of problems by means of prediction when the operating parameters deviate but the alarm is not reached or the process status has not reached a more serious level. The diagnosis process is to find out the abnormal causes according to the characteristics of various quantities (measurable or unmeasurable) in the system that are different from the normal state on the basis of finding the problem[86].

In the abnormal early warning, on the one hand, it is necessary to avoid the missing report of production abnormalities, which may make the operator insufficiently prepared, miss the opportunity of early intervention, and eventually cause serious safety accidents. On the other hand, it is necessary to reduce the false alarms of production abnormalities. Frequent false alarms will bring a lot of unnecessary work and a waste of manpower. At the same time, it may gradually lead to operator's distrust of the alarm system and bring potential safety hazards[87,88]. It is of great significance for the safety production of FCC to use neural networks for improving the performance

of abnormal early warning and minimizing the level of false alarms on the premise that the predicted recall rate of abnormal conditions meets the safety production.

AI technology is applied to the operation process of FCCU to establish a perfect statistical analysis model, which has stronger real-time and predictability than human monitoring. It can predict in advance whether the key points of the production system will be abnormal and monitor the production system in real time from multiple angles to further reduce the occurrence of accidents and economic losses.

4.1 Data-driven methods

Anomaly early warning can be regarded as a regression problem. That is to predict the changing trend of indicator data representing conditions in the future according to a collected set of device parameters[86, 89–93]. Some researchers simplified the abnormal early warning as a classification problem according to the application needs. That is to predict whether there will be abnormal conditions in a future time window[94, 95].

Connectionist AI methods based on data-driven only need data information to complete modeling, which can effectively face with the complex, nonlinear, and time-varying characteristics of petrochemical processes, and are especially suitable for abnormal pattern recognition of complex petrochemical processes such as catalytic cracking[96]. The application of AI technologies, such as statistical machine learning and neural network to early warning, have been studied for many years. In 2000, Kourniotis et al.[97] subdivided a group of chemical accident data by region and time period, and determined the value range of parameter values with the help of Bayesian model, providing a theoretical basis for chemical production accident risk early warning. Salzano et al.[98] used a linear probability function to represent the vulnerability of equipment and carried out early warning research on industrial equipment accidents on the basis of obtaining the equipment reliability threshold. Wang[39] built an early warning index system from the human, machine, and environment, and established a support vector machine based risk early warning model.

However, due to the strong correlation among the process parameters of FCCU, the input features of the chemical model have a lot of redundant information. On the one hand, it will make the model more complex and require more data and larger computing resources to train the model. On the other hand, it will make the solution space unstable, resulting in the weak generalization of the model. Feature filtering or feature dimensionality reduction can be used to reduce redundant features, so as to reduce the complexity of the model and improve the training efficiency and model generalization. Principal Component Analysis (PCA) and Independent Component Analysis (ICA) are two commonly used dimensionality reduction methods. Reference [92] built a chemical process monitoring method based on PCA. Reference [93] applied ICA in chemical process risk analysis. Jiang et al.[99] proposed a method combining PCA and ICA. First, they projected the input features into the dominant subspace via PCA. Second, they extracted independent components from the dominant subspace determined by PCA. Finally, they established a Bayesian fault diagnosis system to identify the chemical process state. Some applications use machine learning methods combining feature space transformation with regression processes, such as Partial Least Squares regression (PLS)[100]. PCA, ICA, and PLS are linear transformations, while the relationship between process parameters of FCCU is highly nonlinear[18]. Therefore, nonlinear methods are often used to process the feature of FCC. Some studies use Mutual Information (MI) to select the features and eliminate the feature variables having low correlation with the predicted results, so as to achieve the purpose of reducing the feature dimension[90, 91]. Jiang et al.[99] grouped features by mutual information. The operation mode of the plant was identified via the T2 statistics after reducing the dimension of features via PCA in each group[101]. Reference [102] constructed a random forest prediction model and selected features according to the weight of features in the random forest model. Reference [103] used Spearman Ranking Correlation Coefficient (SRCC) to improve the accuracy of prediction model for fault diagnosis in the chemical process data.

According to the distribution characteristics of process parameter vectors under abnormal conditions, the Gaussian model is often applied to anomaly detection[104]. Reference [103] used the Gaussian mixture model to detect the anomalies of nonlinear systems. Reference [89] developed several detection algorithms based on the traditional Gaussian process regression, in which the mean function, covariance function, likelihood function, and inference method were specially designed. And the proposed scheme has fewer assumptions and is more suitable for modern

industrial processes compared with traditional detection methods. Zhang et al.[105] used fuzzy neural network to establish the normal behavior model of key equipment parameters, calculated the residual between the sample set data and the normal model, constructed a multivariate Gaussian distribution model of the residual, and set the abnormal state threshold by using the contour of Gaussian probability density. In addition, methods such as SVM[106], least squares SVM[107], random forest[108], and gradient boosting[109] are also often applied to anomaly detection.

With the widespread use of deep learning, more and more studies have used artificial neural networks for anomaly early warning[110]. As chemical process analysis is a typical time series problem, the recurrent neural networks represented by LSTM have been paid attention to in the research of abnormal working condition analysis[111]. In actual production, the number of abnormal conditions is far less than that of normal working conditions, thus, the samples are very uneven. Data enhancement is often required for using LSTM. The dataset is supplemented by dynamic simulation in the mechanism model[112, 113]. First, a simplified unit working mechanism model is constructed and dynamic simulation is carried out to obtain the dataset. The dataset is then fitted with an LSTM, and the potential relationship between variables that have a direct impact on accident hazards is discussed to predict abnormal trends. Reference [114] used an encoder-decoder module with the multi-head self-attention mechanism to deal with the imbalanced time series data. A novel density-based clustering density based clustering algorithms was proposed for incomplete data processing[115]. Zhan et al.[116] proposed a controllable gradient rotation strategy to realize local boundary expansion for positive samples. Reference [117] proposed a data enhancement method based on Generative Adversarial Networks (GANs) to synthesize fault data to solve the imbalance problem when using LSTM to predict the abnormal leakage of heat exchanger tubes.

LSTM mainly learns the global change law of data over a period of time, which is easy to ignore the local features. While CNN can capture the local features of data in a small window. Some studies of chemical anomaly warning add convolution layer to improve the performance of LSTM[118]. Figure 3 shows a network model combining CNN and Bi-LSTM for predicting abnormal conditions of FCCU[119].

The method consists of three parts: (1) convolution layer; (2) Bi-LSTM layer, and (3) attention mechanism module. After one-dimensional convolution, the features filtered by SRCC are used to obtain a hidden feature sequence through a single-layer Bi-LSTM. And then, the attention module is used to weight each element of

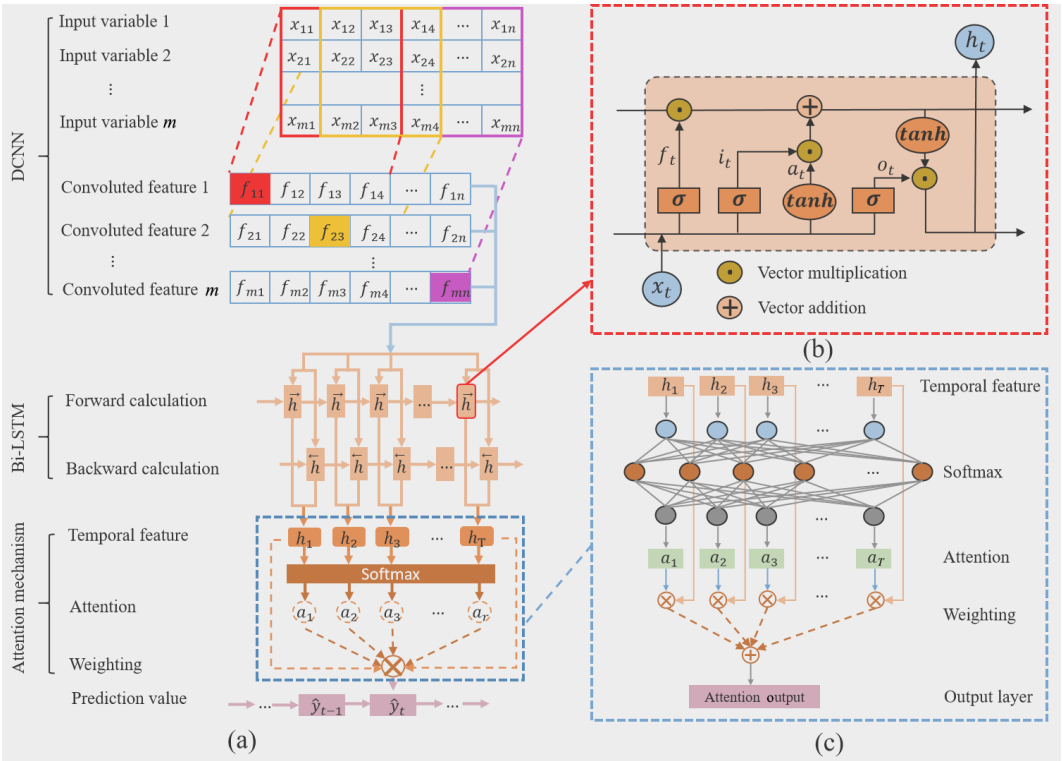


Fig. 3 Model combining CNN and Bi-LSTM for predicting abnormal conditions of FCCU[119].

the sequence to obtain the prediction results of target parameters.

In industrial scenarios, during the running process of the device, the operating environment and condition may change over time, continuing generating data belong to unknown classes with new characteristics and distribution. To address this challenging problem, Chen et al.[120] proposed a generic open set signal classification method, using Fourier transform and variational encoder-classifier network to determine whether samples belong to unknown or not.

4.2 Knowledge-driven methods

Symbolic AI methods based on knowledge-driven bulid a rule system according to the working principle and historical fault data of chemical plants. By hitting the rules with the values of process parameters and change rules, it can predict whether abnormal conditions will occur and infer the causes of abnormal conditions[121].

There are two main types of intelligent diagnosis methods for abnormal conditions, namely, fault tree based methods and Signed Directed Graph (SDG) based methods.

On the basis of the fault tree, a tree structure is created in which the least expected occurrence of a system event is considered the top event, and events that may lead to the top event are referred to as the intermediate events, or bottom events, and the logic gates are used to illustrate the relationship among events[122]. The minimum cut set of each event is found recursively after building the fault tree. The cut set refers to the bottom event set of the fault tree. When these bottom events occur, the top event must occur. For the minimum cut set, the cut set is no longer formed by removing any of the bottom events. A minimum cut set corresponds to a fault mode. And the fault causes can be traced by testing the minimum cut sets one by one[123]. Some studies have proposed probabilistic ranking methods for fault modes to improve the diagnosis speed[124]. The fault tree method itself is a qualitative causal model. And the fault tree used in some recent studies replaces the inevitable relationship between cut sets and events with occurrence probability. Reference [125] constructed the fault probability density function of lower level events under the influence of multiple factors, where the T-S dynamic gate is used to calculate the fault probability distribution function of upper level events based on the sequence rules of lower level events and the output rules of upper level events. Figure 4 shows the fault tree that causes abnormal conditions of FCC carbon

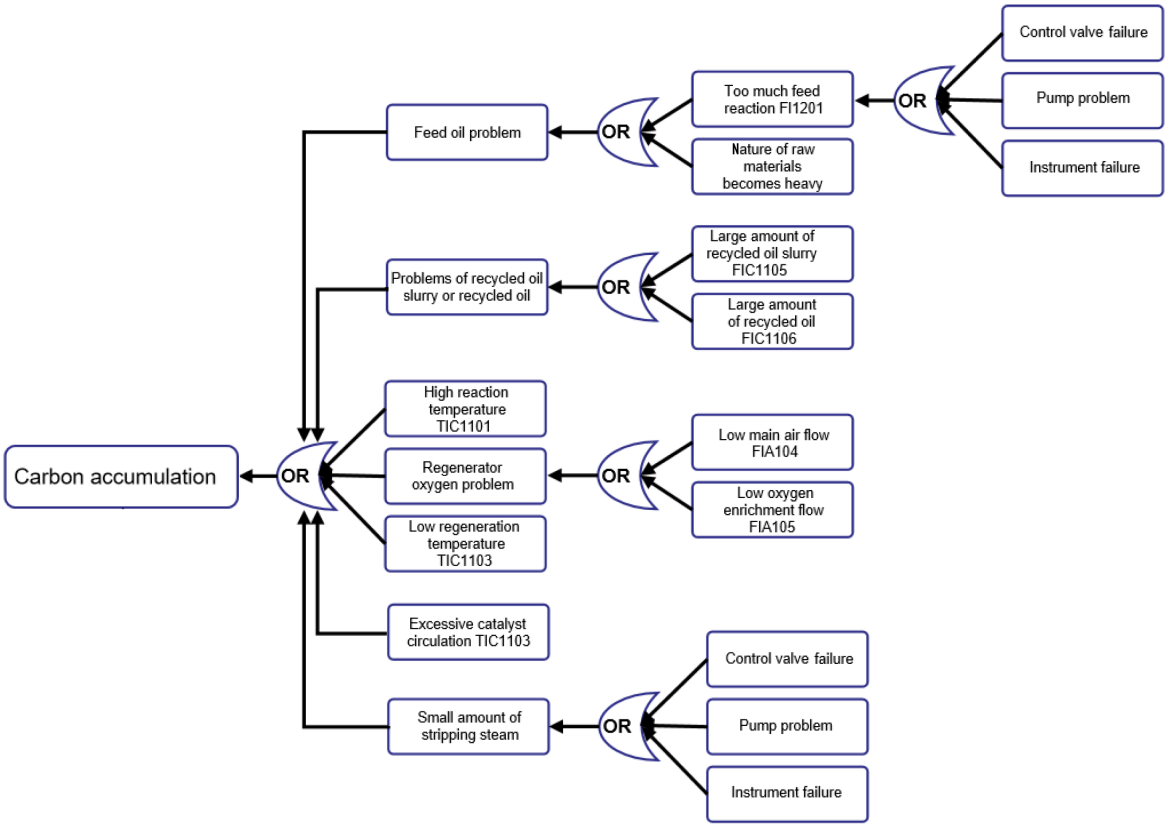


Fig. 4 Fault tree that causes abnormal conditions of FCC carbon accumulation[126].

accumulation[126].

However, the fault tree, strictly depending on the summary of people's experiences, is often suitable for the determination of abnormal conditions with certain types, which usually have high accuracy but insufficient recall. For uncertain abnormal conditions, the cause of abnormal conditions is usually traced through the relationship among complex variables based on SDG[127]. The most likely root cause of abnormal conditions is found in the fault logic relationship database by combining the deviation degree of alarm points with fuzzy rules. SDG is a graph representation of process causal information, in which process variables are represented as graph nodes and causal relationships are represented as directed edges. SDG nodes contain important attributes, i.e., node status, including "0", "+", and "−", which represent normal steady-state value, above steady-state value, and below steady-state value, respectively. The edge points from the cause node to the effect node. And there are two types of SDG edges, solid lines and dashed lines, which indicate the positive and negative correlations between the changing trends of the starting point and the end point, respectively[128]. Figure 5[129] shows a simple SDG model structure, which means that an increase in *A* leads to an increase in *B*, and then leads to a decrease in *C*.

The core of the SGD-based anomaly analysis method is to decompose the anomaly Cause-Result Graph (CEG) into several strongly connected units (a single node is also a strongly connected unit), the largest of which contains the root cause of the anomaly. Thus, the process of anomaly diagnosis is essentially the process of finding the maximum strongly connected units.

Under the assumption that the root cause of equipment abnormality should be unique, the maximum strongly connected unit is unique[130]. However, there may be multiple abnormal sources at the same time during the production process of FCCU. To solve this problem, Reference [128] introduced the concept of minimum cut set in fault tree into SDG, identifying the minimum number of fault origins could help to explain the failure the process. Reference [131] used the reverse reasoning strategy instead of the common forward reasoning strategy to reason multiple abnormal sources.

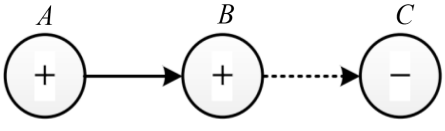


Fig. 5 Structure of SDG.

Similar to the data-driven method, it is often necessary to select and transform the variables that form nodes when constructing SDG model. Reference [132] used the idea of PCA weight to select the key variables with large weight from the multi-layer correlation coefficient set, which reduced the complexity of the SDG model. And the deep-seated correlation feature of SDG can be fully exploited, which has advantages of process monitoring.

Knowledge graph is an increasing popular method of symbolic AI. It uses nodes to represent a research object and edges to represent the relationships among objects. The graph can contain multiple classes of objects and relationships[133]. The powerful knowledge representation capability of the knowledge graph can be used to uniformly model the nodes of the fault tree and the nodes of the SDG into the same graph[134]. In addition, the mechanism process can be introduced into the knowledge graph to enhance the interpretability[135], and then, the visual operation of the knowledge graph can be used to interactively query the pattern graph to find the cause of the abnormality[136].

Knowledge-driven methods for early warning and diagnosis of abnormal conditions have stronger interpretability, higher accuracy, and better generalization ability than data-driven methods. Therefore, these methods are more suitable for abnormal diagnoses to meet people's understanding of the diagnosis causes. However, the high construction cost and the difficulty of implementation are usually the biggest challenges for this type of methods.

4.3 Methods by combining data and knowledge

Knowledge-driven anomaly early warning and diagnosis has strong interpretability, high accuracy, and generalization ability. And it can interact well with operators. Operators can often expand their diagnosis ability according to their own experience. However, there is a problem of low recall due to the difficulty of building and covering all cases of abnormal cause results. And the abnormal causes cannot be found in many scenarios. On the contrary, the data-driven method has advantages in construction difficulty and prediction accuracy but has obvious disadvantages in interpretability and generalization. Table 1 shows the comparison of these two methods. More and more recent studies have combined these two methods to obtain better results.

Since the acquisition and collection of expert

knowledge requires a great cost of time and manpower, some scholars have begun to try to excavate abnormal causal relationships through data-driven methods, such as association rules and decision trees[137].

Reference [93] proposed a method for mapping fault trees to neural networks. First, the fault tree of the equipment is constructed, and then the process parameters of the bottom event satisfying the fault event cut set are randomly generated by the fault tree. These process parameters and fault results are added to the training set of the neural network as samples.

Reference [138] proposed a fuzzy Bayesian network model on reasoning fault diagnosis for complex equipment that is based on the fault tree. First, a fault tree model of complex equipment is established by analyzing the structure of complex equipment. Second, fault tree based Bayesian network topology is constructed by using fault tree transformation method. And then, the parameters were determined such as conditional probability by fuzzy set theory, aiming at the lack of structure data on complex equipment and the uncertainty of expert scoring.

Some scholars combine SDG with data-driven methods. Some of these methods propose possible causes and paths of abnormal phenomena via SDG, and then find the most likely path via time series analysis[139]. And some methods combine SDG with machine learning models, such as PLS[140], SVM[141] or neural network[119], which predict the abnormal states of some key points via machine learning models and find the causes of these abnormal states in SDG.

At present, there are a large number of methods that combine data and knowledge to realize the early warning and diagnosis of abnormal conditions. Many scholars are making various attempts and have achieved good results. It can be seen that such methods are likely to become an important research direction in the future.

Table 1 Comparison of the two methods.

Method	Interpretability	Construction	Generalization ability	Metric
Data-driven	Bad	Bad	Bad	High accuracy, high recall
Knowledge-driven	Good	Good	Good	High accuracy, low recall

5 Conclusion

Product analysis and optimization, FGD prediction, and abnormal condition early warning are several key research directions in FCC process analysis. The paper provides a comprehensive review of FCC process analysis, mainly introducing methods based on traditional mathematical mechanisms and AI. Compared with traditional mathematical mechanism methods, AI methods can effectively solve the difficulties in chemical process modeling, such as high-dimensional, nonlinear, strong correlation of influence factors, time-varying, large lag, and uncertainty: (1) Machine learning itself depends on data-driven, and can solve high-dimensional problems automatically and efficiently; (2) The nonlinear mapping relationship in chemical process can be fitted effectively by applying kernel function and nonlinear machine learning model; (3) The problem of strong correlation among influencing factors can be solved by feature selection or dimensionality reduction methods, such as PCA, ICA, and PLS; (4) The model that conforms to the time-varying and large lag characteristics of chemical processes can be constructed by leveraging the RNN structure; (5) The combination of machine learning with mechanism model can further reduce the uncertainty and improve the prediction performance.

In a word, AI algorithm has become an important approach to chemical process modeling and analysis. In future research, hybrid models combining with mechanism models and AI algorithms are expected to become powerful tools for more comprehensive and accurate analysis of chemical processes and prediction of production results. These methods will play an important role in the future development of the chemical industry and will be of great value.

Acknowledgment

This work was supported in part by the State Key Program of National Science Foundation (No. 61836006), the National Natural Science Fund for Distinguished Young Scholar (No. 61625204), the National Natural Science Foundation of China (Nos. 62106161 and 61602328), as well as the Key Research and Development Project of Sichuan (No. 2019YFG0494).

References

[1] N. L. A. Souza, I. Tkach, E. M. Jr, and K. Krambrock, Vanadium poisoning of FCC catalysts: A quantitative analysis of impregnated and real equilibrium catalysts,

Applied Catalysis A General., vol. 560, no. 12, pp. 206–214, 2018.

[2] F. C. Salvado, F. Teixeira-Dias, S. M. Walley, L. Lea, and J. B. Cardoso, A review on the strain rate dependency of the dynamic viscoplastic response of FCC metals, *Progress in Materials Science*, vol. 88, no. 7, pp. 186–231, 2017.

[3] C. X. Lu, Y. P. Fan, M. X. Liu, and X. Y. Yao, Advances in key equipment technologies of reaction system in RFCC unit, (in Chinese), *Acta Petrolei Sinica* (*Petroleum Processing Section*), vol. 34, no. 3, pp. 441–454, 2018.

[4] C. H. Yang, X. B. Chen, C. Y. Li, and H. Shan, Challenges and opportunities of fluid catalytic cracking technology, (in Chinese), *Journal of China University of Petroleum.*, vol. 41, no. 6, pp. 171–177, 2017.

[5] F. Yang, C. N. Dai, J. Q. Tang, J. Xuan, and J. Cao, A hybrid deep learning and mechanistic kinetics model for the prediction of fluid catalytic cracking performance, *Chem. Eng. Res. Des.*, vol. 155, pp. 202–210, 2020.

[6] Z. Chen, S. Feng, L. Z. Zhang, Q. Shi, Z. M. Xu, S. Q. Zhao, and C. M. Xu, Molecular-level kinetic modelling of fluid catalytic cracking slurry oil hydrotreating, *Chemical Engineering Science*, vol. 195, pp. 619–630, 2018.

[7] P. Cristina, Four-lump kinetic model vs. three-lump kinetic model for the fluid catalytic cracking riser reactor, *Procedia Engineering*, vol. 100, pp. 602–608, 2015.

[8] H. Zhang, H. Z. Wang, J. Z. Li, and L. H. Gao, A generic data analytics system for manufacturing production, *Big Data Mining and Analytics*, vol. 1, no. 2, pp. 160–171, 2018.

[9] Z. H. Yuan, W. Z. Qin, and J. S. Zhao, Smart manufacturing for the oil refining and petrochemical industry, *Engineering*, vol. 2, pp. 66–74, 2017.

[10] P. Nitu, J. Coelho, and P. Madiraju, Improvising personalized travel recommendation system with recency effects, *Big Data Mining and Analytics*, vol. 4, no. 3, pp. 39–154, 2021.

[11] S. K. Patnaik, C. N. Babu, and M. Bhave. Intelligent and adaptive web data extraction system using convolutional and long short-term memory deep learning networks, *Big Data Mining and Analytics*, vol. 4, no. 4, pp. 279–297, 2021.

[12] G. Zhai, Y. Yang, H. Wang, and Du. S, Multi-attention fusion modeling for sentiment analysis of educational big data, *Big Data Mining and Analytics*, vol. 3, no. 4, pp. 311–319, 2020.

[13] A. Guezzaz, Y. Asimi, M. Azrour, and A. Asimi, Mathematical validation of proposed machine learning classifier for heterogeneous traffic and anomaly detection, *Big Data Mining and Analytics*, vol. 4, no. 1, pp. 18–24, 2021.

[14] N. Yuvaraj, K. Srihari, S. Chandragandhi, R. A. Raja, G. Dhiman, and A. Kaur, Analysis of protein-ligand interactions of SARS-Cov-2 against selective drug using deep neural networks, *Big Data Mining and Analytics*, vol. 4, no. 2, pp. 76–83, 2021.

[15] Z. Tong, F. Ye, M. Yan, H. Liu, and S. Basodi, A survey on algorithms for intelligent computing and smart city applications, *Big Data Mining and Analytics*, vol. 4, no. 3, pp. 155–172, 2021.

[16] B. Zhang, J. Zhu, and H. Su, Toward the third generation of artificial intelligence, (in Chinese), *Scientia Sinica* (*Informationis*), vol. 50, no. 9, pp. 1281–1302, 2020.

[17] F. Qian, W. M. Zhong, and W. L. Du, Fundamental theories and key technologies for smart and optimal manufacturing in the process industry, *Engineering*, vol. 3, no. 2, pp. 154–160, 2017.

[18] M. T. Shah, R. P. Utikar, V. K. Pareek, G. M. Evans, and J. B. Joshi, Computational fluid dynamic modelling of FCC riser: A review, *Chemical Engineering Research and Design*, vol. 111, pp. 403–448, 2016.

[19] D. V. Naik, V. Karthik, V. Kumar, B. Prasad, and M. O. Garg, Kinetic modeling for catalytic cracking of pyrolysis oils with VGO in a FCC unit, *Chemical Engineering Science*, vol. 170, pp. 790–798, 2017.

[20] J. Chang, W. Cai, K. Zhang, F. Zhang, L. Wang, and Y. Yang, Computational investigation of the hydrodynamics, heat transfer and kinetic reaction in an FCC gasoline riser, *Chemical Engineering Science*, vol. 111, pp. 170–179, 2014.

[21] M. Ahsan, Computational fluid dynamics (CFD) prediction of mass fraction profiles of gas oil and gasoline in fluid catalytic cracking (FCC) riser, *Ain Shams Engineering Journal*, vol. 3, no. 4, pp. 403–409, 2012.

[22] J. X. Zhang, Y. H. Qi, and J. Z. Qiu, Study on FCC correlation models, (in Chinese), *Computers and Applied Chemistry*, vol. 24, no. 11, pp. 1519–1522, 2007.

[23] P. Cristina, Four-lump kinetic model vs. three-lump kinetic model for the fluid catalytic cracking riser reactor, *Procedia Engineering*, vol. 100, pp. 602–608, 2015.

[24] Z. Chen, S. Feng, L. Zhang, S. Quan, and C. Xu, Molecular-level kinetic modelling of fluid catalytic cracking slurry oil hydrotreating, *Chemical Engineering Science*, vol. 195, pp. 619–630, 2018.

[25] F. Yang, M. Zhou, J. M. Jin, and J. Cao, Research progress on application of intelligent optimization algorithm and artificial neural network in FCC model analysis, (in Chinese), *Acta Petrolei Sinica* (*Petroleum Processing Section*), vol. 36, no. 4, pp. 878–888, 2020.

[26] C. Tang, C. Yu, Y. Gao, J. Chen, and J. Yang, Deep learning in nuclear industry: A survey, *Big Data Mining and Analytics*, vol. 5, no. 2, pp. 140–160, 2022.

[27] H. Wang, Z. Xu, H. Fujita, and S. Liu, Towards felicitous decision making: An overview on challenges and trends of big data, *Information Sciences*, vols. 367&368, pp. 747–765, 2016.

[28] S. Wang, J. Wan, D. Zhang, D. Li, and C. Zhang, Towards smart factory for industry 4.0: A self-organized multi-agent system with big data based feedback and coordination, *Computer Networks*, vol. 101, pp. 158–168, 2016.

[29] D. J. Doong, J. P. Peng, and Y. C. Chen, Development of a warning model for coastal freak wave occurrences using an artificial neural network, *Ocean Engineering*, vol. 169, pp. 270–280, 2018.

[30] J. H. Wei, M. Zhou, X. M. Wei, G. Z. He, L. Q. Song, and J. L. Lei, Prediction of coke strength using linear regression method, (in Chinese), *China Coal.*, vol. 37, no. 6, pp. 90–93, 2011.

[31] C. Q. Lv, A multiple linear regression model to predict chemical fiber output, (in Chinese), *China Synthetic Fiber Industry*, vol. 9, no. 2, pp. 1–25, 2013.

[32] J. Gmeinbauer, C. Ramakrishnan, and A. Reichhold, Prediction of FCC product distributions by means of feed parameters. *Oil Gas European Magazine*, vol. 120, no. 3, pp. 29–33, 2004.

[33] W. A. Anderson, P. M. Reilly, and M. Moo-Young, Application of a Bayesian regression method to the estimation of diffusivity in hydrophilic gels, *Canadian Journal of Chemical Engineering*, vol. 70, no. 3, pp. 499–504, 2010.

[34] Z. C. Sun, H. H. Shan, Y. B. Liu, and Y. Li, Modeling and optimization for catalytic cracking products of heavy oil based on support vector regression, (in Chinese), *Petroleum Processing and Petrochemicals*, vol. 43, no. 5, pp. 76–81, 2012.

[35] P. S. Roy, C. Ryu, S. K. Dong, and C. S. Park, Development of a natural gas methane number prediction model, *Fuel.*, vol. 246, pp. 204–211, 2019.

[36] X. Yuan, Z. Ge, and Z. Song, Locally weighted kernel principal component regression model for soft sensing of nonlinear time-variant processes, *Industrial & Engineering Chemistry Research*, vol. 153, pp. 116–125, 2014.

[37] X. Zhang, M. Kano, and Y. Li, Locally weighted kernel partial least squares regression based on sparse nonlinear features for virtual sensing of nonlinear time-varying processes, *Computers & Chemical Engineering*, vol. 104, pp. 164–171, 2017.

[38] F. Yang, M. Zhou, C. N. Dai, and J. Cao, Construction and analysis of gasoline yield prediction model for FCC unit based on artificial intelligence algorithm, (in Chinese), *Acta Petrolei Sinica* (*Petroleum Processing Section*), vol. 35, no. 4, pp. 807–817, 2019.

[39] X. Wang, Research on safety production early-warning system for ammonia plant, Master dissertation, Anhui University of Science and Technology, China, 2013.

[40] Y. Y. Zhao, Application of data mining in gasoline yield optimization for MIP process, Master dissertation, East China University of Science and Technology, Shanghai, 2018.

[41] J. Wang, D. F. Cao, X. Y. Lan, and J. S. Gao, Select filter-wrapper characteristic variables for yield prediction of fluid catalytic cracking unit, (in Chinese), *CIESC Journal*, vol. 69, no. 1, pp. 464–471, 2018.

[42] W. Wang, K. Wang, F. Yang, C. N. Dai, and J. M. Jin, Construction and analysis of gasoline yield prediction model for Fluid Catalytic Cracking Unit (FCCU) based on GBDT and P-GBDT algorithm, *Acta Petrolei Sinica* (*Petroleum Processing Section*), vol. 1, pp. 179–187, 2020.

[43] C. Y. Lv, J. H. Wu, and H. M. Lin, Operation optimization of neural networks for reaction-reproduction system of fluid catalytic cracking unit, (in Chinese), *Computers and Applied Chemistry*, vol. 19, pp. 447–450, 2002.

[44] G. R. Wang, Application of neural network technology in prediction models for the yield of hydrocracking products, (in Chinese), *Petrochemical Technology &Application*, vol. 4, pp. 349–353, 2015.

[45] H. Jiang, J. Tang, and F. S. Ouyang, A new method for the prediction of the gasoline yield of the MIP process, *Petroleum Science and Technology*, vol. 33, no. 20, pp. 1713–1720, 2015.

[46] C. Shang, F. Yang, D. Huang, and W. L. Yu, Data-driven soft sensor development based on deep learning technique, *Journal of Process Control*, vol. 24, no.3, pp. 223–233, 2014.

[47] S. Hochreiter and J. Schmidhuber, Long short-term memory, *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1996.

[48] W. S. Ke, D. X. Huang, F. Yang, and Y. Jiang, Soft sensor development and applications based on LSTM in deep neural networks, in *Proc. on 2017 IEEE Symposium Series on Computational Intelligence* (*SSCI*), Honolulu, HI, USA, pp. 1–6, 2017.

[49] X. Zhang, Y. Y. Zou, S. Y. Li, and S. Xu, Product yields forecasting for FCCU via deep Bi-directional LSTM network, in *Proc. of Chinese Control Conference* (*CCC*), Wuhan, China, 2018, pp. 8013–8018.

[50] X. Zhang, Y. Zou, S. Li, and S. Xu, A weighted auto regressive LSTM based approach for chemical processes modeling, *Neurocomputing*, vol. 367, pp. 64–74, 2019.

[51] Y. Dauphin, R. Pascanu, C. Gulcehre, K. Cho, S. Ganguli, and Y. Bengio, Identifying and attacking the saddle point problem in high-dimensional non-convex optimization, *Advances in Neural Information Processing Systems*, vol. 4, pp. 2933–2941, 2014.

[52] P. K. Dasila, I. R. Choudhury, D. N. Saraf, V. Kagdiyal, and S. J. Chopra, Estimation of FCC feed composition from routinely measured lab properties through ANN model, *Fuel Process Technol.*, vol. 125, pp. 155–162, 2014.

[53] S. Ruder, An overview of gradient descent optimization algorithms, arXiv:1609.04747, 2016.

[54] Q. Fan, Research on micro-vortex coagulation dosing control model based on genetic algorithm and BP neural network, (in Chinese), Master dissertation, East China Jiaotong University, Shanghai, 2018.

[55] F. S. Ouyang, J. F. You, and W. Fang, Optimizing product distribution of MIP process by BP neural network combined with genetic algorithm, (in Chinese), *Petroleum Processing and Petrochemicals*, vol. 49, no. 8, pp. 98–104, 2018.

[56] F. S. Ouyang and Y. Liu, Prediction of the product yield from catalytic cracking process by lumped kinetic model combined with neural network, (in Chinese), *Petrochemical Technology*, vol. 46, no. 1, pp. 9–16, 2017.

[57] F. S. Ouyang, W. Fang, and J. Tang, Optimizing product distribution of MIP process using BP neural network, (in Chinese), *Petroleum Processing and Petrochemicals*, vol. 47, no. 5, pp. 95–100, 2016.

[58] W. G. Fang, Optimizing product distribution of FCC MIP process by data mining technology, (in Chinese), Master dissertation, East China University of Science and Technology, Shanghai, 2016.

[59] X. Su, H. Pei, and Y. Wu, Predicting coke yield of FCC unit using genetic algorithm optimized BP neural network, (in Chinese), *Chemical Industry and Engineering Progress*, vol. 35, no. 2, pp. 389–396, 2016.

[60] Z. Wang, B. Yang, and C. Chen, Modeling and optimization for the secondary reaction of FCC gasoline based on the fuzzy neural network and genetic algorithm, *Chemical Engineering & Processing Process Intensification*, vol. 46, no. 1, pp. 175–80, 2014.

[61] Y. M. Gao, Y. F. Xing, J. Fu, W. Zhang, and J. H. Zhao, The application of BP neural network based on a particle swarm optimization to the catalytic cracking reaction-regeneration process, (in Chinese), *Computers & Applied Chemistry*, vol. 34, no. 11, pp. 899–903, 2017.

[62] Y. Shang, W. Xu, and X. Gu, Particle swarm optimization with pre-crossover and its application in soft-sensor of C3 concentration of FCCU, (in Chinese), *Acta Electronica Sinica*, vol. 40, no. 9, pp. 1885–1888, 2012.

[63] X. Wang, X. Gu, and Z. Liu, Soft sensing model of C3 concentration of FCCU based on PSO-BP neural network, (in Chinese), *Journal of System Simulation*, vol. 21, no. 4, pp. 973–980, 2009.

[64] J. R. Tang, A new method for modeling and optimization of gasoline yield of MIP process, (in Chinese), Master dissertation, East China University of Science and Technology, Shanghai, 2016.

[65] W. Z. Qin, D. X. Xie, and J. S. Zhao, *Petrochemical Industry Intelligent Manufacturing* (in Chinese). Beijing, China: Chemical Industry Press, 2019.

[66] G. M. Bollas, A. A. Lappas, D. K. Iatridis, and I. A. Vasalos, Five-lump kinetic model with selective catalyst deactivation for the prediction of the product selectivity in the fluid catalytic cracking process, *Catalysis Today*, vol. 127, nos. 1–4, pp. 31–43, 2007.

[67] G. M. Bollas, J. Papadokonstadakis, J. Michalopoulos, G. Arampatzis, A. A. Lappas, I. A. Vasalos, and A. Lygeros, Using hybrid neural networks in scaling up an FCC model from a pilot plant to an industrial unit, *Chemical Engineering and Processing*, vol. 42, nos. 8&9, pp. 697–713, 2003.

[68] Y. J. Liu, Prediction of the product yields from heavy oil catalytic cracking process by lumped kinetic model combined with neural network, (in Chinese), Master dissertation, East China University of Science and Technology, Shanghai, 2017.

[69] H. Zhang, B. Zhang, and J. Bi, More efforts, more benefits: Air pollutant control of coal-fired power plants in China, *Energy.*, vol. 80, pp. 1–9, 2015.

[70] H. N. Tang, Analysis of FCCU wet flue gas scrubbing technologies, (in Chinese), *Petroleum Refinery Engineering*, vol. 42, no. 3, pp. 1–5, 2012.

[71] Y. Zhong, X. Gao, W. Huo, Z. Luo, M. Ni, and K. Chen, A model for performance optimization of wet flue gas desulfurization systems of power plants, *Fuel Processing Technology*, vol. 89, no. 11, pp. 1025–1032, 2008.

[72] W. Y. Tan, Z. X. Zhang, H. Y. Li, Y. X. Li, and Z. W. Shen, Carbonation of gypsum from wet flue gas desulfurization process: Experiments and modeling, *Environmental Science and Pollution Research*, vol. 24, no. 9, pp. 8602–8608, 2016.

[73] B. Wu, W. He, J. Wang, H. Liang, and C. Chen, A convolutional-LSTM model for nitrogen oxide emission forecasting in FCC unit, *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 1, pp. 1537–1545, 2021.

[74] Z. Qiao, X. Wang, H. Gu, Y. Tang, F. Si, C. E. Romero, and X. Z. Yao, An investigation on data mining and operating optimization for wet flue gas desulfurization systems. *Fuel.*, doi:10.1016/j.fuel.2019.116178.

[75] Z. B. Ren, L. Sun, and Y. Deng, Modeling and optimization research of CFB-FGD based on improved genetic algorithms and BP neural network, *Advanced Materials Research*, vols. 610–613, pp. 1601–1604, 2012.

[76] J. Warych and M. Szymanowski, Model of the wet limestone flue gas desulfurization process for cost optimization, *Ind. Eng. Chem. Res.*, vol. 40, no. 12, pp. 2597–2605, 2001.

[77] P. F. Hou, J. Y. Bai, and J. Yin, On-line monitoring and optimization of performance indexes for limestone wet desulfurization technology, *Applied Mechanics and Materials*, vols. 295–298, pp. 1020–1028, 2013.

[78] F. Vieira and H. M. Ramos, Optimization of operational planning for wind/hydro hybrid water supply systems, *Renew Energy*, vol. 34, no. 3, pp. 928–936, 2009.

[79] J. Lang and J. Zhao, Modeling and optimization for oil well production scheduling, *Chinese Journal of Chemical Engineering*, vol. 24, no. 10, pp. 1423–1430, 2016.

[80] V. Kapsalis and L. Hadellis, Optimal operation scheduling of electric water heaters under dynamic pricing, *Sustain. Cities Soc.*, vol. 31, pp. 109–121, 2017.

[81] F. Yang, Y. S. Luo, C. S. Zhang, and T. Liu, Process optimization method, device and equipment and computer readable storage medium, (in Chinese), China Patent CN201911424710.6, June, 2020.

[82] J. Filipe, R. J. Bessa, M. Reis, R. Alves, and P. Povoa, Data-driven predictive energy optimization in a wastewater pumping station, *Applied Energy.*, vol. 252, p. 113423, 2019.

[83] Z. Shao, F. Si, D. Kudenko, P. Wang, and X. Tong, Predictive scheduling of wet flue gas desulfurization system based on reinforcement learning, *Computers & Chemical Engineering*, vol. 141, p. 107000, 2020.

[84] H. Gharahbagheri, S. Imtiaz, and F. Khan, Root cause diagnosis of process fault using KPCA and Bayesian network, *Ind. Eng. Chem. Res.*, vol. 56, no. 8, pp. 2054–2070, 2017.

[85] P. Li, X. J. Zheng, and L. Ming, Application of big data technology in operation analysis of catalytic cracking, (in Chinese), *Chemical Industry and Engineering Progress*, vol. 35, no. 3, pp. 665–670, 2016.

[86] H. Z. Wang, J. F. Zhu, J. S. Zhao, T. Qiu, and B. Chen, Kalmen filter and grey system based abnormal event trend predictio method in chemical progress, (in Chinese), *Computers and Applied Chemistry*, vol. 32, no. 3, pp. 257–260, 2015.

[87] F. Yang and D. Y. Xiao, Research topics of intelligent alarm management, (in Chinese), *Computers and Applied Chemistry*, vol. 28, no. 12, pp. 1485–1491, 2011.

[88] F. Yang, S. L. Shah, D. Xiao, and T. Chen, Improved correlation analysis and visualization of industrial alarm data, *ISA T.*, vol. 51, no. 4, pp. 499–506, 2012.

[89] B. Wang and Z. Mao, Outlier detection based on Gaussian process with application to industrial processes, *Applied Soft Computing*, vol. 76, pp. 505–516, 2019.

[90] Y. Ren, J. Wang, and W. Tian, Fault detection and identification in chemical processes based on feature engineering and kernel extreme learning machine, *J. Chem. Eng. Chin. Univ.*, vol. 33, no. 5, pp. 1271–1284, 2019.

[91] W. Tian, Y. Ren, Y. Dong, S. Wang, and L. Bu, Fault monitoring based on mutual information feature engineering modeling in chemical process, *Chinese Journal of Chemical Engineering*, vol. 27, no. 10, pp. 2491–2497, 2019.

[92] Y. Cao, Y. Yuan, Y. Wang, and W. Gui, Hierarchical hybrid distributed PCA for plant-wide monitoring of chemical processes, *Control Engineering Practice*, vol. 111, p. 104784, 2021.

[93] M. Sarbayev, M. Yang, and H. Wang, Risk assessment of process systems by mapping fault tree into artificial neural network, *Journal of Loss Prevention in the Process Industries*, vol. 60, pp. 203–212, 2019.

[94] F. Yang, J. M. Jin, and Y. H. Wang, Early warning method and device, (in Chinese), China Patent CN201811457374.0, 2019.

[95] J. M. Jin, F. Yang, and P. Yang, Data processing method and device, and electronic equipment, (in Chinese), China Patent CN 201911403955.0, 2020.

[96] Z. Ge, Z. Song, and F. Gao, Review of recent research on data-based process monitoring, *Industrial & Engineering Chemistry Research*, vol. 52, no. 10, pp. 3543–3562, 2013.

[97] S. P. Kourniotis, C. T. Kiranoudis, and N. C. Markatos, Statistical analysis of domino chemical accidents, *Journal of hazardous materials*, vol. 71, nos. 1–3, pp. 239–252, 2000.

[98] E. Salzano, A. G. Agreda, A. D. Carluccio, and G. Fabbrocino, Risk assessment and early warning systems for industrial facilities in seismic zones, *Reliability Engineering & System Safety*, vol. 94, no. 10, pp. 1577–1584, 2009.

[99] Q. Jiang, X. Yan, and J. Li, PCA-ICA integrated with Bayesian method for non-Gaussian fault diagnosis, *Industrial & Engineering Chemistry Research*, vol. 55, no. 17, pp. 4979–4986, 2016.

[100] C. Botre, M. Mansouri, M. N. Karim, H. Nounou, and M. Nounou, Multiscale PLS-based GLRT for fault detection of chemical processes, *Journal of Loss Prevention in the Process Industries*, vol. 46, pp. 143–153, 2017.

[101] Q. Jiang and X. Yan, Monitoring multi-mode plant-wide processes by using mutual information-based multi-block PCA, joint probability, and Bayesian inference, *Chemometrics and Intelligent Laboratory Systems*, vol. 136, pp. 121–137, 2014.

[102] C. Chen, N. Lu, L. Wang, and Y. Xing, Intelligent selection and optimization method of feature variables in fluid catalytic cracking gasoline refining process, *Computers & Chemical Engineering*, vol. 150, p. 107336, 2021.

[103] M. Karami and L. Wang, Fault detection and diagnosis for nonlinear systems: A new adaptive Gaussian mixture modeling approach, *Energy and Buildings*, vol. 166, pp. 477–488, 2018.

[104] A. Guezzaz, Y. Asimi M. Azrour, and A. Asimi, Mathematical validation of proposed machine learning classifier for heterogeneous traffic and anomaly detection, *Big Data Mining and Analytics*, vol. 4, no. 1, pp. 18–24, 2021.

[105] Y. X. Zhang, Y. Zheng, X. Y. Qian, and G. Mohanmmed, Abnormaly monitoring based on fuzzy neural network with hybrid input for wind turbine generator, (in Chinese), *Control Engineering*, vol. 28, no. 4, p. 9, 2021.

[106] M. Onel, C. A. Kieslich, Y. A. Guzman, C. A. Floudas, and E. N. Pistikopoulos, Big data approach to batch process monitoring: Simultaneous fault detection and diagnosis using nonlinear support vector machine-based feature selection, *Computers & Chemical Engineering*, vol. 115, pp. 46–63, 2018.

[107] J. Q. Hu, F. Guo, and L. B. Zhang, Study on ultra-early prediction of chemical abnormal situation based on improved PSO algorithm and LSSVM, (in Chinese), *Journal of Electronic Measurement and Instrumentation*, vol. 32, no. 2, pp 36–41, 2018.

[108] Y. Zhang, L. Luo, X. Ji, and Y. Dai, Improved random forest algorithm based on decision paths for fault diagnosis of chemical process with incomplete data, *Sensors-Basel*, vol. 21, no. 20, p. 6715, 2021.

[109] R. Shrivastava, Comparative study of boosting and bagging based methods for fault detection in a chemical process, in *Proc. on 2021 International Conference on Artificial Intelligence and Smart Systems* (*ICAIS*), Coimbatore, India, 2021, pp. 674–679.

[110] S. Heo and J. H. Lee, Fault detection and classification using artificial neural networks, *IFAC International Symposium on Advanced Control of Chemical Processes*, vol. 51, no. 18, pp. 470–475, 2018.

[111] R. Arunthavanathan, F. Khan, S. Ahmed, and S. Imtiaz, A deep learning model for process fault prognosis, *Process Safety and Environmental Protection*, vol. 154, pp. 467–479, 2021.

[112] Z. Liu, W. Tian, Z. Cui, H. Wei, and C. Li, An intelligent quantitative risk assessment method for ammonia synthesis process, *Chemical Engineering Journal*, vol. 420, p. 129893, 2021.

[113] W. Tian, N. Liu, D. Sui, Z. Cui, Z. Liu, J. Wang, H. Zhou, and Y. Zhao, Early warning of Internal leakage in heat exchanger network based on dynamic mechanism model and long short-term memory method, *Processes*, vol. 9, no. 2, p. 378, 2021.

[114] C. Hou, J. Wu, B. Cao, and J. Fan, A deep-learning prediction model for imbalanced time series data forecasting, *Big Data Mining and Analytics*, vol. 4, no. 4, pp. 266–278, 2021.

[115] Z. Xue and H. Wang, Effective density-based clustering algorithms for incomplete data, *Big Data Mining and Analytics*, vol. 4, no. 3, pp. 183–194, 2021.

[116] A. H. Zhan, Y. Sang, Y. Sun, and J. Lv, A neural network learning algorithm for highly imbalanced data classification, *Information Sciences*, vol. 612, pp. 496–513, 2022.

[117] P. Xu, R. Du, and Z. Zhang, Predicting pipeline leakage in petrochemical system through GAN and LSTM, *Knowledge-Based Systems*, vol. 175, pp. 50–61, 2019.

[118] B. L. Shao, X. L. Hu, G. Q. Bian, and Y. Zhao, A multichannel LSTM-CNN method for fault diagnosis of chemical process, *Mathematical Problems in Engineering*, vol. 2019, pp. 1–14, 2019.

[119] W. Tian, S. Wang, S. Sun, C. Li, and Y. Lin, Intelligent prediction and early warning of abnormal conditions for fluid catalytic cracking process, *Chemical Engineering Research and Design*, vol. 181, pp. 304–320, 2022.

[120] J. Chen, G. Wang, J. Lv, Z. He, T. Yang, and C. Tang, Open set classification for signal diagnosis of machinery sensor in industrial environment, *IEEE Transactions on Industrial Informatics*, doi: 10.1109/TII.2022.3169459.

[121] D. Vitkus, I. Steckeviius, N. Goranin, D. Kalibatien, and A. Enys, Automated expert system knowledge base development method for information security risk analysis, *International Journal of Computers Communications & Control*, vol. 14, no. 6, pp. 743–758, 2020.

[122] D. Q. Zhu and S. L. Yu, Survey of knowledge-based fault diagnosis methods, (in Chinese), *Journal of Anhui University of Technology*, vol. 19, no. 3, pp. 197–204, 2002.

[123] J. A. Carrasco and V. Sune, An algorithm to find minimal cuts of coherent fault-trees with event-classes, using a decision tree, *IEEE Transactions on Reliability*, vol. 48, no. 1, pp. 31–41, 1999.

[124] S. Ni, Y. F. Zhang, H. Yi, and X. F. Liang, Intelligent fault diagnosis method based on fault tree, (in Chinese), *Journal of Shanghai Jiaotong University*, vol. 42, no. 8, pp.1372–1375, 2008.

[125] D. N. Chen, J. Y. Xu, C.Y. Yao, H. Y. Pan, and Y. L. Hu, Continuous-time multi-dimensional T-S dynamic fault tree analysis methodology, (in Chinese), *Journal of Mechanical Engineering*, vol. 57, no. 10, pp. 231–244, 2021.

[126] C. K. Li, C. L. Wang, X. J. Gao, and J. Zhu, Abnormal situations monitoring and early warning system of petrochemical process, (in Chinese), *Safety Health & Environment*, vol. 15, no. 11, pp. 8–13, 2015.

[127] D. Gao, C. G. Wu, B. Zhang, and X. Ma, Signed directed graph and qualitative trend analysis based fault diagnosis in chemical industry, *Chinese Journal of Chemical Engineering*, vol. 18, no. 2, pp. 265–276, 2010.

[128] H. Vedam and V. Venkatasubramanian, Signed digraph based multiple fault diagnosis, *Computers & Chemical Engineering*, vol. 21, pp. S655–S660, 1997.

[129] W. Tian, S. Zhang, Z. Cui, Z. Liu, S. Wang, Y. Zhao, and H. Zhou, A fault identification method in distillation process based on dynamic mechanism analysis and signed directed graph, *Processes*, vol. 9, no. 2, p. 229, 2021.

[130] F. Yang and D. Y. Xiao, Review of SDG modeling and its application, (in Chinese), *Control Theory & Applications*, vol. 22, no. 5, pp. 767–774, 2005.

[131] Z. Zhang, C. Wu, B. K. Zhang, T. Xia, and A. F. Li, SDG multiple fault diagnosis by real-time inverse inference, *Reliability Engineering & System Safety*, vol. 87, no. 2, pp. 173–189, 2005.

[132] Y. X. Dong, L. N. Li, and W. Tian, A novel fault diagnosis method based on multilayer optimized PCC-SDG, *CIESC Journal*, vol. 69, no. 3, pp. 1173–1181, 2018.

[133] IEEE, Framework of knowledge graphs: IEEE2807-2022[S/OL], https://standards.ieee.org/ieee/2807/7525/, 2022.

[134] J. Q. Tang, F. Yang, J. M. Jin, and C. S. Zhang, Causal link analysis method and device and computer readable storage medium, China Patent CN202010618648.0, 2020.

[135] F. Yang, Q. F. Kuang, B. B. Jin, and C. S. Zhang, Chemical reaction path acquisition method and device, electronic device, and storage medium, China Patent CN201810681778.1, 2018.

[136] X. Fan and F. Yang, Graph data query method and device and storage medium, (in Chinese), China Patent CN202010579437.0, 2020.

[137] J. Lee, C. Yoo, S. W. Choi, and P. A. Vanrolleghem, Nonlinear process monitoring using kernel principal component analysis, *Chemical Engineering Science*, vol. 59, no. 1, pp. 223–234, 2004.

[138] H. Z. Chen, A. J. Zhao, T. J. Li, C. C. Cai, S. Cheng, and C. L. Xu, Fuzzy Bayesian network inference fault diagnosis of complex equipment based on fault tree, (in Chinese), *Systems Engineering and Electronics*, vol. 43, no. 5, pp. 1248–1261, 2021.

[139] Z. Gao, T. Breikin, and H. Wang, Discrete-time proportional and integral observer and observer-based controller for systems with both unknown input and output disturbances, *Optimal Control Applications and Methods*, vol. 29, no. 3, pp. 171–189, 2008.

[140] S. J. Ahn, C. J. Lee, Y. Jung, C. Han, E. S. Yoon, and G. Lee, Fault diagnosis of the multi-stage flash desalination process based on signed digraph and dynamic partial least square, *Desalination*, vol. 228, no. 1, pp. 68–83, 2008.

[141] N. Lv and X. Wang, Fault diagnosis based on signed digraph combined with dynamic kernel PLS and SVR, *Industrial & Engineering Chemistry Research*, vol. 47, no. 23, pp. 9447–9456, 2008.

Fan Yang received the PhD degree in electronics and information from Sichuan University, China in 2022. Now he is working as a principal scientist and head of the Algorithm and Big Data Center, New Hope Liuhe Co Ltd, China. His research interests include neural networks, machine learning, big data and their application in industry.

Mao Xu received the MEng degree in information science and technology from Southwest Jiaotong University, China in 2019. He is currently an algorithm researcher at the Algorithm and Big Data Center, New Hope Liuhe Co Ltd, China. His main research interest is artificial intelligence methods for agriculture and animal husbandry.

Jiancheng Lv received the PhD degree in computer science and engineering from the University of Electronic Science and Technology of China, Chengdu, China in 2006. He was a research fellow at the Department of Electrical and Computer Engineering, National University of Singapore, Singapore. He is currently a professor at the Data Intelligence and Computing Art Laboratory, College of Computer Science, Sichuan University, Chengdu, China. His research interests include neural networks, machine learning, and big data.

Wenqiang Lei received the PhD degree in computer science from National University of Singapore, Singapore in 2019. He is currently a professor at Sichuan University, China. His research interests include natural language processing, computational linguistics, computational musicology, and computational social science.