Energy 203 (2020) 117893

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Review

Review: Energy efficiency evaluation of complex petrochemical industries

Yongming Han [a, b], Hao Wu [a, b], Zhiqiang Geng [a, b, *], Qunxiong Zhu [a, b, **], Xiangbai Gu [a, b, c], Bin Yu [d]

[a] College of Information Science and Technology, Beijing University of Chemical Technology, Beijing, 10029, China
[b] Engineering Research Center of Intelligent PSE, Ministry of Education in China, Beijing, 10029, China
[c] Sinopec Engineering (Group) Co., Ltd., Beijing, 100029, China
[d] Hengli Petrochemical Co. LTD, Dalian, Liaoning, 116000, China

A R T I C L E I N F O

Article history:
Received 20 February 2020
Received in revised form
14 May 2020
Accepted 15 May 2020
Available online 18 May 2020

Keywords:
Energy efficiency evaluation
Mechanism model
Data-driven
Domain knowledge
Complex petrochemical industry

A B S T R A C T

As the most effective indicator for energy saving and emission reduction, energy efficiency evaluation is widely used in complex petrochemical industries. It is nowadays common to combine traditional mechanism methods based on momentum transport, energy transport, quality transport (TT) and reaction engineering (RG) (TT-RG), with data-driven artificial intelligence methods. Using the combined method to achieve production optimization and energy saving by analyzing the evaluation indicator of energy efficiency has gradually become an important part in complex petrochemical industries. Therefore, this paper introduced the main methods and the latest research results of energy efficiency evaluation of complex petrochemical industries. These methods are mainly divided into three parts, including the mechanism methods based on TT-RG, the data-driven artificial intelligence methods, and the hybrid methods combining the mechanism and the data-driven. Then, different methods are compared and described in detail. Moreover, the best method for evaluating the energy efficiency can be found to provide theoretical guidance for energy saving and emission reduction of complex petrochemical industries. Finally, the future development direction for energy efficiency evaluation in complex petrochemical industries is given.

Contents

* Corresponding author. College of Information Science and Technology, Beijing University of Chemical Technology, Beijing, 10029, China.
** Corresponding author. College of Information Science and Technology, Beijing University of Chemical Technology, Beijing, 10029, China.
E-mail addresses: gengzhiqiang@mail.buct.edu.cn (Z. Geng), zhuqx@mail.buct.edu.cn (Q. Zhu).

https://doi.org/10.1016/j.energy.2020.117893
0360-5442/

1. Introduction

With the continuous development of economy, China has become the second largest economy nation in the world. As a pillar of modern economy, energy has greatly promoted the economy development. However, problems of environmental pollution and high energy consumption become more and more serious. The comparison of carbon dioxide (CO_2) emissions and energy consumption between China and other countries (or regions) in 2018 are shown in Fig. 1 and Fig. 2, respectively, where the data are from the World Energy Statistics Yearbook. It can be seen from Figs. 1 and 2 that improving the energy efficiency and reducing pollutants emission have been in a highly strategic position for the sustainable development of all countries in the world [1]. In the process of energy utilization, the ratio of the working energy to the actual consumed energy is the energy efficiency. How to improve the energy efficiency is the main goal of the energy development. The energy efficiency evaluation can provide the most effective indicator for this goal, which explores ways to improve the energy efficiency by establishing mathematical models of production processes, or to reach the same or more output with less energy input [2]. The complex petrochemical industry in China is a high energy consumption industry with large scales. Therefore, by evaluating the energy efficiency, finding efficient ways of energy utilization, and then reducing pollutants emission such as CO_2 are hot research topics for energy saving and emission reduction [3]. At the same time, it has far-reaching significance for the sustainable development of economy and society [4].

The complex petrochemical industry, which has characteristics of large-scale equipment, large production scale, strong continuity, high-dimension and noise data, not only is an important industry related to national production and life but is a high pollution and high carbon emissions industry. Due to the continuous expansion of energy application in the complex petrochemical industry, the information is coupled and intersected, and the structure becomes complicated in the system [5]. In addition, because of the multi-operation factors and their interrelationships in the energy transfer system, it is difficult to describe the process in a reasonable way by traditional mechanism methods based on momentum transport, energy transport, quality transport (TT) and reaction engineering (RG) (TT-RG) [6]. The TT-RG methods are mainly based on a direct mathematical description of the actual process including the material balance and thermal balance, the laws of thermodynamics and energy balance and the specific chemical process flow with thermodynamic analysis. Therefore, they are not only increased the difficulty of modeling but have the shortcoming of the long cycle analysis, low efficiency and heavy manual workload [7].

With the rapid development of big data and neural network, and breakthrough of key technologies, more and more scholars at home and abroad have adopted data-driven methods to evaluate the energy efficiency of complex petrochemical industries which included the fuzzy analytic hierarchy process [8], the DEA (data envelopment analysis) method [9], respected classification method [10] and extreme learning machine (ELM) neural network [11]. The data-driven methods can learn and make decisions independently, which make the researcher classify and extract the key factors for logistics, energy flows, and information flows in complex and dynamic petrochemical production systems [12]. The data-driven methods made up these shortcomings of traditional mechanism methods and greatly simplified the energy efficiency analysis process [13]. The data-driven methods mainly include data-driven statistical methods, data-driven artificial intelligence methods and data-driven hybrid methods. The data-driven statistical methods include the regression analysis, the classification and the clustering method. The data-driven artificial intelligence methods usually include various artificial neural networks (ANNs) and data-

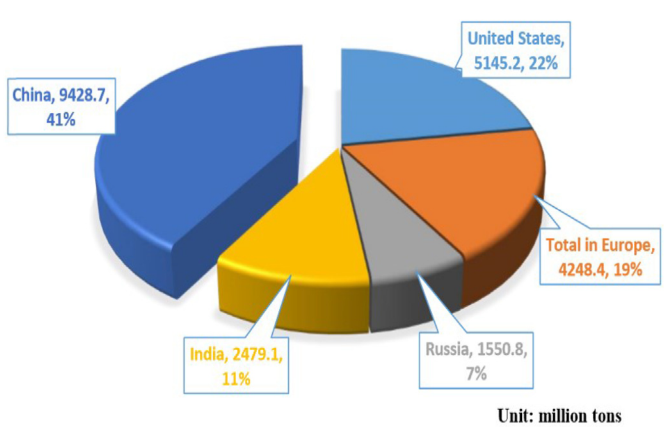

Fig. 1. CO_2 emissions comparison between China and other countries (or regions) in 2018.

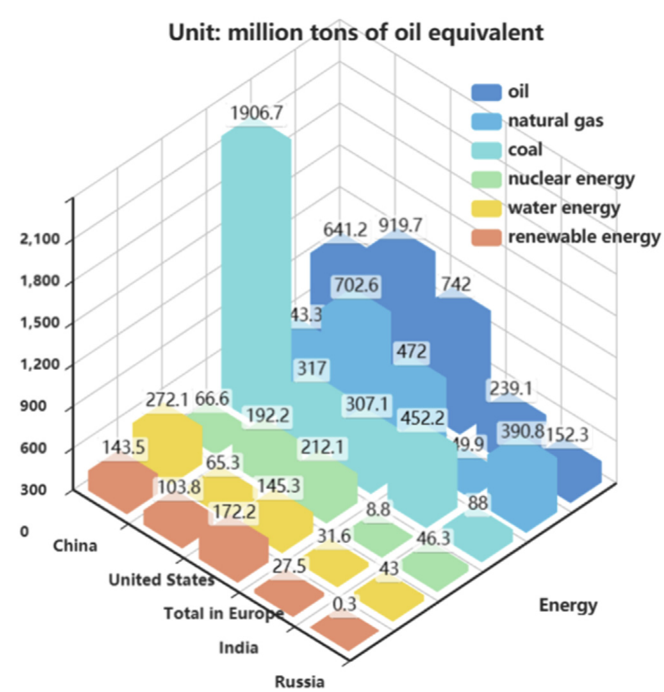

Fig. 2. Energy consumption comparison between China and other countries (or regions) in 2018.

driven hybrid methods which combine statistical methods and ANNs. With the development of the ANN, deep learning methods also began to applied in energy efficiency evaluation for its powerful learning ability, such as the convolutional neural network (CNN), which allows the computer to learn pattern features automatically and incorporates features into the modeling process to reduce the incompleteness caused by traditional feature extraction methods.

However, a large number of data-driven methods are mainly based on the relationship between input and output variables in the experiment and the neural network with simple structures, which are more effective within the scope of the experiment, so it will be not widely extended [14]. Therefore, the methods combined mechanism methods and data-driven methods were adopted by some scholars, which included the prediction of carbon content in converter end point [15], the endpoint control model of converter steelmaking [16], the control technology of converter blowing and limestone residue behavior [17] and the steel industry multi-type energy optimized scheduling [18]. Although energy efficiency evaluation methods in complex petrochemical industries have gradually become a hot topic in the research, it still faces great challenges [19]. This paper summarized the research work of energy efficiency evaluation methods in complex petrochemical industries in recent years. On the basis of mathematical models, these methods are divided into three classes including the mechanism methods based on the TT-RG, the data-driven methods and the mixed methods combined the mechanism and the data-driven. Then these methods are introduced, compared and summarized comprehensively. Finally, summing up the current research situation, this paper presents problems and future directions of energy efficiency evaluation methods in complex petrochemical industries.

The rest of the chapters are arranged as follows: section 2 introduces the energy efficiency evaluation methods including mechanism methods based on TT-RG, data-driven methods and the methods combining mechanism and data-driven. The discussion is described in section 3. section 4 and section 5 present the development in the future and the conclusion, respectively.

2. Energy efficiency evaluation methods

With the development of complex petrochemical industries, the energy consumption scale increases continuously. Energy saving and emission reduction become the main goal of the complex petrochemical industry. Therefore, how to improve the energy production, and reduce the energy consumption and cost are the key problems in domestic and abroad [20]. High-effective energy efficiency evaluation in complex petrochemical industries is undoubtedly an important way to improve the output and reduce the energy consumption [21].

At present, there are many researches on energy efficiency evaluation of the complex petrochemical industry. In this paper, energy efficiency evaluation methods are mainly divided into three categories: mechanism methods based on TT-RG, data-driven methods and methods combining mechanism and data-driven. Based on this classification, the main evaluation methods are analyzed, sorted and summarized as shown in Fig. 3.

2.1. The mechanism methods based on TT-RG

The mechanism method is established by equations which describe the process of mechanism and then verified by experiments [22]. It is a direct mathematical description of the actual process and also reflection of the process essence. Therefore, it is easy to generalize [23]. There are three steps for the mechanism method to describe the petrochemical production process. First, reasonable assumptions are proposed according to the selected method. On the premise of satisfying the application need, understanding of objects and ignoring of the secondary factors are considered. Second, according to the internal mechanism of the petrochemical production process, the balance relation of material, energy and momentum, and the chemical equation constitute the mathematical model. Finally, as long as satisfying the control engineering of the petrochemical industry, the proposed method is simplified as much as possible [24]. In the complex petrochemical industry, the mechanism method generally has the following aspects with one based on the material balance and thermal balance to establish a mathematical relationship, one based on the laws of thermodynamics and energy balance to establish a mathematical relationship, and another combining specific chemical process flow with thermodynamic analysis to establish the mathematical relationship. The characteristics, advantages and disadvantages of these above methods are shown in Table 1.

Elisa et al. [25] proposed a thermal fluid dynamics model to reduce effectively primary energy consumption. Combining dry quenching systems with waste heat utilization systems, Cheng et al. [26] calculated and analyzed the coking process of coke oven based on material balance and heat balance to realize the recycling of leftover heat and waste heat, improve coke products and byproducts energy, and further increase economic and environmental benefits. Han et al. [27] studied the method of non-catalytic thermal conversion reaction of aromatic hydrocarbons in aromatic liquids adopted by domestic and foreign scholars in recent years, summarized the law of thermal conversion reaction of aromatic hydrocarbons, and reduced raw materials consumption. As a typical industry of the petrochemical industry, the ethylene industry is a complex process with multiple inputs, as shown in Fig. 4. A simple proportional relationship could not fully evaluate its energy efficiency [28]. Therefore, Hua et al. [29] proposed a "three-link" analysis method based on the thermodynamic law, which mean that the craft flow process was divided into the energy conversion subsystem, utilization subsystem and recycling subsystem. At the same time, the energy balance equation was established by analyzing the subsystem data with the thermodynamic law, and further the energy loss was analyzed. From the perspective of the ethylene production plant, Yan et al. [30] proposed relevant optimization measures to reduce the energy consumption and improve the overall utilization rate of the ethylene plant. Since then, based on the extensive status of CO_2 in the smelting flue gas of magnesite, some scholars carried out thermodynamic analysis on the thermal parameters of each node in the process, found the greatest irreversible loss in the process, and pointed out the direction for energy-saving and optimization of the process industry, such as the CO_2 capture system [31], the CO_2 hydrogenation to the methanol system [32] and characteristic and model of the thermal decomposition [33]. Zhang et al. [34] summarized the catalytic cracking application in the production of diesel oil and gasoline, reducing the olefin content in petroleum products based on the process flow of catalytic cracking in petroleum refining.

The mechanism method is an accurate mathematical method based on the internal mechanism of objects, the production process or the transfer mechanism of materials flow. It has advantages of clear physical meaning of parameters, strong adaptability, and can be largely extended. However, this method also has disadvantages, such as difficulty to determine the mathematical expression of some complex objects, and needs of a large number of parameters. If these parameters cannot be well obtained, the modeling result will be affected [35]. Besides, the energy efficiency evaluation based on mechanism methods spend much time and large manual workload, which make this method inefficient in evaluating the

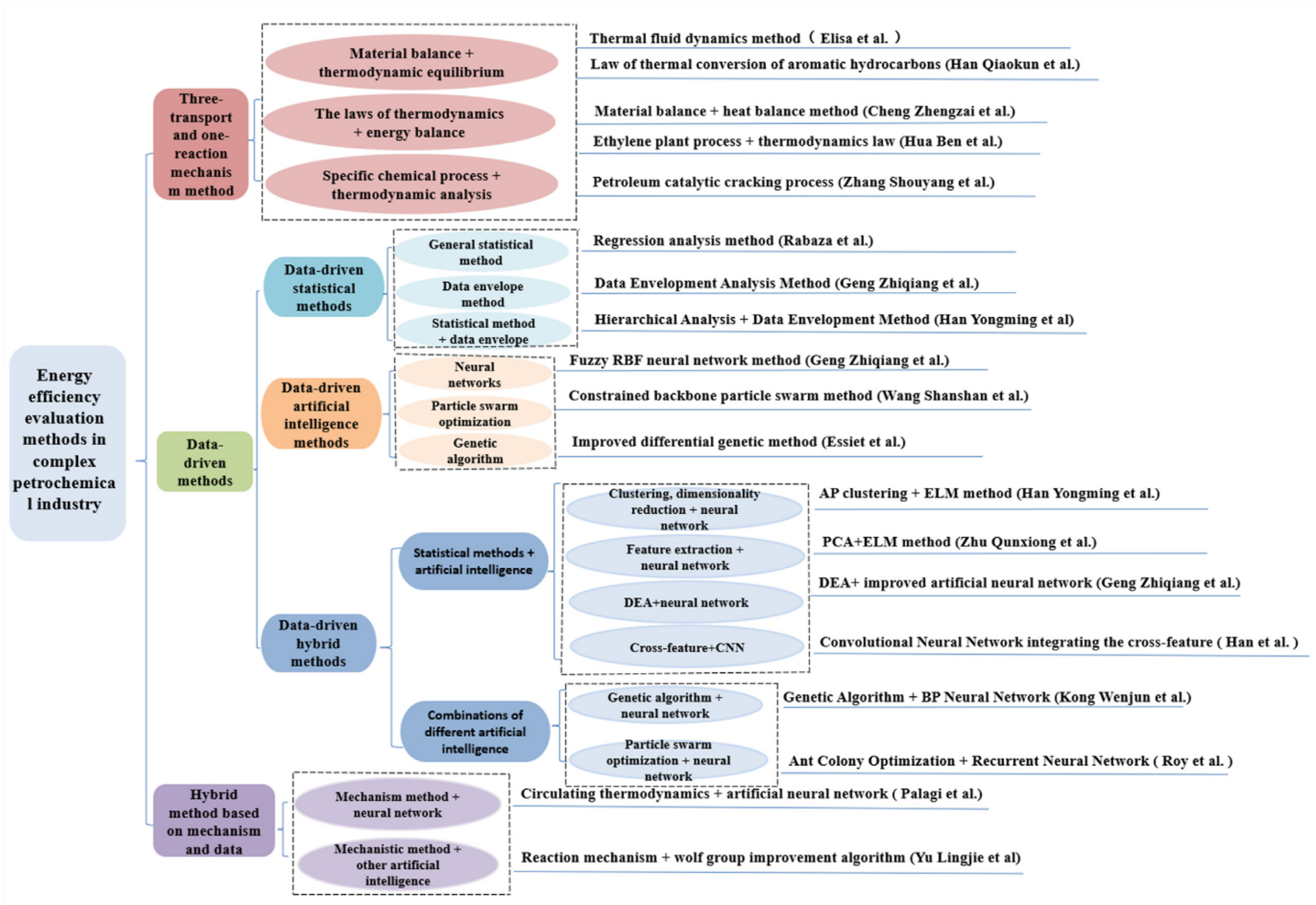

Fig. 3. The classification of energy efficiency evaluation methods in complex petrochemical industries.

Table 1
Comparison of mechanism evaluation methods.

Mechanism method	Core idea	Advantage	Disadvantage	Stability
Based on material balance and thermodynamic equilibrium	All substances and products in the reaction process are mass and heat distributed according to the reaction	Quantitatively grasp important parameters in the reaction process and discover weakness	High cost and difficulty.	Depend on whether there is a spontaneous reaction trend during the reaction
Based on the laws of thermodynamics and energy balance	Analyze subsystem data using thermodynamic laws and establish energy balance equations	Wide range of application	Energy balance is obtained based on complete material balance	Depend on the correctness of physical property data such as heat capacity and specific heat capacity
Combing specific chemical process with thermodynamic analysis	Parameters are determined according to the process flow and analyzed by the law of heat	Reduce the energy consumption to a certain extent under certain load	Accurately control the progress of the reaction	Generally stable

energy efficiency. Therefore, these mechanism methods are suitable for the condition of clear parameters and simple processes.

2.2. The data-driven methods

With the development of artificial intelligence and big data, data-driven methods have gradually become a very common modeling tool, which build the model based on the actual data [36]. At present, data-driven methods mainly include data-driven statistical methods, data-driven artificial intelligence methods and data-driven hybrid methods. The data-driven hybrid methods include statistical methods integrating with the neural network and the neural network integrating other artificial intelligence methods. The characteristics, advantages and disadvantages of these methods are shown in Table 2. In the following, detailed classification and comparison of these above three methods will be analyzed, and emphasis on the important method will be laid.

2.2.1. Data-driven statistical methods

Data-driven statistical methods are suitable for nonlinear and

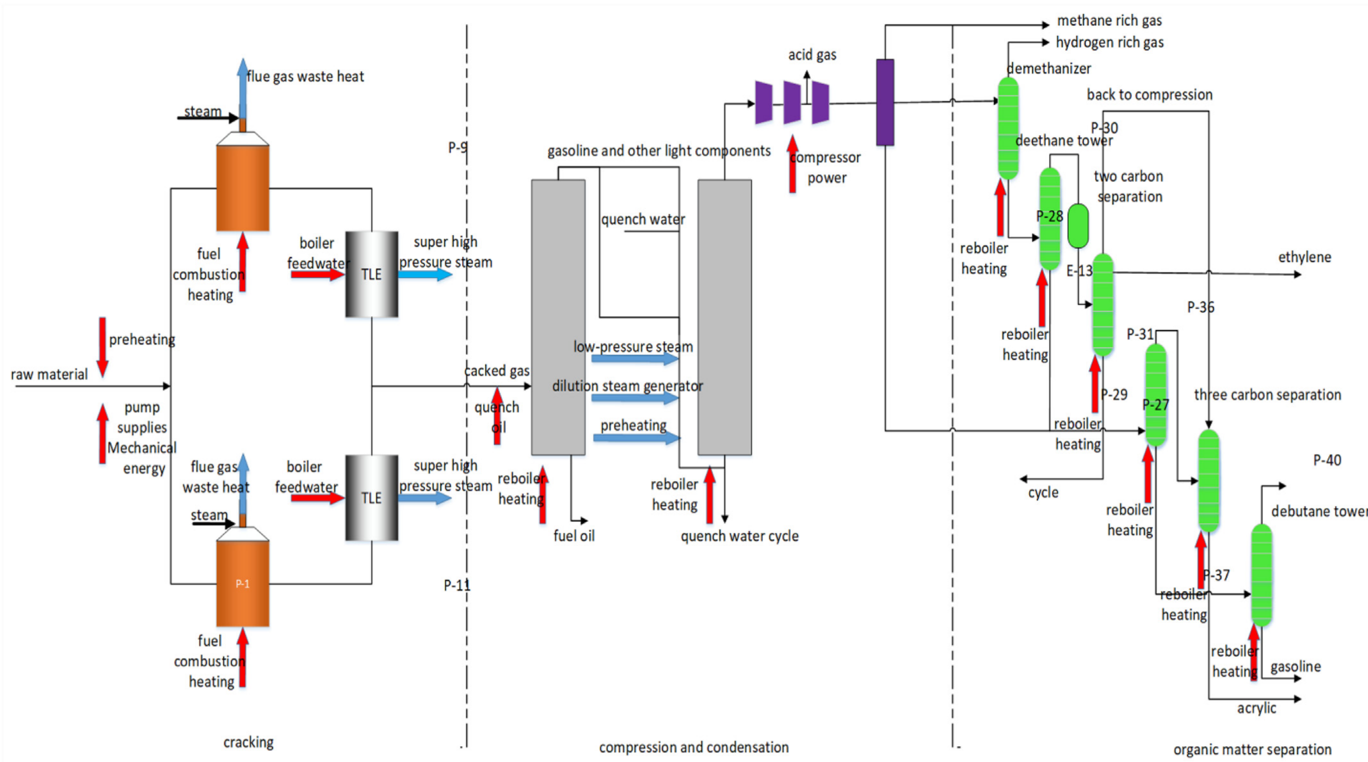

Fig. 4. The process of the ethylene production system.

Table 2
Comparison of data-driven evaluation methods.

Data-driven method	Idea	Advantage	Disadvantage	Stability
Data-driven statistical method	Easily process the data by statistical methods	Suitable for nonlinear and uncertain systems with lower data dimension	Difficult to handle high-dimensional, redundant or messy data	Weak
Data-driven artificial intelligence method	Directly use artificial intelligence methods such as neural networks for modeling and analysis	Process high-dimensional redundant data	When the data is too large, redundant or messy, its performance will be affected.	Medium strong
Data-driven hybrid method	Filter and extract the data by statistical methods and then modeling by neural networks	Handle high dimensional redundant data better	The effect is within the scope of the experiment, and it should not be promoted to a large extent	Strong

uncertain systems. In the complex petrochemical industry, the statistical methods are mainly divided into three aspects with the general statistical method, the DEA methods and the DEA combined with statistical methods. General statistical methods include the regression analysis method, support vector machine (SVM) method and so on. Regression analysis includes linear and nonlinear regression. Linear regression means the input and output variables are linear correlation, that is to say, the relationship between input and output variables are satisfying a multivariate linear equation which can be shown in Eq. (1). Where the $\hat{y}$ is prediction value, x is the sample data, w and b are the regression coefficient. And then, by constructing the loss function, the parameters w and b can be got when the loss function is minimum. The express of loss function as Eq. (2) described. Where the y is actual value. Finally, by least square method or gradient descent method to minimize the loss function and then get the w and b.

$$\hat{y} = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n + b \tag{1}$$

$$L = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \tag{2}$$

There are few statistical methods related with energy efficiency functions and energy implementation based on regression analysis presented by scholars [37,38]. However, the linear regression cannot describe the nonlinear characteristics of complex processes comprehensively. As usual, the nonlinear regression problem is transformed into a linear problem to be solved, which results in poor robustness in fitting complex nonlinear relations [39].

The SVM method has been widely applied in the petrochemical field at home and abroad in recent years for modeling of steam cracking of naphtha [40], predicting water source heat pump energy demand [41] and Forecasting of consumers heat load in district heating systems [42]. This method transforms the machine learning problem into a quadratic programming problem, which can obtain the best solution. The SVM can get different classifiers and function approximators by selecting different kernel functions and parameters [43], more flexible in design and has good modeling effect. In the sample classification space, the hyperplane equation of the SVM is described as Eq. (3).

$$w^T \cdot x + b = 0 \tag{3}$$

Where, the w is the normal vector of the hyperplane, b is the offset from the test origin to the hyperplane, and x is the sample data. The hyperplane can be described as Eq. (4). Therefore, when the distance between two support vector planes is 2b, the plane separation is maximum.

$$-1 \le w^T \cdot x + b \le 1 \tag{4}$$

Yao et al. [44] proposed an improved SVM based on the feature expansion to simplify the selection of kernel functions, and used the seismic data from the Guanyin field in Sichuan to predict the oil and gas. Compared with the traditional prediction method, the result showed that the error rate was reduced by about 50%. Taking coal characteristics and the oven heating system as reference, Cui et al. [45] used the SVM to predict the coke quality, testifying that this method has high accuracy and strong generalization ability for coke quality prediction. However, these methods are difficult to train the large-scale data. Considering the whole production system of the complex petrochemical industry, the energy efficiency is a multi-index evaluation result that integrates various inputs and outputs [46]. Taking the ethylene production as an example, the weights of input and output cannot be determined in calculating the energy efficiency, making the general multi-index evaluation method based on the SVM not suitable. The DEA, serving as a non-parametric statistical and classical performance evaluation method [47], is widely used in the complex petrochemical industry. It evaluates decision making units (DMUs) of multi-input and multi-output of the same type based on the relative efficiency [48]. The principle of the DEA method is to evaluate and calculate weights of inputs and outputs of each DMU, so as to determine whether each DMU is effective [49].

The calculation steps are shown as follows: supposing there are n DMUs, and each DMU has m inputs and s outputs, respectively. For the jth DMU, the input vector is $x_j = (x_{1j}, x_{2j}, \ldots x_{mj}) > 0$, and x_{ij} represents the ith input of DMU_j. The output vector is $y_j = (y_{1j}, y_{2j}, \ldots y_{sj}) > 0$ $y_j = (y_{1j}, y_{2j}, \ldots y_{sj}) > 0$, and y_{ks} represents the kth output of DMU_j. θ_j is the efficiency value of DMU_j as shown in Eq. (5).

$$\theta_j = \frac{\sum_{k=1}^{s} u_k y_{kj}}{\sum_{i=1}^{m} v_i x_{ij}} \tag{5}$$

Where, u_k and $v_i v_i$ are the weight coefficients of the kth output and the ith input of DMU_j, respectively. There is no fixed expression of the weight coefficient, which is a maximum value under the actual situation. If the efficiency value θ_j cannot reach the optimal value of 1, this DMU_j is considered invalid. Therefore, the DEA method can improve the energy efficiency of complex petrochemical industries by adjusting the production structure timely [50].

The DEA structure model is shown in Fig. 5. However, in the complex petrochemical industry, it is difficult to analyze the effectiveness of multiple DMUs based on the traditional DEA method, which does not consider the impact of the uncertain data. Therefore, many experts and scholars have improved the DEA method and combined it with statistical methods to evaluate the energy efficiency.

Azadeh et al. [51] used the principal component analysis (PCA) and other methods to verify the effectiveness of the DEA in energy efficiency evaluation and energy saving, and introduced structural indicators to analyze the energy consumption for energy-intensive manufacturing industry. Bian et al. [52] extended a radial random DEA method based on probability constraints to a non-radial method, to evaluate the energy efficiency and CO_2 emission in China. Fan et al. [53] used the log-mean divisia index (LMDI) decomposition method to evaluate the main influencing factors of carbon emission of China petrochemical industry, and pointed out that the economic growth was the main factor and the industrial structure had a certain impact on carbon emission. Then the carbon emission could be reduced by reasonably balancing the allocation of these influencing factors. The DEA evaluation method combined with the statistical method is widely studied and analyzed by scholars which including the novel DEA [54], the DEA cross-model integrated analytic hierarchy process [55] and the DEA integrated analytic hierarchy process [56]. These evaluation methods can guarantee the consistency and precision of the data, and improve the efficiency of the DEA for energy efficiency evaluation. In addition to the evaluation methods combining general statistical methods with the DEA, there are other evaluation methods combining different statistical methods. For example, Wang et al. [57] preprocessed the raw data with the PCA method to effectively extract the characteristic information of the data. The collinearity between variables was eliminated, and the optimization and evaluation model of the energy efficiency was established with the SVM.

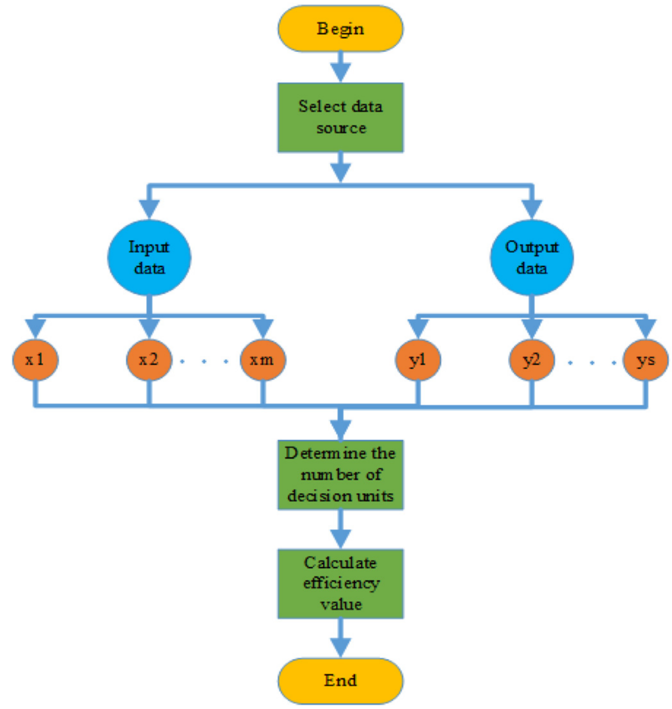

Fig. 5. The DEA structure model.

The statistical method is a general data analysis method applied in all fields, especially in the complex petrochemical industry. Due to the large production scale, a large amount of data is generated. The data-driven statistical method has a good effect on the modeling process in the complex petrochemical industry [58]. However, due to the strong intersection and high spatial dimension of the raw data in complex petrochemical industries, these statistical methods have slow response speed and poor approximation ability to nonlinear models.

2.2.2. Data-driven artificial intelligence methods

Due to the popularity of artificial intelligence, more and more artificial intelligence methods have been used in mechanical engineering fields [59]. In the complex petrochemical industry, data-driven artificial intelligence methods are mainly based on neural networks to evaluate the energy efficiency [60]. There are many characteristics such as high dimensionality and nonlinearity in petrochemical processes, which makes the modeling process complex [61]. The ANN has been widely used in energy efficiency evaluation of petrochemical processes [62] for its fast response speed [63], without considering internal mechanism [64] and strong nonlinear approximation ability [65]. In complex petrochemical industries, taking the ethylene production for example, the resources that produce ethylene and other products are served as inputs of the neural network, and the ethylene yield or other by-products yields are as outputs, forming the evaluation model of the energy efficiency. By analyzing the predicted value and the real value of the ethylene yield, the input of raw material is adjusted, and the energy efficiency is analyzed and evaluated [66].

Among all the neural networks, the back propagation (BP) neural network and radial basis function (RBF) neural network are classical neural network, which have been applied in the complex petrochemical industry for energy efficiency evaluation including the optimization of the reaction dehydration to ethylene [67], the modeling of atmospheric and vacuum distillation unit [68], the prediction of 350 °C content in atmospheric distillation of refinery [69], the multi-objective optimization of rectification process [70] and the modeling of production prediction and energy-saving [71]. The BP neural network adopts gradient descent algorithm to gradually approach the minimum error of actually value and prediction value by continuously adjusting the weights of neurons [72]. Its learning rate is fixed, so the convergence speed of the network is slow and easy to fall into local extremum [73]. In addition, for some complex problems, the training time of the network will be increased [74]. The RBF neural network is an efficient feedforward network, which has the best approximation performance and global optimal characteristics that other feedforward networks do not have. In addition, its structure is simple and the training speed is fast [75], but its generalization ability is weak. Due to the insufficiency of the BP neural network and the RBF neural network, the evaluation is not accurate in energy efficiency analysis in the petrochemical industry and has limitations in improving the energy efficiency [76]. In 2004, Huang et al. [77] proposed the ELM neural network, which could solve the above problems well and was widely used in various fields.

The ELM neural network can approximate any linear and nonlinear functions almost with zero error [78], and the network has excellent generalization ability, simple structure and fast training speed [79]. It is a simple and effective feedforward neural network with a single hidden layer. The weight between the input layer and the hidden layer is directly obtained by Gaussian distribution in the training process, while the weight between the hidden layer and the output layer is calculated by generalized inverse of the vector [80]. Therefore, the ELM neural network has advantages of easy parameter selection [81], fast learning speed and strong robustness [82]. The training process of the ELM neural network is shown as follows: suppose the input layer contains n vectors, the hidden layer contains L nodes (L is much smaller than n), and the output layer contains m vectors. The optimal solution of the ELM can be obtained by setting the number of nodes in the hidden layer. The connection weights between the input layer and the hidden layer, the hidden layer and the output layer are W and V, respectively.

$$W=\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \vdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{L1} & w_{L2} & \cdots & w_{Ln} \end{bmatrix}_{L\times n} \tag{6}$$

$$\mathrm{W}=\begin{bmatrix} \mathrm{w}_1 \\ \mathrm{w}_2 \\ . \\ . \\ . \\ \mathrm{w}_L \end{bmatrix} = \begin{bmatrix} \mathrm{w}_{11} & \mathrm{w}_{12} & . & . & . & \mathrm{w}_{1n} \\ \mathrm{w}_{21} & \mathrm{w}_{22} & . & . & . & \mathrm{w}_{2n} \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ \mathrm{w}_{L1} & \mathrm{w}_{L2} & .. & . & . & \mathrm{w}_{Ln} \end{bmatrix}_{L\times n} \tag{7}$$

$$V=\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_L \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1m} \\ v_{21} & v_{22} & \cdots & v_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ v_{L1} & v_{L2} & \cdots & v_{Lm} \end{bmatrix}_{L\times m}$$

The threshold of the hidden layer is:

$$\mathrm{b}=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_L \end{bmatrix} \tag{8}$$

The activation function is $g(x)$, and the output including n samples can be expressed as:

$$Y=HV \quad \mathrm{Y}=\mathrm{HV} \tag{9}$$

$Y\in R^{m\times n}$. H: The output matrix of the hidden layer, and the expression of H is shown as follows:

$$\mathrm{H}=\begin{bmatrix} g(w_1x_1+b_1) & g(w_2x_1+b_2) & . & . & . & g(w_Lx_1+b_L) \\ g(w_1x_2+b_1) & g(w_2x_2+b_2) & . & . & . & g(w_Lx_2+b_L) \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ g(w_1x_m+b_1) & g(w_2x_m+b_2) & . & . & . & g(w_Lx_m+b_L) \end{bmatrix}$$

$$H=\begin{bmatrix} g(w_1x_1+b_1) & \cdots & g(w_Lx_1+b_L) \\ \vdots & \ddots & \vdots \\ g(w_1x_m+b_1) & \cdots & g(w_Lx_m+b_L) \end{bmatrix} \tag{10}$$

W and b are arbitrarily specified before training, and stay unchanged during the training process. It can be known from Eqs. (7), (9) and (10) that the output weight of the hidden layer is:

$$\hat{\mathrm{v}} = H^{+}Y \tag{11}$$

Where $\mathrm{H}^{+}H^{+}$ is the generalized inverse of the output matrix of the hidden layer. The ELM simplifies the complex training process of the neural network and changes it into the problem of matrix inverse, which greatly improves the learning speed [83]. Because of the simple network structure and strong generalization ability of the ELM, it is widely used in energy efficiency evaluation of the complex petrochemical industry and other process industries including the prediction of flood gas velocity in packed tower liquid

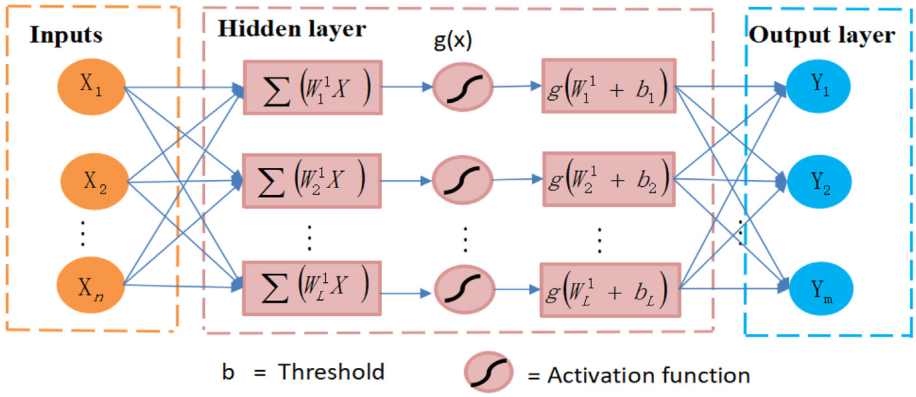

Fig. 6. The structure of the ELM.

[84], the modeling of near infrared spectroscopy [85], the quantitative analysis of near infrared spectroscopy [86] and the kinetic separation of propylene over propane in a microporous meta [87]. The structure of the ELM is shown in Fig. 6.

Data-driven artificial intelligence methods also include the genetic algorithm (GA) and the particle swarm optimization (PSO) method. The GA is an iterative optimization algorithm, which takes variables to be optimized in practical problems as the initial population, and then calculates the fitness of each variable in the population [88]. After selection, crossover and mutation, the new population is generated, and then the new fitness value is calculated again. After multiple iterations, the individual with the maximum fitness is finally obtained, which is also the optimal solution [89]. The PSO is also an iterative optimization algorithm [90]. First, a group of random solutions is initialized and the corresponding fitness values are calculated. According to the fitness values, the optimal local and global values are found, and the velocity and the position of each particle are updated. After the iteration, the optimal solution is obtained [91]. The procedure of data-driven GA or PSO method in energy efficiency evaluation of complex petrochemical industries is shown as follows: all the raw materials of a certain chemical product are used as the initial population or a group of particles to be optimized. Then the individual (i.e. the raw material) with the greatest fitness or the global optimal value is found. Finally, the selected materials of a certain chemical product are used to predict and evaluate the energy efficiency, and judge whether the energy efficiency is significantly improved.

Essiet et al. [92] proposed an improved GA to optimize energy utilization by balancing load scheduling and the contribution of renewable energy. Allahyarzadeh et al. [93] selected the most important parameters that affected the given output based on the GA to obtain fuel consumption and hydrocarbon liquid recovery, which greatly improved product recovery and energy efficiency. Wang et al. [94] constructed a dynamic constrained multi-objective optimization algorithm based on the PSO to solve the optimization problem of batch reactors in the chemical process. Geng et al. [95] proposed an adaptive multi-objective PSO algorithm based on dynamic analytic hierarchy process (AHP) integrating the fuzzy consistent matrix to select the global optimal solution and ensure correct direction of the evolution, which effectively solved the multi-target operation optimization problem and provided a viable solution for the ethylene cracking furnace.

At present, the data-driven artificial intelligence methods gradually replace the traditional statistical methods because of the fast response speed, no needs to consider the internal mechanism, and strong nonlinear approximation ability [96]. However, due to the cross-redundancy of the production data and the complicated production structure in the complex petrochemical industry, the generalization performance of simple artificial intelligence algorithms (such as neural networks) is still not satisfactory.

2.2.3. Data-driven hybrid methods

Data-driven hybrid methods are the most widely used and efficient methods for energy efficiency evaluation in the complex petrochemical industry [97,98]. The actual data of the complex petrochemical industry are processed by statistical methods, such as clustering, dimension reduction and feature extraction. Then the processed data is taken as the input of neural networks to establish the energy efficiency model. The flow chart of data-driven hybrid methods in the complex petrochemical industry is shown in Fig. 7. Due to the complexity of the petrochemical industry, there are usually a large number of independent variables. Every independent variable not only contains useful information for modeling, but also interference information. Moreover, these variables also interact with each other [99]. Therefore, taking all the variables directly as the input of the network not only increases the training time of the neural network, but also affects the training accuracy [100]. Combine the statistical methods with neural networks can not only reduce the dimension of input variables of neural networks, but also simplify the structure and improve the learning efficiency [101], as well as the accuracy of the energy efficiency evaluation. In recent years, there are many energy efficiency evaluation methods based on data-driven hybrid methods, such as neural network methods based on data attribute partitioning [102], based on principal component analysis [103], based on the clustering [104] and based on the traditional DEA [105].

Sanchez et al. [106] proposed an ELM based on the variable neighborhood search algorithm to obtain a set of features most relevant to the initial features. After the features were selected, the ELM was used to build an exponential prediction model. The results showed that this method had strong robustness. In view that the ELM cannot effectively solve the high-dimensional data modeling in petrochemical processes, Peng et al. [107] combined it with the self-associative neural network to filter the redundant information existed in the input data and extract the features. Then, the extracted features were taken as inputs of the ELM neural network. Finally, a self-associative ELM based on the feature extraction was proposed. Han et al. [108] proposed an ANN method based on the DEA. This method firstly divided all DMUs into valid DMUs and invalid DMUs by processing the data based on the DEA method, then optimized the invalid DMUs according to valid DMUs, and finally combined them with the ANN to establish a new energy efficiency evaluation model. Geng et al. [109] proposed an energy

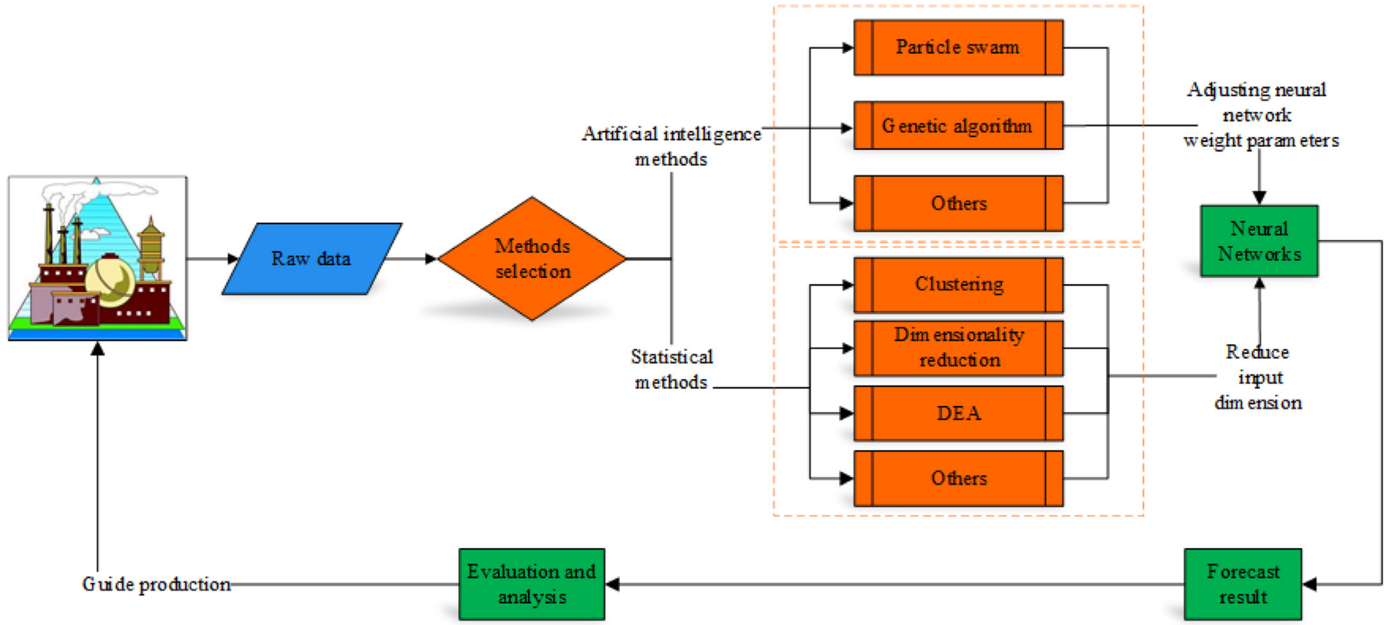

Fig. 7. The flow chart of data-driven hybrid methods.

efficiency evaluation method based on the RBF neural network integrating fuzzy C-means and the PCA algorithm. The PCA algorithm was used to eliminate the noise and reduce the data dimension and thus to reduce the training time. The fuzzy C-means was used to separate each fuzzy class in the input space and determine the number of neurons in the hidden layer of the RBF neural network, improving the accuracy of the energy efficiency evaluation in complex petrochemical industries. In addition, Geng et al. [110] proposed an ELM method based on fuzzy C-means and analytic hierarchy process. The fuzzy C-means algorithm was used to cluster the input attributes of the high-dimensional data. Through the analytic hierarchy process based on the entropy weight, the redundant information was filtered and the feature components were extracted. Finally, the fusion data was used as the input of the ELM to predict the complex petrochemical yield and improve the energy efficiency. Zhu et al. [111] proposed an energy efficiency evaluation method based on the ELM and exponential decomposition analysis. This method used exponential decomposition analysis to decompose the high-dimensional data into three energy performance indicators with activity effect, structure effect and intensity. Then, these indicators and complex petrochemical output were taken as inputs and outputs of the ELM, respectively. Finally, energy efficiency prediction and evaluation of ethylene production and the purified terephthalic acid (PTA) production in complex petrochemical process were obtained. Zhu et al. [112] proposed a robust extreme learning machine (RELM) based on principal component extraction (PCE). By extracting the principal component features of the hidden layer of the ELM, the linear correlation between variables was removed, which reduced the influence of the number of hidden layer nodes on the accuracy of the model and realized rapid selection of the number of hidden layer nodes, and improved the robustness of the ELM. Then, this method was applied in the PTA production process, and the results showed that the PCE-based RELM method was effective. Han et al. [113] proposed an improved evaluation method based on the ELM and the interpretative structural model (ISM), which used the ISM based on the partial correlation coefficient to analyze key parameters that affected the system energy and carbon emissions. The model was used to remove noise and reduce the dimensionality of the data, and the processed data were taken as inputs of the ELM to establish an energy efficiency evaluation model for the ethylene production in the complex petrochemical industry. Later, Han et al. [114] proposed a novel capacity analysis and energy efficiency prediction evaluation method. This method clustered the multi-dimensional data of the complex petrochemical industry with the affinity propagation (AP) to extract the main influencing factors. These factors were taken as inputs of the ELM to predict the energy output. This method was compared with the k-means clustering method. The results showed that this method greatly improved the prediction accuracy.

The above researches show that data-driven hybrid methods can tackle the shortcomings of single neural networks in energy efficiency evaluation, with high accuracy of modeling and strong stability. In order to further analyze data-driven hybrid evaluation methods, the following five typical methods including the ELM method based on the AP clustering (AP-ELM), the ELM method based on k-means clustering (K-means-ELM), the single ELM method, the single BP method, and the single RBF method are compared based on the real ethylene data. The experimental data are from 2009 to 2013 of ethylene production industries [114]. The ratio of raw materials and ingredients, which include crude oil, fuel, steam, water and electricity. The main materials affected the ethylene production is crude oil, which include Naphtha (NAP), carbon3, carbon4, carbon5 (C345), Residual fluid (REF), Hydrogenated tail oil (HDL), Light diesel oil (LDO), Light hydrogenation tail oil (LHY), and other materials (OTH). The output is the ethylene yield. The results in Fig. 8 and Table 3 showed that for prediction of ethylene yields under the same training set and test set, the prediction accuracy of the ELM is higher than that of the BP neural network and the RBF neural network, and the prediction accuracy of the neural network combined with the statistical method is higher than that of the single neural network.

However, the most data-driven methods are based on the neural network with the simple structure, such as the BP, which are limited to extracting the more effective features of complex petrochemical data. Therefore, deep learning methods began to be

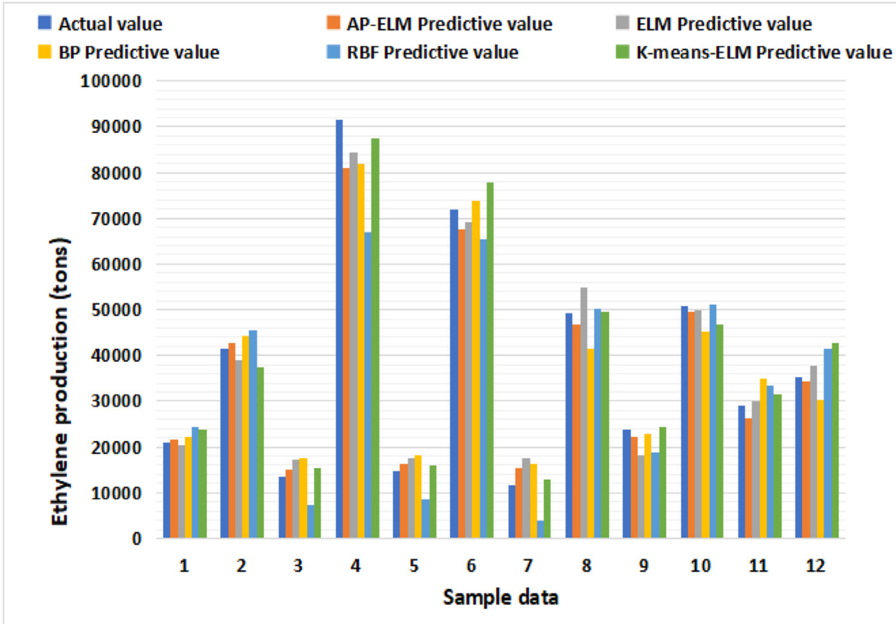

Fig. 8. The ethylene yield prediction based on five typical methods.

Table 3
The error comparison of five typical methods for ethylene yield prediction.

	BP	RBF	ELM	k-means-ELM	AP-ELM
Number of training sets	199	199	199	199	199
Number of test sets	12	12	12	12	12
Average error	0.073	0.110	0.071	0.068	0.060

used for energy efficiency evaluation. For the purpose of improving the energy efficiency in the complex petrochemical industry, an energy optimization and prediction model based on the improved CNN integrating the cross-feature (CF) is proposed by Geng et al. [115]. The CF can combine the correlation between features to obtain the input of the CNN, which can avoid the over-fitting problem caused by fewer features. Then the CNN is designed as a three-layer structure and the Rectified Linear Unit (RLU) is introduced to achieve better generalization capability and stability. Han et al. [116] proposes a production capacity analysis and energy saving model using long short-term memory (LSTM) based on attention mechanism (AM) to directly control raw materials consumption and effectively measure the product quality. Therefore, the deep learning methods are suitable for large-scale data and can extract the better features of these data.

Among data-driven hybrid methods, in addition to the combination of statistical methods and neural networks, there is also a kind of hybrid method that combine neural networks with other artificial intelligence algorithms, such as neural networks with the PSO [117,118], and neural networks with the GA [119] or ant colony algorithm [120]. They are designed to select the most optimal influencing factors, rather than simple feature extraction. Therefore, this kind of method is also one of the most popular energy efficiency evaluation method at present [121]. Kong et al. [122] proposed an evaluation method combining the GA with the BP neural network, to get the raw material ratio under the condition of given quality index and optimal product performance by establishing the relationship between the decision variable and the objective function, and selecting the best decision variable based on the GA. Data-driven hybrid methods can solve the problem of data cross-redundancy well in petrochemical industries and optimize the generalization performance of artificial intelligence methods, which perform well under experimental conditions.

2.3. The methods combining mechanism method and data-driven method

For the complex petrochemical industry, the mechanism method based on the traditional TT-RG is easy to be limited by practical conditions, and the application scope is limited. Moreover, the data-driven method is only effective in experiments and not suitable to be widely promoted. Therefore, the methods combining mechanism method and data-driven method have been proposed. The methods not only inherit advantages of mechanism methods, but also simplify the energy efficiency evaluation method by appropriately ignoring some parameters which are difficult to determine. The flow chart of the mixed method is shown in Fig. 9. This kind of energy efficiency evaluation method includes mechanism methods integrating ANNs or mechanism methods integrating other artificial intelligence methods.

Recently, Palagi et al. proposed an organic ranking cycle multi-objective optimization algorithm based on the ANN, which can select the best cyclic thermodynamic parameters, and has been proved very suitable for solving typical high-nonlinear problems in energy systems [123]. With the continuous growth of the energy demand, in order to improve the energy efficiency, Arshad et al. [124] proposed a thermodynamic analysis and optimization method for double effect absorption refrigeration system using the GA. Due to the low accuracy of the distillation column model, He et al. [125] proposed a measurement method of the alcohol distillation column based on the mechanism method and neural network compensation, not only improved the distillation accuracy and efficiency, but also the extrapolation ability of the neural network. In order to improve the estimation accuracy of model parameters, Yu et al. [126] took a methanol synthesis tower as the research object and proposed an improved wolf group algorithm through in-depth analysis of the reaction mechanism of the methanol synthesis process. The modified wolf pack algorithm was

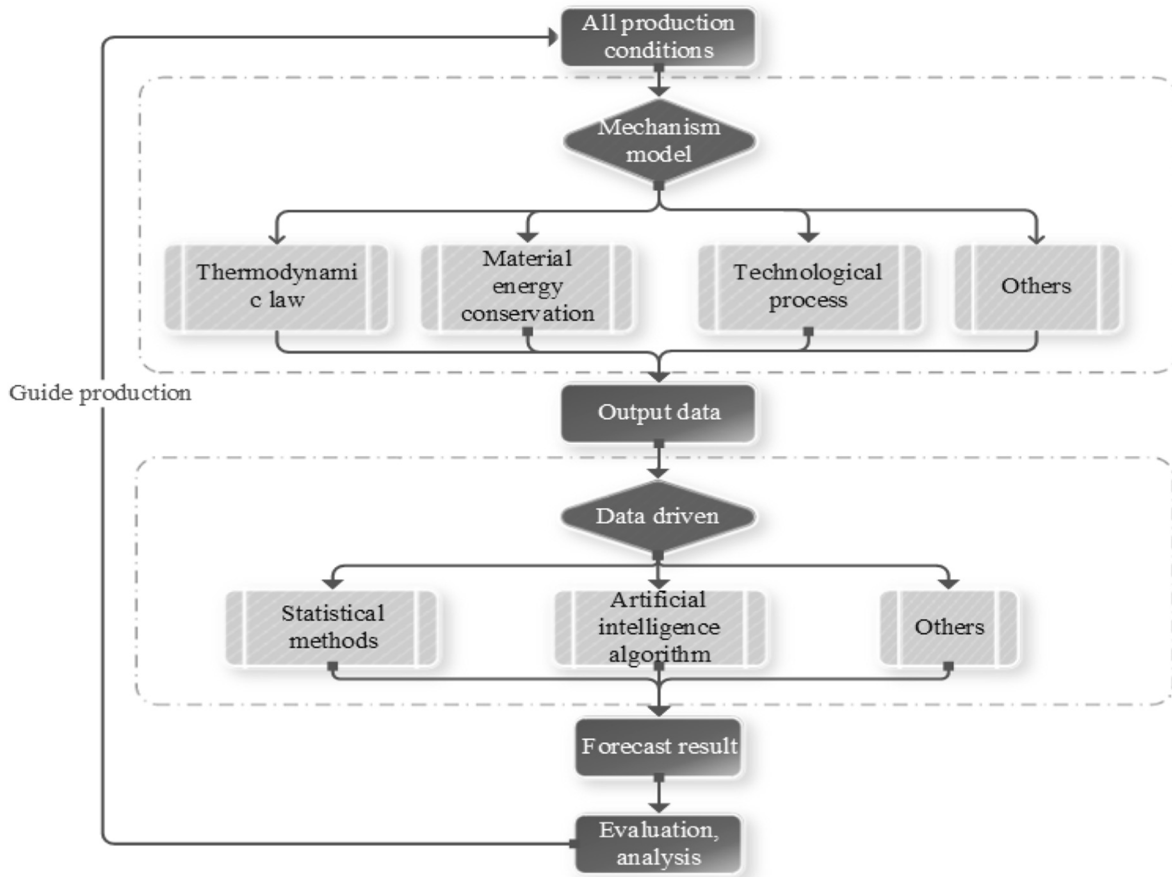

Fig. 9. The flow chart of the combined method.

used to estimate parameters of the methanol conversion mechanism method, which could more accurately predict the methanol conversion at the exit of the synthesis tower and improved the energy efficiency. In order to realize the refined control of the energy flow network, Liu et al. [127] proposed an input-output method of the converter based on the mechanism and data-driven method, to analyze the input and output in the converter process, and obtain relevant parameters of the converter refining by using mathematical statistics and regression methods. Then, they calculated the output parameters according to the mechanism method, and used the neural network to predict the temperature of the molten steel end point for improving the accuracy of the model. In addition, based on the actual production data, Xu et al. [128] proposed analysis methods of oxygen consumption and gas recovery, based on the mechanism method combining with the data-driven method, for the steelmaking-continuous casting process. The method reduced the process consumption and realized secondary energy recycling. The blast furnace smelting process is characterized by strong nonlinearity, time delay and under regulation. Most of its internal parts are coupled complex dynamic systems, so it is difficult to achieve a good effect by constructing its model by the single mechanism. Combined with the blast furnace smelting mechanism, operational data, and expert experience, Li et al. [129] used the measured data of the sensor to construct a method of the blast furnace in the complex industry.

At present, the combined evaluation method is not widely used in the production system of petrochemical industries, but its effect is improved obviously. Although the combined method has achieved good results in the energy efficiency evaluation of the petrochemical industry, and is easy to be promoted. However, because of the high cost and complexity, it is not easy to establish appropriate production and energy efficiency evaluation model. Therefore, this combined methods are suitable for the petrochemical industry with the urgent needing to extend but not getting all meanings of parameters.

3. Discussion

In this paper, we have introduced many energy efficiency evaluation methods in complex petrochemical industries. These methods are mainly divided into three parts with the mechanism methods based on TT-RG, the data-driven artificial intelligence methods, and the hybrid methods combining the mechanism and the data-driven. The mechanism methods based on TT-RG, which is established by describing the mechanism process, is a direct mathematical description and the essence reflection of the actual production process. Therefore, it is not only easy to extend but also has the advantages of clear physical meaning of parameters, strong adaptability. However, these mechanism methods also have visible disadvantages for difficulty to determine the mathematical expression of some complex objects with a large number of parameters. What's more, if the parameters cannot be well obtained, the modeling result will be inaccurate.

With the development of artificial intelligence and big data, data-driven methods have gradually become a very common evaluation modeling tool based on the actual data. The data-driven method makes up the shortcomings of mechanism methods based on TT-RG. Through addressing the high dimensionality and nonlinearity in petrochemical processes effectively, the modeling process is become more simpler [70]. Besides, the data-driven methods based on the ANN have the advantages of its fast response speed, without considering internal mechanism, and strong nonlinear approximation ability. However, it is difficult to build the accurate evaluation model of the complex petrochemical industry with multi-dimensional and multi-level characteristics due to the limitation of the ANN itself. Therefore, some scholars began to used data-driven hybrid methods that combine statistical analysis with the ANN to build energy efficiency evaluation model, which improved the prediction accuracy. However, many statistical methods have certain requirements on the data, such as requiring the data to obey normal distribution. Regression analysis is mainly based on the dependency relationship and the internal laws between data. If the correlation of the data is not large, this method loses its value. The PCA can extract the main impact factors and ignore the secondary factors, which may loss some useful information after data processing. The cluster analysis mainly uses different rules to make similar data into clusters, and then selects the cluster center from different clusters as the representative of each category, which finally achieves effective data fusion.

And then, these five typical methods including the AP-ELM method, the k-means-ELM method, the single ELM method, the single BP method, and the single RBF method are compared based on the real ethylene data. Compared with the K-means-ELM, the ELM, the BP, and the RBF under the same condition, the average error of the AP-ELM decreased 0.8%, 1.1%, 1.3% and 5%, respectively. Furthermore, because the ANN with the simple structure cannot extract the more effective features of complex petrochemical data, the deep learning method with the CNN and the recurrent neural network (RNN) is becoming more and more popular for evaluating the energy efficiency in the complex petrochemical industry.

For the complex petrochemical industry, the application scope of the mechanism method based on the traditional TT-RG is limited by practical conditions. And the data-driven method is more effective in experiments and not suitable to be widely promoted. Therefore, the hybrid methods combining the mechanism method and the data-driven method have been proposed and become another trend for energy efficiency evaluation. The hybrid methods not only inherit advantages of mechanism methods, but also simplify the energy efficiency evaluation process by appropriately ignoring some parameters which are difficult to determine. However, this method still has some limitations for the difficult in modeling and some parameters are difficult to determine of the energy efficiency evaluation. Despite the above defects, the mechanism methods based on the TT-RG and data-driven methods still should be studied and focused as the support and foundation for the hybrid method.

4. Development in the future

Through comprehensive introduction, comparison and summary for existing energy efficiency evaluation methods in the complex petrochemical industry, we find that the better and effective energy efficiency evaluation methods constructed in the petrochemical production system are the following two methods. The data-driven hybrid method which combining statistical methods and deep learning methods such as the CNN can abstract high-level features with having better effect in evaluate energy efficiency modeling. And the hybrid methods combining mechanism methods and data-driven methods can simplify the energy efficiency evaluation process and easy to widely promoted. However, if these evaluation methods are analyzed and used well, there are mainly focused on three aspects. First, how to extract and fuse main factors that affect the energy efficiency from the raw data with high coupling and complexity by clustering and reducing the dimension. Second, how to use less training data and less numbers of iteration, so as to obtain better generalization ability. Finally, how to meet the actual production process as much as possible, and make these methods extrapolated substantially. Therefore, the future research is shown as follows.

4.1. More intelligence data analysis method

In this paper, we have introduced many statistical and artificial intelligence methods. Generally, the statistical method means that the raw data is simplified by classifying and analyzing the high-dimensional data of the complex petrochemical industry. Traditional data analysis methods include clustering algorithms, dimension reduction algorithms, feature extraction methods and the DEA. However, with application of the computer technology, the intelligent manufacturing and the cloud platform, a large number of multi-dimensional, multi-level and high-coupled data are produced in complex petrochemical production systems. Therefore, efficient data-processed methods can help to deal with these intricate data well [130]. Currently, the more efficient methods are optimization algorithms including traditional optimization algorithms and improved optimization algorithms. The traditional optimization algorithm is based on the single solution, such as the hill climbing algorithm, the tabu search (TS) algorithm, the greedy algorithm and the simulated annealing (SA) algorithm. However, the outputs of traditional optimization algorithms are the same and start from a solution to look for the best, which easy to fall into the local best. The improved optimization algorithms based on the population can process multiple individuals in the group simultaneously and evaluate multiple solutions in the search space, which reduce the risk of falling into the local solution, and the algorithm itself is easy to parallelize. Generally, the improved optimization algorithms including the GA and the swarm intelligence optimization algorithm. The swarm intelligence optimization algorithm including the PSO, the Ant Colony Algorithm (ACO), the Artificial Bee Colony Algorithm (ABC), the Artificial Fish School Algorithm (AFSA), the Shuffled Frog Leaping Algorithm (SFLA) and the Bacterial foraging optimization (BFO). These improved optimization algorithms have grate advantage in dealing with redundant and complex data in complex petrochemical industries.

Moreover, the effective energy efficiency evaluation knowledge can be obtained through novel optimization algorithm and knowledge discovery, which can build the energy efficiency model more easily. Above all, the advanced data processing method is one of necessary means for future development of energy efficiency evaluation in complex petrochemical industries.

4.2. Deep learning method

Deep learning is the gradual development of the ANN, which contains multiple hidden layers. The deep learning can combine low-level features to form more abstract high-level features that represent attribute categories, and then to discover distributed features of the data. Therefore, the deep learning which contains multi-layer perceptions, has great advantages for solving complex data problems and good generalization ability. The CNN is one typical deep learning method with multi-layer structures for the feature extraction. And the CNN uses spatial relative relationship to reduce the number of parameters, which not only improves the training performance, but also improves the precision of multi-dimensional and multi-level energy efficiency evaluation model in complex petrochemical processes. And the RNN can take the time sequence data as the input to learn the nonlinear features, and has excellent characteristics of the memory and parameter sharing. Because the LSTM introduces a memory unit to replace the hidden layer node in the traditional RNN, the LSTM can avoid the problem of gradient explosion or gradient disappearance that may occur when the RNN returns the gradient. And the bidirectional recurrent neural network (BIRNN) considers the information of the next time step and the previous time step simultaneously to make the decision of the current time step, so the BIRNN has been applied in the sequence labeling task of Natural language processing (NLP) successfully. Therefore, the deep learning will be used to extract more features of energy efficiency data in complex petrochemical industries.

4.3. Method model extension

The extension of the energy efficiency evaluation method in complex petrochemical industries mainly includes two aspects. On one hand, the energy efficiency evaluation method can be extended from experiments to the real production processes. The predictive performance learned from one task can be migrated to another task. On the other hand, the method combining the data-driven statistical method with the deep learning has low cost and can avoid many inconveniences in the actual production process. However, the simulation environment is still different from the actual production, and the effect of the method is not very good in the real production. Therefore, how to apply the method with the good performance from experiments to the actual petrochemical industry is one of the key issues in energy efficiency analysis and optimization.

5. Conclusion

Compared with traditional process methods in complex petrochemical industries, the methods combining data-driven with the artificial intelligence have advantages of data fusion and energy optimization. It can dynamically adapt to the unstructured production environment of various energy sources in complex petrochemical industries. Furthermore, it overcomes defects of lacking index analysis of the intermediate process and simple calculation of energy consumption indexes in previous energy efficiency analysis. It can also analyze and evaluate the specific device of different technologies and sizes from multi-dimensional and multi-objective perspective. Moreover, it provides technical reference and industry benchmark for decision makers, which can find the main factors and better direction of improving the energy efficiency to achieve good results and economic benefits.

The development of artificial intelligence technology has provided some new technologies and opened up some new idea for energy efficiency evaluation for the complex petrochemical industry. Compared with the light industry and the infant industry, the complex petrochemical industry requires higher price and costs. Currently, how to combine the new data-driven method with the mechanism method and the deep learning method to improve the energy efficiency and reduce pollutant emissions of the complex petrochemical industry is a new direction of the research. Therefore, by reviewing these energy efficiency evaluation methods, not only the most suitable evaluation methods can be found to evaluate the energy efficiency of complex petrochemical industry, but also can provide theoretical guidance for energy conservation and emission reduction further. Besides, there have been a certain understanding of energy efficiency evaluation methods in the petrochemical industry. Such as, which data fusion method are appropriate to deal with the petrochemical data and what is the easiest evaluation methods for analyzing the energy efficiency.

In addition, for the current research situation and characteristics of big data in the petrochemical process, the density of data collection should be increased to achieve real-time analysis for petrochemical production processes.

Declaration of interest statement

The authors declared that they have no conflicts of interest to this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted. Your attention and consideration of our manuscript is honestly appreciated, and hope that the revision will meet with requirements.

Acknowledgement

This work is supported by National Natural Science Foundation of China (21978013 and 61533003), the Fundamental Research Funds for the Central Universities, China (XK1802-4), Science and Technology Major Project of Guizhou Province (Guizhou Branch [2018]3002) and Industrial Internet Innovation and Development Project (TC19084DY).

References

[1] Wu Q, Wu CY. Research on evaluation model of energy efficiency based on DEA. Journal of Management Science 2009;22:103–12.
[2] Liao H, Du YF, Huang ZM, Wei YM. Measuring energy economic efficiency: a mathematical programming approach. Appl Energy 2016;179:479–87.
[3] Zafer U, Arif H. A review and assessment of the energy utilization efficiency in the Turkish industrial sector using energy and exergy analysis method. Renew Sustain Energy Rev 2007;11(7):1438–59.
[4] Baileya JA, Gordona R, Burtonb D, Yiridoe EK. Energy conservation on nova scotia farms: baseline energy data. Energy 2008;33:1144–54.
[5] Jiang C, Zhong WM, Li Z, Peng X, Yang ML. Real-time semi-supervised predictive modeling strategy for industrial continuous catalytic reforming process with incomplete data using slow feature analysis. Ind Eng Chem Res 2019;58(37):17406–23.
[6] Wang SH, He XW. Ethylene process and technology [M]. Beijing: China Petrochemical Press; 2000.
[7] Chang LZ, Li ZB. Static model of converter based on BP neural network. Steelmaking 2006;22:41–4.
[8] Han YM, Geng ZQ, Zhu QX, Wang Z, Cui YF. Energy consumption analysis and evaluation of petrochemical industries using an improved fuzzy analytic hierarchy process approach. J Intell Fuzzy Syst 2017;32(6):4183–95.
[9] Geng ZQ, Wang ZK, Zhu QX, Han YM. Energy efficiency evaluation method based on IDA-DEA and its applications in ethylene industries. CIE J 2017;68(3):910–5.
[10] Gong S, Shao C, Zhu L. Energy efficiency evaluation in ethylene production process with respect to operation classification (Article). Energy 2017;118: 1370–9.
[11] Geng ZQ, Li HD, Zhu QX. Production prediction and energy-saving model based on Extreme Learning Machine integrated ISM-AHP: application in complex chemical processes. Energy 2018;106:898–909.
[12] Nan Y, Di YE, Jie L, Huang Yu, Dong BT, Hu WB, Liu SK. Research on data-driven intelligent security-constrained unit commitment dispatching method with self-learning ability. Proceedings of the CSEE 2019;(10): 2934–46.
[13] Yu P, Gu XB, Qi ZQ. EPI-based ethylene industry efficiency analysis. J Chem Ind Eng 2012;63:2931–5.
[14] Peter VC, Edward RD, Roger RH. Big data need big theory too. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences; 2016.
[15] Yan LT, Li M, Yang DY. Prediction of carbon content in converter end point based on GA-KPLSR. Contr Eng 2017;24:923–6.
[16] Hu Y. Research on knowledge discovery and endpoint control model of converter steelmaking based on rough set. Chongqing: Chongqing University; 2013.
[17] Zhang L, Pan YF, Yuan ZF. Research on control technology of converter blowing and limestone residue behavior. Nonferrous Metals Science and Engineering 2015;6:16–21.
[18] Sun YG, Liang QY, Li W-B, Jia TY. Steel industry multi-type energy optimized scheduling with energy flow network simulation. Acta utomatica Sinica 2017;43:1065–79.
[19] Peter VC, Edward RD, Roger RH. Big data need big theory too. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and

Engineering Sciences; 2016. p. 374.
[20] Kumar BD, Monjur M. Forecasting methods in energy planning models. Renew Sustain Energy Rev 2018;88:297–325.
[21] Zhao ZT, Chong KT, Jiang JY, Wilson K, Zhang XC, Wang F. Low-carbon roadmap of chemical production: a case study of ethylene in China. Renew Sustain Energy Rev 2018;97:580–91.
[22] Zhang B, Yang WM, Qian F, Wu ZY, He WJ, Zhang WD. First principle modeling and optimization for dehydrogenation of ethylbenzene to styrene. Control Theory & Appl 2010;(7):903–8.
[23] Wang W, Zheng YL. A new method of system identification using neural network. Autom Instrumentacion 2002;(6):18–21.
[24] Gao QJ. Mechanism modeling and analysis of coaxial catalytic cracking reaction system. Liaoning Chemical Industry 2018;47:575–7.
[25] Elisa G, Adriano S, Vittorio V. Thermo-fluid dynamic model of large district heating networks for the analysis of primary energy savings. Energy 2019;184:34–44.
[26] Cheng ZZ, Gong K, Wang Y, Wang GH, Li WB, Mao L. Analysis of coke oven heat balance calculation and energy saving and consumption reduction. Ind Furn 2013;35:27–31.
[27] Han QK, Fan QM, Shen HP. Research progress on the thermal conversion mechanism of aromatic hydrocarbon model compounds. Prog Chem Eng 2017;36:133–41.
[28] Wang Q, Xu XY. The influence of ethane in the United States on China's modern coal chemical industry. Journal of China University of Petroleum (Edition of Social Sciences), Vol.34(4): 7-12.
[29] Hua W. Process energy analysis and synthesis [M]. Beijing: Hydrocarbon Processing Press; 1989.
[30] Yan MY. Optimization of ethylene plant process flow to reduce energy consumption. Chemical Industry Management 2017;28:126.
[31] Zhang LH, Zhao L, Dong H. Construction and energy efficiency analysis of a CO2 capture system. Metallurgical Energy 2017;36:17–9.
[32] Wang X, Cui JJ, Zhao L, Zhang W, Dong H. Energy efficiency analysis of CO_2 hydrogenation to methanol system in magnesite smelting flue gas. Metallurgical Energy 2018;37:6–10.
[33] Jiao K, Yuan ZF, Zhang LN, Xie SS, Li LS, Sui DP. Characteristic and model of thermal decomposition of magnesite. Nonferrous Met 2018;(1):64–8.
[34] Zhang SY. Application of catalytic cracking process in petroleum refining process. China Petroleum and Chemical Standards and Quality 2019;39: 218–9.
[35] Zhong W, Wang SQ. Online soft measurement of crude gasoline dry point. J Chem Ind Eng 1998;49(2):251–5.
[36] Zhao Q, Cao JL, Hu YL. Salient region detection based on Gaussian multi-scale transform and color complexity measure. Chin J Sci Instrum 2012;(2): 405–12.
[37] Rabaza O, Gomez LD, Perez OF, Pena G. A simple and accurate model for the design of public lighting with energy efficiency functions based on regression analysis. Energy 2016;107:831–42.
[38] Ivan S, Dragoljub M, Svetlana S. Possibilities for wider investment in solar energy implementation. Energy; 2019.
[39] Zhao Y, Wang PH, Su ZG, Li YG, Zhu XJ. TS modeling based on robust fuzzy C-regressions and its application for thermal process. Proceedings of the CSEE 2018;38(7):2063–9.
[40] Zhang LJ, Wang GQ. Modeling of steam cracking of naphtha by SVM. Petrochem Technol 2017;46(8):1022–7.
[41] Ahmad T, Chen HX, Shair J. Water source heat pump energy demand prognosticate using disparate data-mining based approaches. Energy 2018;152: 788–803.
[42] Protić M, Shamshirband S, Petković D, Abbasi A, Mat Kiah ML, Unar JA, Živković L, Raos. Forecasting of consumers heat load in district heating systems using the support vector machine with a discrete wavelet transform algorithm. Energy 2015;87:343–51.
[43] Dai XH, Wang YC, Zhu Y, Dai ZH, Song YL. Natural gas pipeline network load forecast by SVM-based method. Natural Gas and Oil 2009;(2):13–5.
[44] Yao KF, Lu WK, Ding WL, Zhang SW, Xiao HQ, Li YD. An oil and gas prediction method based on SVM feature selection. Nat Gas Ind 2004;7:36–8.
[45] Cui QA, He W, Cui FX. Coke quality prediction model based on support vector machine. Chemical Automation & Instrument 2006:28–31. 01.
[46] Cui PC, Wang XL, Ma ZL. Ethylene energy efficiency analysis system based on data envelopment analysis. Computers and Applied Chemistry 2010;27: 1182–6.
[47] Lin XY, Zhu XP, Han YM, Geng ZQ, Liu L. Economy and carbon dioxide emissions effects of energy structures in the world: evidence based on SBM-DEA model. Science of The Total Environment; 2020. p. 138947.
[48] Cook Wade D, Harrison JL, Raha I, Paul R, Joe Z. Data envelopment analysis with nonhomogeneous DMUs. Oper Res 2013;61(3):666–76.
[49] Raha I, Wade D, Cook A, Sonia VB, Joe c. Partial input to output impacts in DEA: the case of DMU-specific impacts. Eur J Oper Res 2015;244(3):837–44.
[50] Zhou P, Ang BW, Poh KW. A survey of data envelopment analysis in energy and environmental studies. Eur J Oper Res 2008;189:1–18.
[51] Azadeh A, Amalnick MS, Ghaderia SF, Asadzadeha SM. An integrated DEA PCA numerical taxonomy approach for energy efficiency assessment and consumption optimization in energy intensive manufacturing sectors. Energy Pol 2007;35:3792–806.
[52] Zha Y, Zhao LL. Bian WY, measuring regional efficiency of energy and earbon dioxide emissions in China: a chance constrained DEA approach, computers. Oper Res 2016;66:351–61.
[53] Fan TJ, Luo RL, Xia HY, Li XP. Using LMDI method to analyze the influencing factors of carbon emissions in China's petrochemical industries. Nat Hazards 2015;75:S319–32.
[54] Chen YX, Han YM, Zhu QX. Energy and environmental efficiency evaluation based on a novel data envelopment analysis: an application in petrochemical industries. Appl Therm Eng 2017;119:156–64.
[55] Han YM, Geng ZQ, Wang Z, Mu P. Performance analysis and optimal temperature selection of ethylene cracking furnaces: a data envelopment analysis cross-model integrated analytic hierarchy process. J Anal Appl Pyrol 2016;122:35–44.
[56] Han YM, Geng ZQ, Liu QY. Energy efficiency evaluation based on data envelopment analysis integrated analytic hierarchy process in ethylene production. Chin J Chem Eng 2014;22:1279–84.
[57] Wang HZ, Yu JS. Software measurement modeling and application based on PCA-SVM. Autom Instrumentacion 2004;25:16–9.
[58] Feng DC, Lu H. Application of data-based techniques for petrochemical industry process. Automation in Petro-Chemical Industry 2010;(6):28–35.
[59] Liu Y. The review of intelligent mechanical engineering based on artificial neural network. Proceedings of the 2015 international conference on intelligent systems research and mechatronics engineering 2015;121:1969–73.
[60] Dong W, Zhou J. Application of predictive model based on optimized BP neural network in energy management of spinning enterprises. Industrial Control Computer 2016;29:150–2.
[61] Yang L. Several problems of artificial neural network for nonlinear system modeling. Gansu Science and Technology 2002;18(7). 126-113.
[62] Sliskovic D, Grbic R, Hocenski Z. Methods for plant databased process modeling in soft-sensor development. Automatica 2012;52:306–18.
[63] Ahmad Z, Zhang J. Selective combination of multiple neural networks for improving model prediction in nonlinear systems modeling through forward selection and backward elimination. Neurocomputing 2009;72:1198–204.
[64] Geng ZQ, Li YN, Han YM, Zhu QX. A novel self-organizing cosine similarity learning network: an application to production prediction of petrochemical systems. Energy 2018;142:400–10.
[65] Cui YF, Han YM, Geng ZQ, Zhu QX, Fan JZ. Production optimization and energy saving of complex chemical processes using novel competing evolutionary membrane algorithm: emphasis on ethylene cracking. Energy Convers Manag 2019;196:311–9.
[66] Lu XD, Zhao L. Cracking severity real time optimization of ZRCC ethylene plant. Green Petroleum & Petrochemicals 2016;1(2):24–9.
[67] Suo HB, Jiang X, Hu Y, Su GD. Optimization of the reaction dehydration to ethylene based on RBF neural network simulation. Pet Process Petrochem 2010;(3):69–73.
[68] Wang WX, Pan LD, Li R, Xu YX, Wen GH. Modeling and application of double model structure RBF neural network in atmospheric and vacuum distillation unit. Journal of Beijing University of Chemical Technology. Natural Science Edition 2004;(4):91–4.
[69] Li Y, Huang DP, Li GY, Jiang RB, Zhong ZP. Prediction of 350 °C content in atmospheric distillation of refinery based on neural network. Computer Engineering and Applications 2006;(30):224.
[70] Liu L, Shi XH. Multi-objective optimization of RBF proxy model rectification process based on adaptive sampling strategy. Computer and Applied Chemistry 2018;35(6):469–79.
[71] Geng ZQ, Li HD, Zhu QX, Han YM. Production prediction and energy-saving model based on extreme learning machine integrated ISM-AHP: application in complex chemical processes. Energy 2018;160:898–909.
[72] Yang YL. Application of BP neural network in coordinate transformation. Construction Engineering Technology and Design 2015;(15).
[73] Hou Y, Zhao L, Lu HW. Fuzzy neural network optimization and network traffic forecasting based on improved differential evolution. Future Generat Comput Syst 2018;81:425–32.
[74] Qian GW, Zhang L. A simple feedforward convolutional conceptor neural network for classification. Applied Soft Computing Journal 2018;70: 1034–41.
[75] Taki M, Rohani A, Soheili-Fard F, Abdeshahi A. Assessment of energy consumption and modeling of output energy for wheat production by neural network (MLP and RBF) and Gaussian process regression (GPR) models. Journal of Cleaner Process 2018;172:3028–41.
[76] Gu HC, Yan P, Li JW. Modeling of ethylene cracking furnace based on cross-iterative BLSTM network. J Chem Ind Eng 2019;70(2):548–55.
[77] Huang GB, Zhu QY, Siew CK. Extreme learning machine: a new learning scheme of feedforward neural networks. IEEE international joint conference on neural networks. Proc IEEE 2004;2:985–90. 2005.
[78] Zou WD, Xia YQ. Energy consumption prediction of incremental ELM virtual machine based on compressed momentum term. Acta Autom Sin 2019;45(7):1290–7.
[79] Yao W. A novel regular algorithm based on genetic algorithm. Computer Knowledge and Technology 2018;14(22):179–81.
[80] Chen WB, Song MJ. A data visualization method based on extreme learning machine. Comput Eng Sci 2017:912–8. 05.
[81] Huang GB, Zhu QY. Extreme leaning machine: a new leaning scheme of feedforward neural networks. Proceeding of International joint conference neural networks 2004:985–90.
[82] Siew CK, Huang GB, Zhu QY. Extreme leaning machine: theory and applications. Neurocomputing 2006;70:489–501.

[83] Zhang YT, Ma C, Li ZN, Fan HB. On-line modeling of nuclear extreme learning machine based on fast leave-one cross validation. J Shanghai Jiaot Univ 2014;(5):641–6.
[84] Zhou LC, Yan X, Liu Y, Gao ZL, Jin FJ. Application of instantaneous local modeling in prediction of flood gas velocity in packed tower liquid. J Chem Ind Eng 2016;67:1070–5. 03.
[85] Chen Sb, Hu Z. Determination of diesel coagulation point by ELM correction model of near infrared spectroscopy. Contemp Chem Ind 2017;46:1010–3. 05.
[86] Zhanf HG, Lu JG. Improved ELM algorithm for quantitative analysis of near infrared spectroscopy. Spectrosc Spectr Anal 2016;36:2784–8. 09.
[87] Li LB, Lin RB, Wang XQ, Zhou W, Jia LT, Li JP, Chen BL. Kinetic separation of propylene over propane in a microporous metal-organic framework. Chem Eng J 2018;354:977–82.
[88] Tao JZ, Yin GF, Wang FG. Signal recognition technology based on genetic algorithm. Vibration. Testing and Diagnosis 2004;(3):6–8.
[89] Chand S, Dutta A. Reliable shape optimization of structures subjected to transient dynamic loading using genetic algorithms. Shock Vib 2005;12(6): 407–24.
[90] Gao H, Xu WB, Sun J. A quantum-particle swarm optimization algorithm for optimizing high-dimensional functions. J Comput Appl 2007;(12):2885–7.
[91] Vanessa LG, Manuel S, Rolf WK. Towards a better understanding of the local attractor in particle swarm optimization: speed and solution quality. Adptive and intelligent systems, ICAIS 2014;8779:90–9.
[92] Essiet I, Sun YX, Wang ZH. Optimized energy consumption model for smart home using improved differential evolution algorithm. Energy 2019;172: 354–65.
[93] Allahyarzadeh BA, Dezan DJ, Salviano LO, Oliveira JS, Yanagihara JI. FPSO fuel consumption and hydrocarbon liquids recovery optimization over the lifetime of a deep-water oil field. Energy 2019;181:927–42.
[94] Wang SS, Du WL, Chen X, Xu B, Qian F. Dynamic multi-objective optimization of chemical process based on constrained backbone particle swarm optimization algorithm. J East China Univ Sci Technol 2014;40:449–57. 04.
[95] Geng ZQ, Wang Z, Zhu QX, Han YM. Multi-objective operation optimization of ethylene cracking furnace based on AMOPSO algorithm. Chem Eng Sci 2016;153:21–33.
[96] Jiang DF. Research on term structure of interest rate based on neural network. Qinghai Finance 2007;(5):21–3.
[97] Priyanka S, Pragya D. A novel hybrid model based on neural network and multi-objective optimization for effective load forecast. Energy 2019;182: 606–22.
[98] Oussama L, Mohamed TK, Lyudmila M. Toward efficient energy systems based on natural gas consumption prediction with LSTM Recurrent Neural Networks. Energy 2019;177:530–42.
[99] Yuan ZQ, Yin BT, Shang HY, Zhu QX. Soft measurement model of PTA solvent dehydration tower based on PCA-BP algorithm. Computers and Applied Chemistry 2006;23(7):623–6.
[100] Li K, Jin Y, Akram MW, Han R, Chen J. Facial expression recognition with convolutional neural networks via a new face cropping and rotation strategy. Vis Comput 2019;36(2):391–404.
[101] Zhang GY, Ge BH, Fang BS. Neural network based on principal component analysis and its application in modeling optimal pH of xylanase. Computers and Applied Chemistry 2005;22:749–52.
[102] Gao HH, He YL, Peng Y, Zhu QX. Hierarchical ELM research and chemical application based on data attribute partitioning. J Chem Ind Eng 2013;64: 4348–53.
[103] He YL, Wang X, Zhu QX. Soft measurement of refined terephthalic acid content based on improved ultimate learning machine method based on principal component analysis. Control Theory & Appl 2015;32:80–5.
[104] Han YM, Long C, Geng ZQ, Zhu QX, Zhong YH. A novel DEACM integrating affinity propagate on for performance evaluation and energy optimization modeling: application to complex petrochemical industries. Energy Convers Manag 2019;183:349–59.
[105] Geng ZQ, Zeng RF, Han YM, Zhong YH, Fu H. Energy efficiency evaluation and energy saving based on DEA integrated affinity propagation clustering: case study of complex petrochemical industries. Energy 2019;179:863–75.
[106] Sanchez OJ, Duarte A, Salcedo SS. Robust total energy demand estimation with a hybrid variable neighborhood search extreme learning machine algorithm. Energy Convers Manag 2016;123:445–52.
[107] Peng Y, He YL, Xu Y, Zhu QX. AANN-ELM research and chemical application based on data feature extraction. CIE J 2012;63:2920–5.
[108] Han YM, Geng ZQ, Zhu QX. Energy optimization and prediction of complex petrochemical industries using an improved artificial neural network approach integrating data envelopment analysis. Energy Convers Manag 2016;124:73–83.
[109] Geng ZQ, Chen J, Han YM. Energy efficiency prediction based on PCA-FRBF model: a case study of ethylene industries. IEEE transactions on systems 2017;47:1763–73.
[110] Geng ZQ, Qin L, Han YM, Zhu QX. Energy saving and prediction modeling of petrochemical industries: a novel ELM based on FAHP. Energy 2017;122: 350–62.
[111] Geng ZQ, Yang X, Han YM, Zhu QX. Energy optimization and analysis modeling based on extreme learning machine integrated index decomposition analysis: application to complex chemical processes. Energy 2017;120:67–78.
[112] Zhang XT, Wang PJ, Gu XB, Xu Y, He YL, Zhu QX. Research on robust extreme learning machine based on principal component extraction and its application in chemical engineering modeling. J Chem Ind Eng 2019;70(2): 475–80.
[113] Han YM, Zhu QX, Geng ZQ, Xu Y. Energy and carbon emissions analysis and prediction of complex petrochemical systems based on an improved extreme learning machine integrated interpretative structural model. Appl Therm Eng 2017;115:280–91.
[114] Han YM, Wu H, Jia MH, Geng ZQ, Zhong YH. Production capacity analysis and energy optimization of complex petrochemical industries using novel extreme learning machine integrating affinity propagation. Energy Convers Manag 2019;180:240–9.
[115] Geng ZQ, Zhang YH, Li CF, Han YM, Cui YF, Yu B. Energy optimization and prediction modeling of petrochemical industries: an improved convolutional neural network based on cross-feature. Energy 2020;(194). 0360-5442.
[116] Han YM, Fan CY, Xu M, Geng ZQ, Zhong YH. Production capacity analysis and energy saving of complex chemical processes using LSTM based on attention mechanism. Appl Therm Eng 2019;(160):114072.
[117] Geng ZQ, Zhu QX, Gu XB, Lin XY. Optimal control of cracking depth based on multi-swarm competitive PSO-RBFNN for ethylene cracking furnace. J Chem Ind Eng 2010;(8):1942–8.
[118] Li CF, Zhu QX. Fuzzy particle swarm artificial neural network and its application. Computers and Applied Chemistry 2007;24:1359–62.
[119] Liao TT, Li GQ, Li RX. Operation optimization of absorption & stabilization system based on ANN and GA with seeking constraint. Acta Pet Sin 2017;33(6):1138–45.
[120] Lang XM, Qu BC, Yang Y, Liu XM. Application of ant colony algorithm with immune algorithm in optimization of reactive distillation. Petrochemical Technology & Application 2010;(2).
[121] Reynolds J, Ahmad MW, Rezgui Y, Hippolyte JL. Operational supply and demand optimisation of a multi-vector district energy system using artificial neural networks and a genetic algorithm. Appl Energy 2019;235:699–713.
[122] Kong WJ, Liu SY, Du RS. Application of neural network and genetic algorithm in traditional chemical enterprises. Comput Eng 2003;(17):193–4.
[123] Palagi L, Sciubba E, Tocci Lorenzo. A neural network approach to the combined multi-objective optimization of the thermodynamic cycle and the radial inflow turbine for Organic Rankine cycle applications. Appl Energy 2019;237:210–26.
[124] Arshad MU, Ghani MU, Ullah A, Güngör A, Zaman Muhammad. Thermodynamic analysis and optimization of double effect absorption refrigeration system using genetic algorithm. Energy Convers Manag 2019;192:292–307.
[125] He XY, Li J, Qi L, Han YX, Duan JX. Mechanism of distillation tower-neural network hybrid construction. Contr Eng 2009;16:211–3.
[126] Yu LJ, Zhang LB, Gu XS. Modeling and parameter estimation of conversion mechanism of methanol synthesis tower based on improved wolf group Algorithm. J East China Univ Sci Technol 2017;43:815–23.
[127] Liu YY, He DF, Feng K, Lu XX. Input-output hybrid model of converter based on mechanism and data. Nonferrous Metal Science and Engineering 2018;9: 13–8.
[128] Xu D, Liu SX. Study on energy consumption modeling of steelmaking-continuous casting with data driving and mechanism model. Hebei Metallurgy 2018;10:14–9.
[129] Li JP, Hua CC, Guan XP. Modeling of large-scale blast furnace smelting process based on mechanism, data and knowledge. J Shanghai Jiaot Univ 2018;52:1142–54.
[130] Sun L, Hua Q, Shen J. Multi-objective optimization for advanced superheater steam temperature control in a 300 MW power plant. Appl Energy 2017;208:592–606.