Energy Conversion and Management 180 (2019) 240–249

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Production capacity analysis and energy optimization of complex petrochemical industries using novel extreme learning machine integrating affinity propagation

Yongming Han[a,b], Hao Wu[a,b], Minghui Jia[a], Zhiqiang Geng[a,b,*], Yanhua Zhong[c]

[a] College of Information Science & Technology, Beijing University of Chemical Technology, Beijing 100029, China
[b] Engineering Research Center of Intelligent PSE, Ministry of Education in China, Beijing 100029, China
[c] Jiangmen Polytechnic, Jiangmen, Guangdong 529020, China

ARTICLE INFO

Keywords:
Affinity propagation
Extreme learning machine
Production capacity analysis
Energy optimization
Emissions reduction
Complex petrochemical industries

ABSTRACT

With the rapid political and economic developments, energy saving and carbon dioxide emission reduction are now recognized as the most important goals worldwide. Therefore, this paper presents a production capacity analysis and energy optimization model that uses novel extreme learning machine integrating affinity propagation clustering. By using the affinity propagation method, clusters of raw data can be obtained automatically to reduce multi-dimensional data and fuse the high-similarity data. Then, the clustering results are taken as the training and testing sets of the extreme learning machine to improve prediction accuracy and analyze production capacity. Through comparisons with the extreme learning machine, back propagation, radical basis function, and extreme learning machine integrated K-means clustering algorithm, the accuracy and validity of the proposed method are verified by using the University of California Irvine data sets. Finally, the proposed method is applied to build the production capacity analysis and energy optimization model of the ethylene and purified terephthalic acid production systems in complex petrochemical industries. The prediction accuracy of the ethylene and purified terephthalic acid production is approximately 99%, thus improving the energy efficiency of these complex petrochemical processes and achieving energy saving and carbon dioxide emission reduction.

1. Introduction

The petrochemical industry is a large energy-intensive industry. As "the food of the petrochemical industry [1]," ethylene is the basic organic chemical raw material for the manufacture of synthetic materials, synthetic fibers, and other products. The yield of the ethylene production process corresponds to the level of a country's petrochemical industry. In 2015, the ethylene production and average fuel and power consumption of the China Petrochemical Corporation were, respectively, 11005.2 kt/a and 559.06 kg/ton of ethylene [2], whereas the China National Petroleum Corporation attained 5032 kt/a and 594 kg/ton of ethylene [3]. However, the energy efficiency of the ethylene production in China is far lower than that achieved in other countries. Therefore, it is necessary to improve the energy efficiency of the ethylene production processes in petrochemical enterprises [4].

The purified terephthalic acid (PTA) is also an important organic raw material and is widely used in all aspects of the national economy. In recent years, the PTA production in China has been constantly expanding [5]. However, there is still a great potential for improvement in energy saving and carbon dioxide (CO_2) emission reduction [6]. Therefore, if the production capacity can be accurately analyzed and the energy configuration can be optimized in the petrochemical industry, significant energy saving and CO_2 emission reduction can be achieved.

To analyze and optimize the energy status of the petrochemical industry, this study proposes a production capacity analysis and energy optimization model based on a novel extreme learning machine (ELM) integrating affinity propagation (AP) clustering (AP-ELM). Clusters of raw data can be obtained automatically by using the AP clustering method to reduce multi-dimensional data and fuse high similarity data. Moreover, the clustering results are taken as the training and testing sets of the ELM to improve the prediction accuracy and analyze the production capacity and optimal energy configuration. Through comparisons with the ELM, back propagation (BP), radical basis function (RBF), and ELM integrated K-means clustering algorithm (K-means-ELM), the accuracy and validity of the proposed method are verified by

* Corresponding author at: College of Information Science & Technology, Beijing University of Chemical Technology, Beijing 100029, China.
E-mail address: gengzhiqiang@mail.buct.edu.cn (Z. Geng).

https://doi.org/10.1016/j.enconman.2018.11.001
Received 12 May 2018; Received in revised form 1 November 2018; Accepted 1 November 2018

Nomenclature

CO_2	carbon dioxide
ELM	extreme learning machine
AP	affinity propagation
BP	Back Propagation
RBF	Radical Basis Function
PTA	the purified terephthalic acid
K-means-ELM	extreme learning machine integrated K-means clustering algorithm
UCI	University of California Irvine
SDSM	sample data selection method
MLP	Multi-Layer Perceptron
GA	the genetic algorithm
PCA	principal component analysis
FCM	the fuzzy C-means
FAHP	Fuzzy C-Means (FCM) integrating analytic hierarchy process
EOS-ELM	Ensemble of Online Sequential Extreme Learning Machine
CEOS-ELM	constructive enhancement for the online sequential Extreme Learning Machine
PV	photovoltaic
SVM	support vector machine
TAP	transfer affinity propagation
S (i, k)	the similarity matrix
P (i)	the bias parameter
R (i, k)	responsibility
A (i, k)	availability
W	the weights of the input layer and the hidden layer
V	the weights of the hidden layer and the output layer
b	the threshold
A	the adjacency matrix
$g(x)$	the activation function
H	hidden layer output matrix
Y	the output
H^+	the Moore-Penrose of the hidden layer output matrix
$\widehat{V}$	the hidden layer's output weight
NAP	naphtha
CTA	crude terephthalic acid
RAF	raffinate
C3	carbon 3
C4	carbon 4
C5	carbon 5
NBA	N-butyl acetate
HDL	hydrogenation tail oil
LDL	light diesel oil
LHY	light hydrogenation tail oil
OTH	other force
PX	purity paraxylene
TA	terephthalic acid
FC	feed composition
FC1501	Feed quantity
FC1502	water reflux
FC1503	NBA main reflux
FC1504	NBA side reflux
FC1507	steam flow
FI1511	produced quantity of the top tower
TI1504	feed temperature
TI1510	reflux temperature
Tl1511	temperature of the top tower
TI1515	temperature point above the 35th tray
TI1516	temperature point between the 35th tray and the 40th tray
Tl1517	temperature point between the 44th tray and the 50th tray
TC1503a	tray temperature near the up sensitive plate
TC1503b	tray temperature near the low sensitive plate
TC1501	temperature point between the 53rd tray and the 58th tray
LC1503a	reflux tank level

using the University of California Irvine (UCI) data sets [7]. Finally, the proposed method is used to build the production capacity analysis and energy optimization model of the ethylene and PTA production systems in complex petrochemical industries. The prediction accuracy of the ethylene and PTA production is approximately 99% and the energy saving potential in the ethylene and PTA production systems is approximately 69% and 68%, respectively. Furthermore, the CO_2 emission reduction in the ethylene production system is approximately 73,802 ton.

2. Related works

In recent years, with the improvement in computer technology, many artificial intelligence methods, especially neural network methods such as BP and RBF, have been proposed to analyze and predict the production capacity of the petrochemical industry. Yuan et al. proposed a novel sample data selection method (SDSM) combining the similar days selection with the virtual samples generation to improve the prediction accuracy [8]. For accurately calculating the CO_2 emission in utility boilers, the BP neural network was proposed by Yin et al. to predict the carbon content of burning coal [9]. Wang et al. proposed a short-term solar irradiance prediction model based on the BP neural network [10]. Combining the chaos-search genetic algorithm and the simulated annealing algorithm, Badar et al. proposed a modified BP neural network to predict the short-term electrical energy demand in the deregulated power industry [11]. Wen et al. established a reasonable model to study the impact factors of carbon emissions in Beijing and carry out forecasts for pollution prevention [12].

The RBF also has many advantages such as simple modeling and rapid training. Hejase et al. proposed an RBF method integrating the multi-layer perceptron (MLP) to estimate the global horizontal irradiance for three major cities in the United Arab Emirates [13]. By optimizing the number of hidden neurons and selecting centers in a wide range of data, the RBF can be trained to deliver maximum fitting and computational performance [14]. Tang et al. studied the relation between experimental results and experimental conditions in the electrosynthesis process by using the RBF and genetic algorithm (GA) [15]. Shi et al. developed an improved RBF to differentiate spill oil samples [16], which was particularly suitable for real-time spill oil identification and could also be easily applied in oil logging and analysis of multi-fluorescent mixtures. When determining the hidden-layer centers with an unsupervised learning algorithm, Hu et al. presented a new distance measure using Gaussian basis [17]. Guo et al. proposed a novel adaptive RBF sliding mode strategy to control Lorenz chaos with parametric uncertainties and external disturbances [18]. Geng et al. proposed an improved RBF based on the fuzzy C-means (FCM) algorithm integrating the principal component analysis (PCA) technology to optimize the energy structure of ethylene plants [19]. The training time of the conventional BP and RBF methods is long, their convergence speed is low, and the result easily falls into a local minimum. Therefore, the ELM was presented.

The ELM was proposed by Huang et al. in 2004 [20]. The predictive method based on the ELM integrating the fuzzy C-means (FCM) and analytic hierarchy process was proposed by Geng et al. to analyze the energy efficiency of the ethylene plant [21]. An online condition forecasting method for natural circulation systems based on the ensemble of online sequential ELM (EOS-ELM) was proposed by Chen et al. [22]. A new model based upon the ELM was presented by Kasra et al. to

estimate the wind power density [23]. Because the online property prediction in industrial rubber mixing processes was not an easy task, Gao et al. proposed a fast property prediction method in an industrial rubber mixing process using the local ELM model [24]. However, it is difficult to perform effective energy modeling owing to the characteristics of uncertainty and high nonlinearity in practical production. Therefore, Zhu et al. presented an energy-saving potential analysis method using a novel ELM logic network [25]. Lan et al. proposed a constructive enhancement for the online sequential ELM (CEOS-ELM) [26], which could add random hidden nodes one-by-one or group-by-group with fixed or varying group size. Although the ELM offers a significant improvement in predictive accuracy and speed, its prediction accuracy is still not very high. Therefore, some improved ELM methods based on clusters, such as the K-means clustering and AP clustering algorithm, have been presented. Ali et al. combined the K-means algorithm and a type of network to forecast the hourly global horizontal solar radiation [27]. Tehrani et al. introduced a novel approach based on the minimum conditional entropy clustering and showed its advantages in terms of data analytics by modifying the conventional sequential K-means algorithm [28]. Mahesh et al. proposed an efficient initial seed selection method to improve the performance of the K-means filtering method by locating the seed points [29]. Li et al. presented a novel dynamic clustering equivalent modeling method for a two-staged photovoltaic (PV) station cluster [30], which is a key tool to analyze the dynamic responses of the distribution network. However, the K-means clustering method needs that the number of clusters be set, which implies poor objectivity. Therefore, the AP clustering algorithm has been proposed, which does not require that the number of clusters be specified. To improve accuracy and reduce time consumption, a method based on AP clustering that integrates the principal component analysis (PCA) was proposed by Lei et al. [31]. A new identification method of genetically modified material by the support vector machine (SVM) based on AP clustering was proposed by Liu et al. for conducting cluster analysis and labeling on unlabeled training samples [32]. Guerra et al. presented a computer-aided interpretation of the M-N plot based on a new hybrid algorithm that refined the AP clustering technique [33]. A novel transfer AP algorithm known as TAP was presented by Hang et al., which could modify the update rules for the two-message propagations [34]. Because the computational complexity of the traditional source localization was high, Wang et al. proposed an AP algorithm based on the localization method [35]. Distinct from the standard clustering-based partition approaches, a modified exemplar based on AP was presented by Zhao et al. to address the problem [36]. Therefore, this study proposes a production capacity analysis and energy optimization model based on a novel AP-ELM method to save energy and reduce carbon emissions in complex petrochemical industries.

3. Extreme learning machine based on affinity propagation

The novel AP-ELM method is proposed in this section. The clusters of raw data can be obtained automatically based on the AP clustering method to reduce multi-dimensional data and fuse high similarity data. Then the clustering results are taken as the training and testing sets of the ELM to improve the prediction accuracy.

3.1. Affinity propagation

AP clustering is a new adjacent propagation clustering proposed by Frey and Dueck in 2007 [37]. Its purpose is to find the optimal set of class representative points so that the sum of the similarities of all data points to the nearest class representative point is the largest [38]. Compared with K-means clustering and other clustering algorithms, AP clustering has many unique features:

(1) There is no need to specify the number of clusters.
(2) There is no need to determine the cluster center in advance.
(3) Symmetry of the distance matrix is not required and the scope of data is very large.
(4) It is insensitive to the initial value. There is no need to randomly select initial values.
(5) The deviation of the AP clustering is lower than that of other clustering methods.

AP clustering has two important parameters. One is the similarity matrix S (i, k). S (i, k) represents the suitability of the data point X_k as the cluster center of the data point X_i. The similarity matrix can be symmetric or asymmetric so that the algorithm could expand the range of the clustering data. When i = k, S (i, k) represents the bias parameter P (i). The larger P (i) is, the greater the probability that the X_k point will be the center of the cluster. During initialization, all P (i) have the same value, that is to say, the probability of all data points to become clustering centers is the same [39]. Another parameter is the messaging function. The AP algorithm has two important messaging functions: responsibility (R (i, k)) and availability (A (i, k)). The expression of R (i, k) is shown as Eq. (1).

$$R_{t+1}(i,k) = (1-\lambda)\cdot R_{t+1}(i,k) + \lambda \cdot R_t(i,k) \tag{1}$$

$$R_{t+1}(i,k) = \begin{cases} S(i,k) - \max\limits_{j\neq k}\{A_t(i,j) + R_t(i,j)\}, & i \neq k \\ S(i,k) - \max\limits_{j\neq k}\{S(i,j)\}, & i = k \end{cases} \tag{2}$$

The expression of A (i, k) is shown as Eq. (3).

$$A_{t+1}(i,k) = (1-\lambda)\cdot A_{t+1}(i,k) + \lambda \cdot A_t(i,k) \tag{3}$$

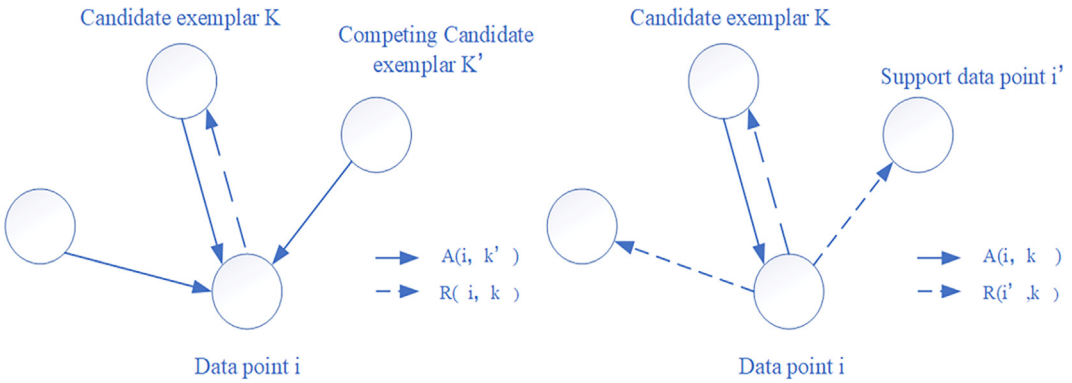

Fig. 1. The process of two-message transfer function.

$$A_{t+1}(i,k)=\begin{cases}\min\left\{0, R_{t+1}(k,k)+\sum_{j\notin\{i,k\}}\max\{0, R_{t+1}(j,k)\}\right\}, & i\neq k\\ \sum_{j\neq k}\max\{0, R_{t+1}(j,k)\}, & i=k\end{cases} \tag{4}$$

Here, R (i, k) describes the suitability degree to which data object k is suitable as a clustering center for data object i, which represents the message from i to k. A (i, k) describes the suitability of data object i that selects data object k as its center of clustering, which represents the message from k to i. Then, the iteration starts from R (i, k). The process of determining the center of clustering by using two message transfer functions is described in Fig. 1. The larger R (i, k) and A (i, k) are, the more suitable data object k is to be the center of the cluster.

The algorithm takes all data points as potential clustering centers in the clustering process, which avoids the initial selection of cluster centers and improves the reliability of the clustering results. Supposing that N data are clustered, the similarity amount of these N data determines the similarity matrix S (N * N). The AP algorithm iteratively updates the responsibility and availability of each data object. After a certain number of iterations, m high-quality exemplars (cluster centers) are obtained. At the same time, the remaining data objects are assigned to the corresponding clusters. During update R (i, k) and A (i, k), the AP algorithm introduces a damping coefficient lam ∈ [0, 1) to avoid oscillation. The larger the damping coefficient, the better the effect of eliminating oscillations, but it will slow down the convergence rate of the AP algorithm, and typically the value is 0.9.

The implementation steps of AP clustering are described as follows:

Step 1: Determine the similarity matrix S and set all the data points as potential clustering centers.
Step 2: Iteratively update the responsibility and the availability of each data object.
Step 3: Determine the cluster center X_k of the data point X.
Step 4: The algorithm is over when the selected class representative point exceeds the set threshold during successive iterations.

3.2. Extreme learning machine

The ELM can approach arbitrary linear and non-linear functions with zero error. Besides, the ELM has excellent generalization ability, and its learning speed is much faster than that of BP.

The ELM is a simple and effective single hidden layer feedforward neural network. Compared with the BP and the RBF, the weight matrix of the input and middle layers of the ELM is directly generated during the training process by the Gaussian distribution. The weight matrix between the hidden layer and the output layer needs to be calculated [40]. Thus, it has the advantages of easy parameter selection, fast learning, good generalization ability, and strong robustness [41,42]. The ELM consists of the input, middle, and output layers. The input layer contains n-dimensional vectors, the hidden layer contains L nodes (typically, L is much smaller than n), and the output contains an m-dimensional vector. The characteristic of the ELM is that only the number of hidden-layer nodes needs to be set, and then the only optimal solution can be obtained. The output weights of the ELM are obtained by calculating a matrix of Moore–Penrose. That is to say, the parameters of the ELM do not require any iterative steps in the process. Therefore, the network parameter adjustment time is greatly reduced. The ELM network structure is shown in Fig. 2.

The connection weights of the input and hidden layers and the hidden and output layers are set as follows:

$$W=\begin{bmatrix}W_1\\W_2\\\vdots\\W_L\end{bmatrix}=\begin{bmatrix}W_{11} & W_{12} & \cdots & W_{1n}\\W_{21} & W_{22} & \cdots & W_{2n}\\\cdots & \cdots & \ddots & \vdots\\W_{L1} & W_{L2} & \cdots & W_{Ln}\end{bmatrix}_{L\times n} \tag{5}$$

$$V=\begin{bmatrix}V_1\\V_2\\\vdots\\V_L\end{bmatrix}=\begin{bmatrix}V_{11} & V_{12} & \cdots & V_{1m}\\V_{21} & V_{22} & \cdots & V_{2m}\\\cdots & \cdots & \ddots & \vdots\\V_{L1} & V_{L2} & \cdots & V_{Lm}\end{bmatrix}_{L\times m} \tag{6}$$

The threshold b for the hidden layer is obtained by Eq. (7):

$$b=\begin{bmatrix}b_1\\b_2\\\vdots\\b_L\end{bmatrix} \tag{7}$$

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

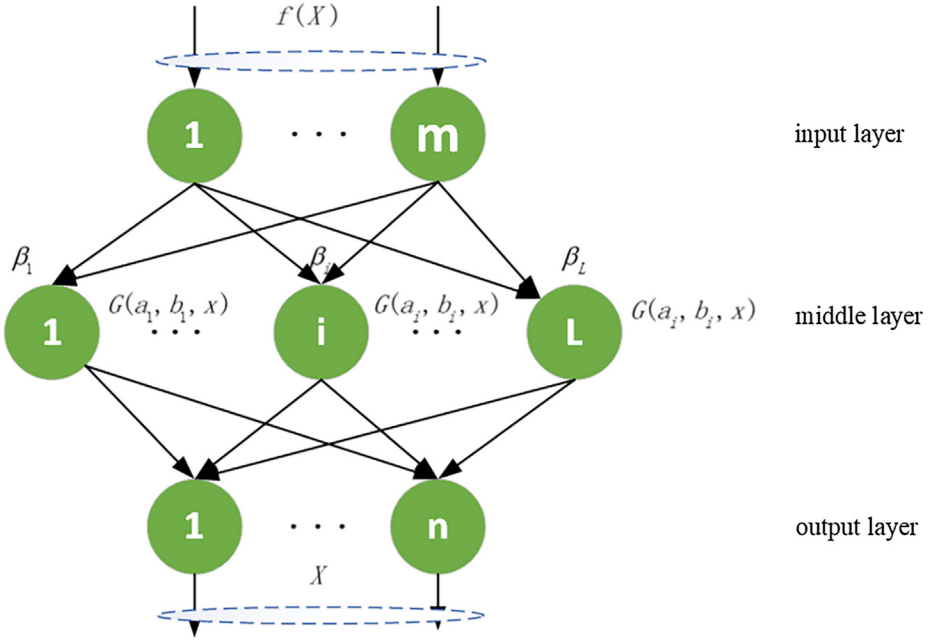

Fig. 2. ELM network structure.

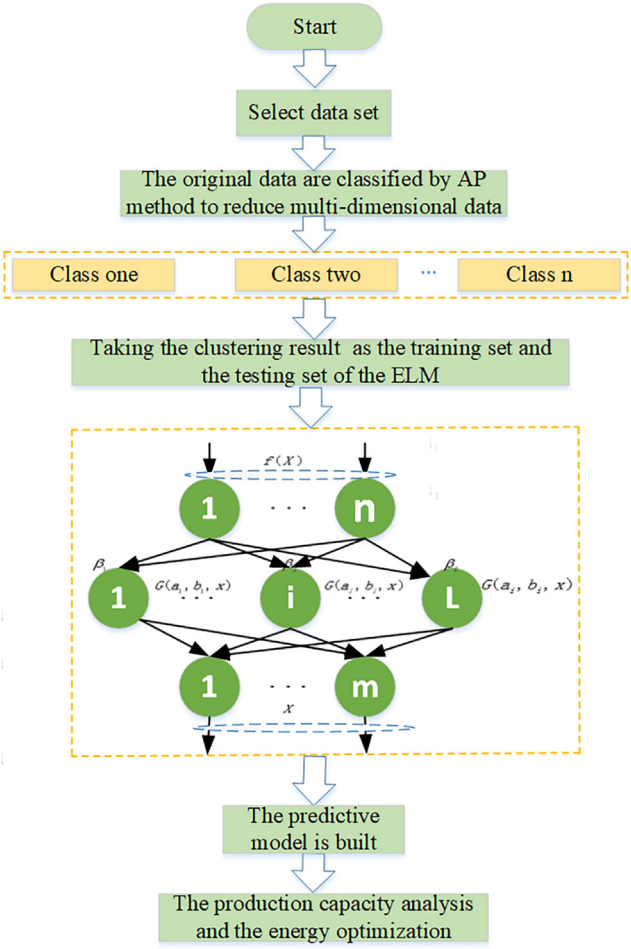

Fig. 3. Flow chart of the AP-ELM algorithm.

Table 1
UCI data set samples clustered by AP-ELM.

Data Sets	Training	Testing	Inputs	Outputs
Wine	119	10	3	1
Residential-Building	199	12	3	1

$$Y = HV \tag{8}$$

H: Hidden layer output matrix.
The expression of H: is as follows:

$$H = \begin{bmatrix} g(w_1x_1 + b_1) & g(w_2x_1 + b_2) & \cdots & g(w_Lx_1 + b_L) \\ g(w_1x_2 + b_1) & g(w_2x_2 + b_2) & \cdots & g(w_Lx_2 + b_L) \\ \cdots & \cdots & \ddots & \vdots \\ g(w_1x_m + b_1) & g(w_2x_m + b_2) & \cdots & g(w_Lx_m + b_L) \end{bmatrix} \tag{9}$$

where w and b are arbitrarily specified before training and remain unchanged during the training process. Eqs. (6), (8), and (9) show that the output weight of the hidden layer is

$$\hat{V} = H^{+}Y \tag{10}$$

where H^{+} is the Moore–Penrose of the hidden layer output matrix. Therefore, the ELM simplifies the training process of complex neural networks to the problem of matrix inversion, which greatly improves the learning speed [43].

3.3. Framework of the extreme learning machine integrating affinity propagation

The steps of the AP-ELM method are as follows:

Step 1: Select the data sets that need to be processed.
Step 2: After the data are clustered by using the AP clustering algorithm, a number of new data sets with high similarity are obtained.
Step 3: Take the clustering results as the training and testing sets of the ELM.
Step 4: Train the ELM neural network by using the training set.
Step 5: Use the testing samples, which are different from the training samples, as a generalization sample of the AP-ELM. Comparing the generalization output and the expected output, the generalization relative error and the generalization standard deviation are calculated.
Step 6: Analyze the results.

The flow chart of the AP-ELM method is shown in Fig. 3.

4. The University of California Irvine test validation

The UCI Machine Learning Repository is a commonly utilized standard test data set, which can be used for classification, regression, clustering, and recommendation of the system task. The UCI includes 378 different sizes and types of data sets. In order to verify the effectiveness and accuracy of the proposed model, this study extracted two groups of UCI data sets [44], as presented in Table 1.

Firstly, the Wine and Residential-Building data are clustered by AP clustering, and then the Wine and Residential-Building inputs are clustered to three classes, which are divided into the training and testing sets of the AP-ELM. The comparative results of the Wine and Residential-Building are presented in Tables 2 and 3, respectively.

In the Wine, the results of the AP-ELM are clearly the best. Compared with the ELM, K-means-ELM, BP, and RBF, the prediction accuracy of the AP-ELM improved 11%, 6%, 14% and 29%, respectively. In the Residential-Building, the results of the AP-ELM are also clearly the best. The prediction accuracy of the AP-ELM improved 1.1%, 0.8%, 1.3% and 5%, respectively. The results verify the efficiency and accuracy of the proposed method.

5. Case study: production capacity analysis and energy optimization of complex petrochemical processes

In the complex petrochemical industries, ethylene is an important indicator of the condition of the national economy and PTA is an important organic raw material. Therefore, a major concern has been how to increase the production of ethylene and PTA and reduce the consumption of energy. The proposed model is used to predict production and save energy in the ethylene and PTA production processes.

Table 2
Comparative results of the wine test.

	The AP-ELM	The K-means-ELM	The ELM	The BP	The RBF
The number of training sets	119	119	119	119	119
The number of testing sets	10	10	10	10	10
Generalization error	0.02	0.08	0.13	0.16	0.31

Table 3
Comparative results of the residential-building test.

	The AP-ELM	The K-means-ELM	The ELM	The BP	The RBF
The number of training sets	199	199	199	199	199
The number of testing sets	12	12	12	12	12
Generalization error	0.060	0.068	0.071	0.073	0.31

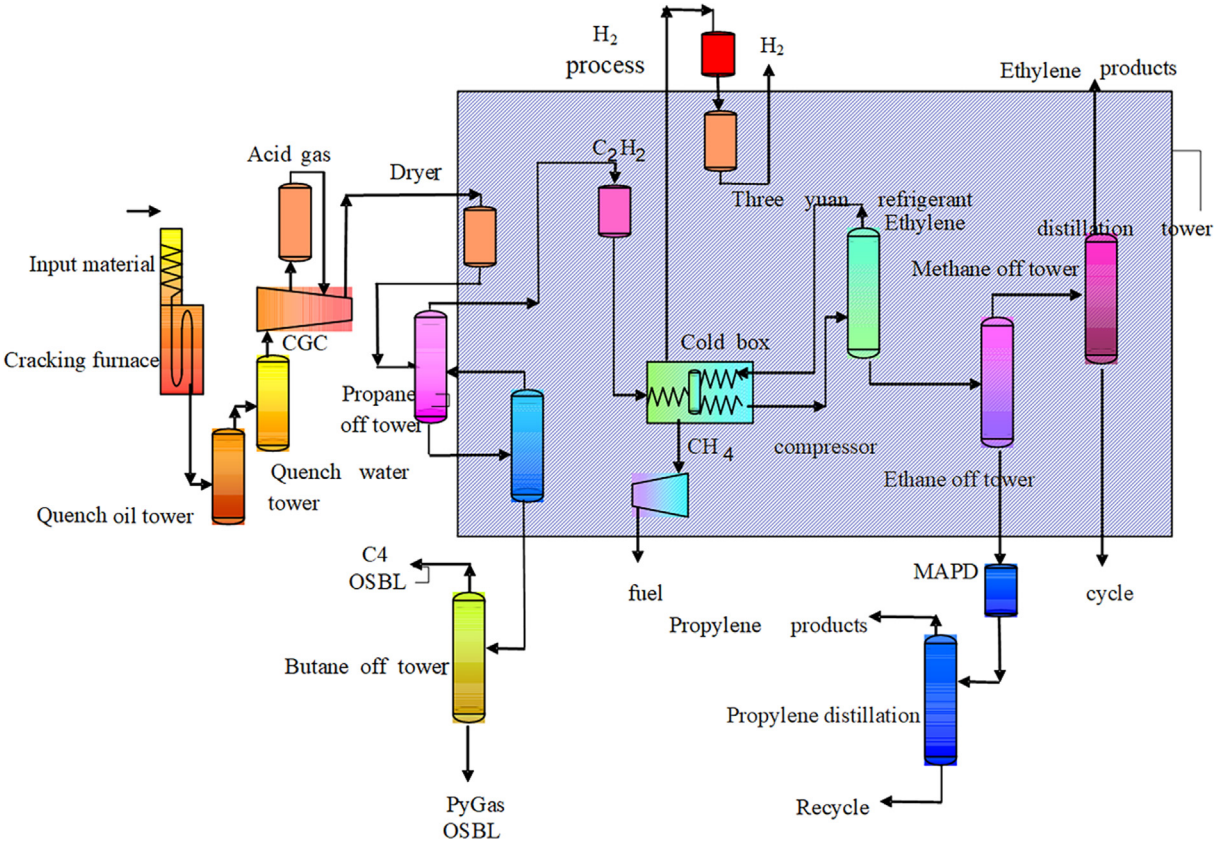

Fig. 4. Flow chart of the ethylene production plant.

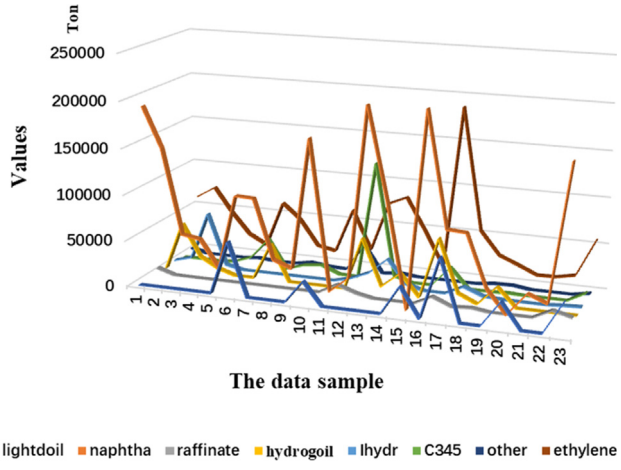

Fig. 5. Raw materials and yield of ethylene in production plants in 2009–2013.

5.1. Production prediction and energy optimization in the ethylene industry

In recent years, China's ethylene industry has had a rapid development and its production of ethylene is the second worldwide. In the ethylene production process, the yield and consumption of the raw materials closely related to the cost of ethylene have been the subject of international concern. The main raw materials of the ethylene production in China are naphtha (NAP); light diesel oil (LDL); raffinate (RAF); hydrogenation tail oil (HDL); carbon3, carbon4, and carbon5 (C345); light hydrogenation tail oil (LHY); and other materials [45].

5.1.1. The ethylene production plant

The ethylene production plant contains cracking and separation parts. The cracking phase is the main core of the entire ethylene plant, and it consumes most of the required energy. The separation section mainly includes the rapid cooling, compression, and separation processes. With respect to the distribution of energy consumption in the ethylene production process, many domestic and foreign scholars have conducted numerous studies on the ethylene cracking furnaces, mainly on two aspects: first, the energy supply. Under the same output and conditions, the raw material consumption is reduced. Second, the valid product output. Under the same energy consumption conditions, the ethylene yields are improved. In Fig. 4, the flow chart of an ethylene production plant is depicted.

5.1.2. Ethylene data analysis

The experimental data of the ethylene production industry is the monthly data from 2009 to 2013. The ratio of raw materials and ingredients, which include crude oil, fuel, steam, water, and electricity, to

Table 4
Average error analysis.

	The AP-ELM	The K-means-ELM	The ELM	The BP	The RBF
The number of training sets	499	499	499	499	499
The number of testing sets	12	12	12	12	12
Average error	0.07	0.078	0.09	0.12	0.17

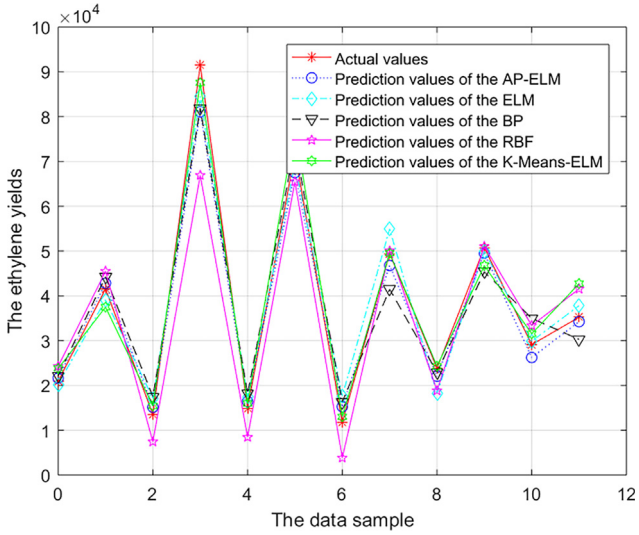

Fig. 6. Actual and predicted values of ethylene production plants.

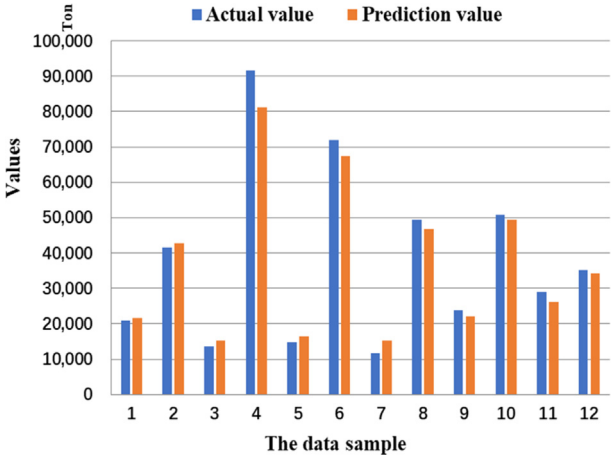

Fig. 7. Yields of ethylene production plants predicted by the AP-ELM.

ethylene production is very complex. As the inputs of ethylene production, the main raw materials are seven: NAP, C345, RAF, HDL, LDL, LHY, and other materials (OTH). The output is the ethylene yield. Fig. 5 shows the use of raw materials and the yield of ethylene in production plants in 2009–2013.

5.1.3. Energy saving and carbon dioxide emission reduction in ethylene production plants

The experiment contains 860 sets of the ethylene industrial data. The training and testing sets of the K-means-ELM, RBF, BP, ELM, and AP-ELM contain 499 and 12 sets, respectively. The average error analysis of the BP, K-means-ELM, RBF, ELM, and AP-ELM in predicting ethylene production is presented in Table 4.

The comparative results between the actual and predicted values of the ethylene production plants are shown in Fig. 6.

It can be seen from Fig. 6 and Table 4 that the AP-ELM method is the best to predict the ethylene yield. The outputs of ethylene production

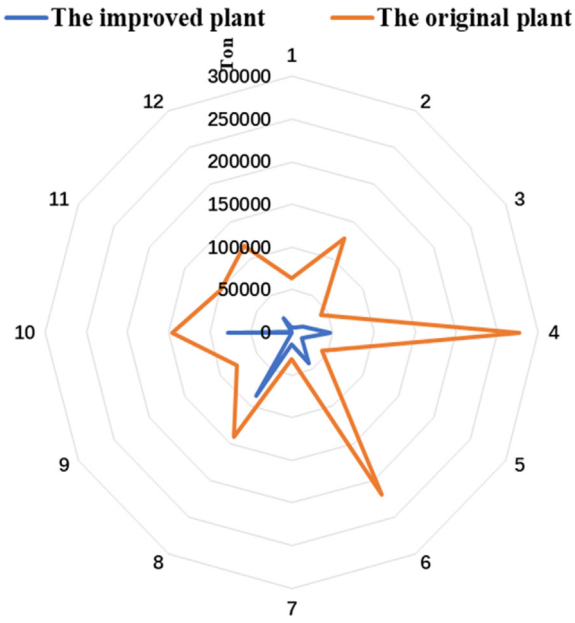

Fig. 8. Comparison of total material consumption in ethylene production plant.

plants based on the AP-ELM method are shown in Fig. 7.

Based on the energy efficiency data from 2009 to 2013, in which the ID of data is equal to 0, this study built the predictive model and predicted the ethylene yield. It can be seen from Table 4 that the average error of the AP-ELM decreased 0.8%, 2%, 5%, and 10%, compared with that of the K-means-ELM, ELM, BP, and RBF, respectively. Moreover, from Fig. 7, it can be seen that the AP-ELM method can reach better performance in predicting the ethylene yield. The comparison of total material consumption of the ethylene production between the improved and original plants is shown in Fig. 8, which shows that the consumption of raw materials of the improved plant is much lower than that of the original plant. Taking sample 2 as example, the improved plant uses 3 ton of RAF and 6246 ton of LHY to produce the ethylene. However, the original plant uses 81,080 ton of NAP, 3 ton of RAF, 39,957 ton of HDL, and 6246 ton of LHY. The actual value of the ethylene yield is 41,528 ton, and the predicted value is 42,804 ton. Meanwhile, the input of crude oil is reduced by 121,938 ton. Furthermore, according to the CO_2 emission factor for different types of fuels [46], the improved plant will reduce by 73,802 ton of CO_2. Based on the result analysis, the original plant should adjust the inputs for ethylene production to improve energy efficiency in the future.

5.2. Production prediction and energy saving in the purified terephthalic acid industry

With the rapid development of polyester and fibers in our country, the deficiencies in PTA production have become increasingly prominent. In recent years, a number of large-scale PTA production units have been completed, and the production capacity of PTA production plants has expanded rapidly. The PTA production quality and stability are improving, which places the PTA industry in China into an important stage.

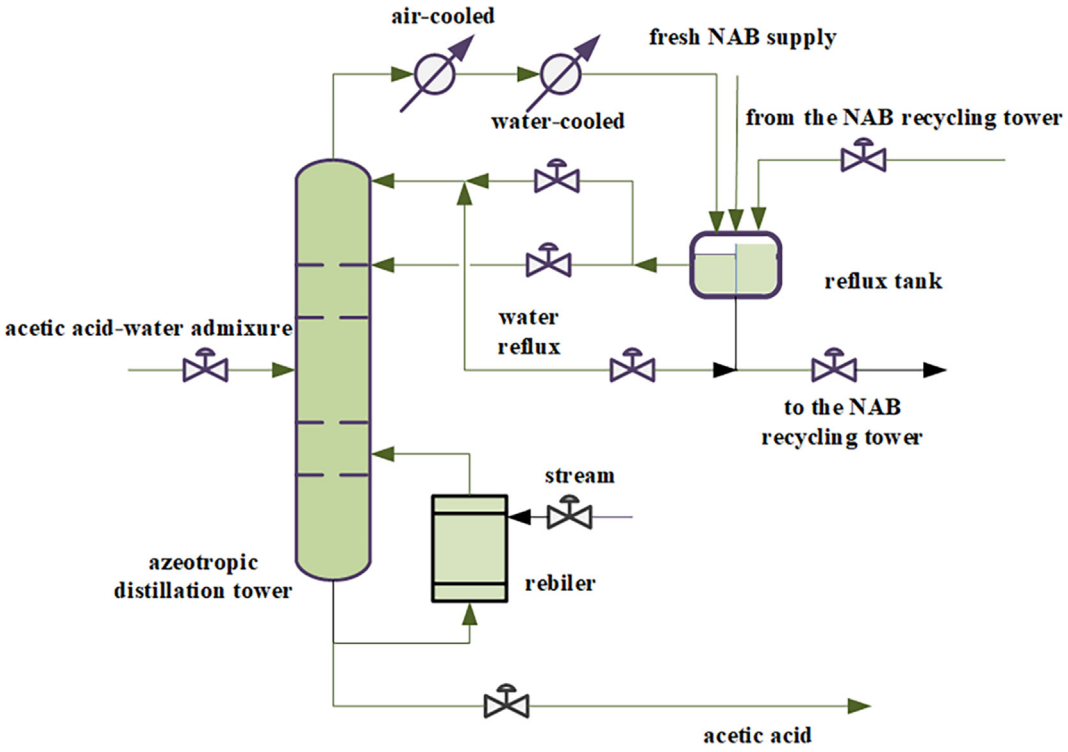

Fig. 9. Flow chart of the PTA production plant.

Table 5
Input variables in PTA production plants, with abbreviations.

No.	Input variable	Noun abbreviation
1	Feed composition	FC
2	Feed quantity	FC1501
3	Water reflux	FC1502
4	N-butyl acetate (NBA) main reflux	FC1503
5	N-butyl acetate (NBA) side reflux	FC1504
6	Steam flow	FC1507
7	Produced quantity of the top tower	FI1511
8	Feed temperature	TI1504
9	Reflux temperature	TI1510
10	Temperature of the top tower	Tl1511
11	Temperature point above the 35th tray	TI1515
12	Temperature point between the 35th tray and the 40th tray	TI1516
13	Temperature point between the 44th tray and the 50th tray	Tl1517
14	Tray temperature near the up sensitive plate	TC1503a
15	Tray temperature near the low sensitive plate	TC1503b
16	Temperature point between the 53rd tray and the 58th tray	TC1501
17	Reflux tank level	LC1503a

5.2.1. The purified terephthalic acid production plant

In China, the PTA production is divided in two parts. The first step includes high purity paraxylene (PX) as raw material, acetic acid as solvent, cobalt acetate and manganese acetate as catalyst, tetrabromoethane or hydrobromic acid, which promotes the catalytic oxidation of air, and after separation and drying, the crude terephthalic acid (TA) is obtained. In the second step, through the principle of hydrogenation reduction and under high temperature and pressure, the crude terephthalic acid impurities are removed by the catalyst. Finally, the higher PTA is obtained.

The acetic acid does not participate in the main reaction but in the oxidation process, as it is an important solvent in oxidation reactors. Therefore, the process conditions should be optimized so that, when the crude TA meets the requirements in the 4-CBA index, it can minimize the need of acetic acid. The different patented agents in the refining part of the production process are basically similar, and the main difference is in the oxidation part. The level of the oxidation process directly determines the material and energy consumption levels for the PTA production. The flow chart of the PTA production plant is shown in Fig. 9.

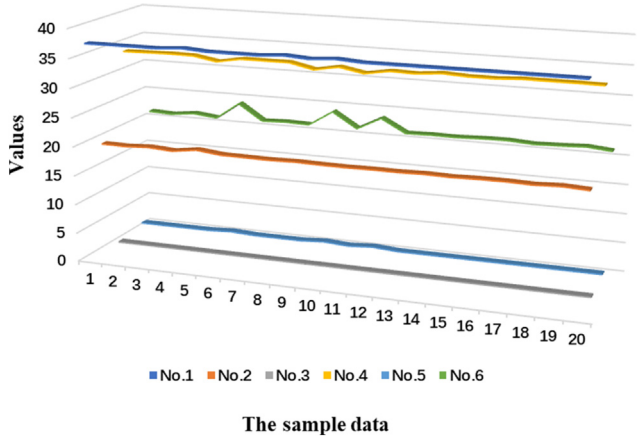

Fig. 10. Partial data samples of PTA production plants.

5.2.2. Data analysis of the purified terephthalic acid

The main factors affecting the PTA solvent system are listed in Table 5. The following 17 variables are used as the input, and the PTA yield is the only output. Based on the PTA production plant analysis, the feed amount FC1501, temperature TI1504, reflux amount FC1502, FC1503, FC1504, temperature TI15010, reboiler steam flow FC1507, and tower internal temperature TI1511–TI1519 and TC1501 are the influencing factors of the acetic acid content in PTA production plants [47]. The partial materials in PTA production plants are shown in Fig. 10.

Table 6
Average error analysis.

	The AP-ELM	The K-means-ELM	The ELM	The BP	The RBF
The number of training sets	179	179	179	179	179
The number of testing sets	9	9	9	9	9
Average error	0.0016	0.0023	0.0032	0.0037	0.0045

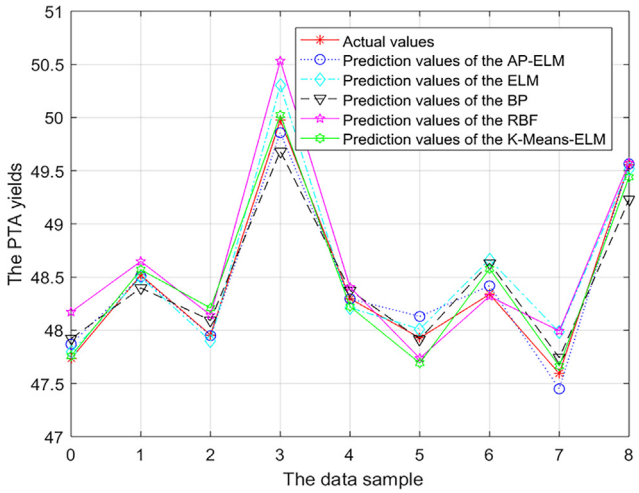

Fig. 11. Actual and predicted values of PTA production plants.

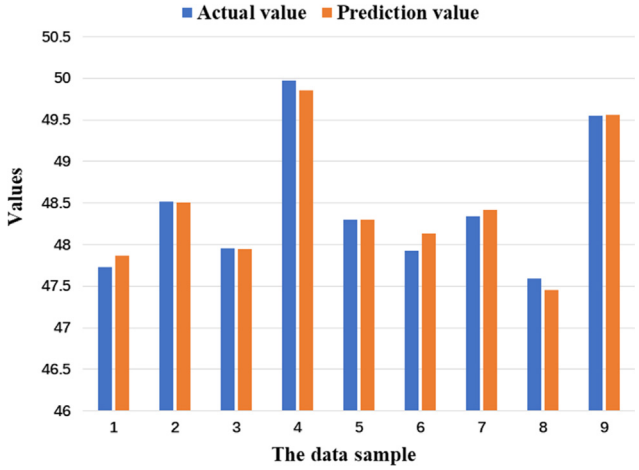

Fig. 12. Outputs of the AP-ELM method in PTA production plants.

5.2.3. Energy saving and optimization of purified terephthalic acid production plants

The average error analyses of the PTA yield based on the AP-ELM, K-means-ELM, ELM, BP, and RBF are presented in Table 6. The comparison between the actual and predicted values of the PTA production plant is shown in Fig. 11. This experiment set contains 260 sets of PTA industrial data. The training and testing sets of the BP, RBF, ELM, K-means-ELM, and AP-ELM methods contain 178 and 9 sets, respectively.

It can be seen from Fig. 11 and Table 6 that the AP-ELM method is the best to predict the PTA yield. The outputs of PTA production plants based on the AP-ELM method are shown in Fig. 12.

Based on the PTA energy efficiency data, this study built the predictive model and predicted the PTA output. It can be seen from Table 6 that, compared with the K-means-ELM, ELM, BP, and RBF, the average error of the AP-ELM decreased by 0.07%, 0.16%, 0.21% and 0.29%, respectively. Moreover, from Fig. 12, it can be seen that the AP-ELM method can reach better performance in predicting the PTA output. The PTA inputs in the improved and original plants are shown in Fig. 13,

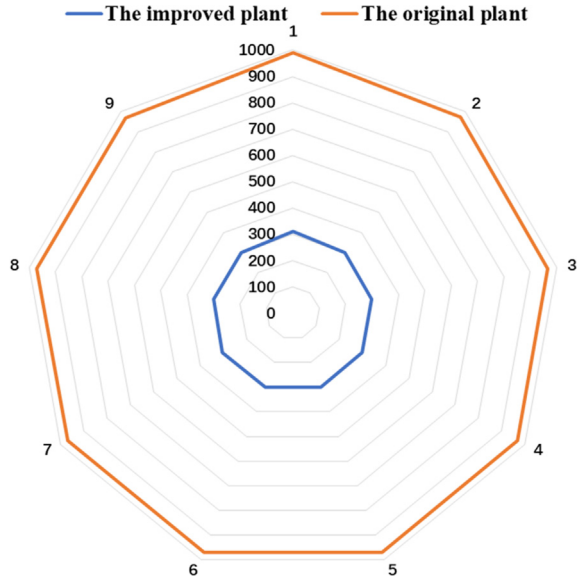

Fig. 13. Inputs of PTA in the improved and original plants.

which shows that the amounts of raw materials in the improved plant are much lower than those in the original plant. Taking sample 1 as example, the improved plant uses approximately 0.56 of material No.3, 32.93 of No.4, 12.22 of No.7, 108.50 of No.13, 97.22 of No.15, and 60.32 of No.17 to produce the PTA. However, the original plant uses 37.96 of material No.1, 19.60 of No.2, 0.56 of No.3, 32.93 of No.4, 1.73 of No.5, 24.08 of No.6, 12.22 of No.7, 67.12 of No.8, 35.99 of No.9, 84.31 of No.10, 97.66 of No.11, 99.20 of No.12, 108.50 of No.13, 96.19 of No.14, 97.22 of No.15, 115.29 of No.16, and 60.32 of No.17 to produce the PTA. The actual value of the output is 47.73, and the predicted value is 47.87. At the same time, the input of materials is reduced by 679. Based on the result analysis, the original plant can adjust the PTA inputs to improve the energy efficiency in the future.

6. Discussion

First, a novel AP-ELM method is proposed. The raw data are clustered by using the AP clustering algorithm to reduce multi-dimensional data and fuse high similarity data. Then, the clustering results are used as the input of the ELM. The accuracy and validity of the proposed method are verified by comparison with the BP, RBF, ELM, and K-means-ELM methods using UCI data sets.

Second, the proposed method is applied to build the production capacity analysis and energy optimization model of the ethylene and PTA production systems in complex petrochemical industries. The prediction accuracy in the ethylene and PTA production reaches approximately 99%. Moreover, the crude oil of the ethylene production plant is reduced by 121,938 ton and the material of the PTA production plant is reduced by 679. Furthermore, the CO_2 emission reduction of the ethylene production system is about 73,802 ton.

Third, the model also has some disadvantages. The clustering results are poor when the differences between the data are not significant. Therefore, some artificial intelligence methods, such as particle swarm optimization and deep learning, will be integrated to adaptively adjust

the clustering results. Moreover, the proposed model can be applied widely in energy saving and CO_2 emission reduction of other complex industries.

7. Conclusion

The novel AP-ELM method is proposed in this paper. The sample data can be classified automatically. Then, the multi-dimensional and high similarity data are reduced as the training and testing sets of the ELM. Finally, the production capacity analysis and energy optimization models of the ethylene and PTA production systems are built by using the proposed method. These experimental results show that the prediction accuracy of the ethylene and PTA production reaches approximately 99%, thus verifying the effectiveness and accuracy of the proposed method. Meanwhile, the proposed model can guide the production department to rationally allocate the inputs of raw materials, improve the production efficiency, save energy, and reduce the CO_2 emission. Moreover, the crude oil of the ethylene production plant is reduced by 121,938 ton and the raw materials in the PTA production plant are reduced by 679. Furthermore, the CO_2 emission reduction of the ethylene production system is approximately 73,802 ton.

In the further works, the pollutant emissions should be taken into consideration. Furthermore, Monte Carlo tools will be used to obtain a large amount of data and analyze the error variation of the AP-ELM method.

Acknowledgement

This work is partly financial supported by the National Key Research and Development Program of China (2017YFC1601800), the National Natural Science Foundation of China (61603025, 61673046 and 61533003), the Natural Science Foundation of Beijing, China (4162045) and the Fundamental Research Funds for the Central Universities (XK1802-4).

References

[1] Chen JM, Yu BY, Wei YM. Energy technology roadmap for ethylene industry in China. Appl Energy 2018;224:160–74.
[2] Yu RW, Ma GF, Xu YH. Review of Sinopec's ethylene production in 2015. Ethylene Indus 2016;28:1–6.
[3] Zhang LJ, Wang ZY, Xing YC. Review of Petrochina's ethylene production in 2015. Ethylene Indus 2016;28:7–11.
[4] Wieslaw G, Wojciech S. Influence of power source type on energy effectiveness and environmental impact of cooling system with adsorption refrigerator. Energy Convers Manage 2014;87:1107–15.
[5] 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 Manage 2016;124:73–83.
[6] Geng ZQ, Gao HC, Wang YQ, Han YM, Zhu QX. Energy saving analysis and management modeling based on index decomposition analysis integrated energy saving potential method: application to complex chemical processes. Energy Convers Manage 2017;145:41–52.
[7] Taeheung K, Byung Do C, Jong-Seok L. Incorporating receiver operating characteristics into naive Bayes for unbalanced data classification. Computing 2016;99:1–16.
[8] Yuan TH, Zhu N, Shi YF. Sample data selection method for improving the prediction accuracy of the 3 heating energy consumption. Energy Build 2018;158:234–43.
[9] Yin LB, Liu GC, Zhou JL. A calculation method for CO_2 emission in utility boilers based on BP neural network and carbon balance. Energy Proc 2017;105:3173–8.
[10] Wang Z, Wang F, Su S. Solar irradiance short-term prediction model based on BP neural network. Energy Proc 2011;12:488–94.
[11] Badar I, Zuhairi B, Perumal N. Development of chaotically improved meta-heuristics and modified BP neural network-based model for electrical energy demand prediction in smart grid. Neural Comput Appl 2007;28:877–91.
[12] Wen L, Liu YJ. A research about Beijing's carbon emissions based on the IPSO-BP model. Environ Prog Sustain Energy 2017;3(6):42 8–434.
[13] Hejase HAN, Al-Shamisi MH, Assi AH. Modeling of global horizontal irradiance in the United Arab Emirates with artificial neural networks. Energy 2014;77:542–52.
[14] Rezaei J, Shahbakhti M, Bahri B. Performance prediction of HCCI engines with oxygenated fuels using artificial neural networks. Appl Energy 2015;138:460–73.
[15] Zhang H, He YH, Tang D. In-cell selective ultrasonic electrosynthesis of methyl benzaldehyde based on RBF neural network and genetic algorithm. Chem J Chin Univ-Chin 2014;35:1199–203.
[16] Liu QQ, Wang CY, Shi XF. Identification of spill oil species based on low concentration synchronous fluorescence spectra and RBF neural network. Spectrosc Spect Anal 2012;32:1012–5.
[17] Zhang JF, Hu SS. Chaotic time series prediction based on RBF neural networks with a new clustering algorithm. Acta Phys Sin 2007;56:713–9.
[18] Guo HJ, Liu JH. Chaos control of Lorenz system via RBF neural network sliding mode controller. Acta Phys Sin 2004;53:4080–6.
[19] Geng ZQ, Chen J, Han YM. Energy efficiency prediction based on PCA-FRBF model: a case study of ethylene industries. IEEE Trans Syst Man Cybernet-Syst 2017;47:1763–73.
[20] Huang GB, Zhu QY, Siew CK. Extreme learning machine: a new learning scheme of feedforward neural networks. IEEE international joint conference on neural networks, 2004. proceedings, vol. 2. IEEE; 2005. p. 985–90.
[21] Geng ZQ, Lin Q, Han YM, Zhu QX. Energy saving and prediction modeling of petrochemical industries: a novel ELM based on FAHP. Energy 2017;12:350–62.
[22] Chen HY, Gao PZ, Tan SC. Online sequential condition prediction method of natural circulation systems based on EOS-ELM and phase space reconstruction. Ann Nucl Energy 2017;110:1107–20.
[23] Kasra M, Shahaboddin S, Lip YP. Predicting the wind power density based upon extreme learning machine. Energy 2015;86:232–9.
[24] Jin WY, Liu Y, Gao ZL. Fast property prediction in an industrial rubber mixing process with local ELM model. J Appl Polym Sci 2017;134(41).
[25] Han YM, Zeng Q, Geng ZQ, Zhu QX. Energy management and optimization modeling based on a novel fuzzy extreme learning machine: case study of complex petrochemical industries. Energy Convers Manage 2018;145:41–52.
[26] Lan Y, Soh YC, Huang GB. A constructive enhancement for online sequential extreme learning machine. International joint conference on neural networks. IEEE Press; 2009. p. 1708–13.
[27] Khalil B, Ali C. Forecasting hourly global solar radiation using hybrid k-means and nonlinear autoregressive neural network models. Energy Convers Manage 2013;75:561–9.
[28] Ali Fallah T, Diane A. Modified sequential k-means clustering by utilizing response: a case study for fashion products. Exp Syst 2017;34.
[29] Mahesh K K, Rama Mohan R A. An efficient k-means clustering filtering algorithm using density based initial cluster centers. Inform Sci 2017.
[30] Li PX, Gu W, Wang LF. Int J Elect Power Energy Syst 2018.
[31] Chen L, Xiao CB, Yu J. Fault detection based on AP clustering and PCA. Int J Pattern Recognit Artif Intell 2018;32.
[32] Liu JJ, Kan JQ. Recognition of genetically modified product based on affinity propagation clustering and terahertz spectroscopy. Mol Biomol Spectrosc 2018;194:14–20.
[33] Carlos Eduardo G, Nayara Safira DSC, Andre Jose NA. Lithology mapping by a hybridization of the firefly and affinity propagation algorithms. J Petrol Sci Eng 2017.
[34] Hang WL, Chung FL, Wang ST. Transfer affinity propagation-based clustering. Inform Sci 2016;348:337–56.
[35] Wang Y, Cheng L, Zhang J. Affinity propagation algorithm based multi-source localization method for binary detection. IEICE Trans Inform Syst 2017.
[36] Zhao JL, Qu H, Zhao JH. Towards controller placement problem for software-defined network using affinity propagation. Electron Lett 2017;53:928–9.
[37] Brendan JF, Delbert D. Clustering by passing messages between data points. Science 2007;315:972–6.
[38] Xia CM, Ni ZW. Affinity propagation clustering algorithm based on density adjustment and manifold distan. Comput Sci 2017;10:187–92.
[39] Yang FW, Zeng ZG, Liu Q, Liu LH. Research and application of image segmentation based on AP-clustering algorithm. College Comput Commun 2007.
[40] Chen WB, Song MJ. A data visualization method based on extreme learning machine. Comput Eng Sci 2017;05:912–8.
[41] Huang GB, Zhu QY. Extreme leaning machine: a new leaning scheme of feedforward neural networks. Proceeding of International joint conference neural networks. 2004. p. 985–90.
[42] Siew CK, Huang GB, Zhu QY. Extreme leaning machine: theory and applications. Neurocomputing 2006;70:489–501.
[43] Qiu RH, Liu KL, Tan HL. Classification algorithm based on extreme learning machine and its application in fault identification of Tennessee Eastman process. Eng Sci 2016;50.
[44] Abdelkarim BA, Mohamed BH, Adel MA. Cluster forests based fuzzy c-means for data clustering. International conference on computational intelligence in security for information systems. 2016.
[45] Han YM, Geng ZQ, Gu XB, Zhu QX. Energy efficiency analysis based on DEA integrated ISM: a case study for Chinese ethylene industries. Eng Appl Artif Intell 2015;45:80–9.
[46] Zhao M, Zhang WG, Yu LZ. Carbon emissions from energy consumption in Shanghai city. Res Environ Sci 2009;22:984–9.
[47] He YL, Geng ZQ, Zhu QX. Positive and negative correlation input attributes oriented subnets based double parallel extreme learning machine (PNIAOS-DPELM) and its application to monitoring chemical processes in steady state. Neurocomputing 2015;165:171–81.