Energy Conversion and Management 165 (2018) 163–171

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Energy management and optimization modeling based on a novel fuzzy extreme learning machine: Case study of complex petrochemical industries

Yongming Han, Qing Zeng, Zhiqiang Geng*, Qunxiong Zhu*

College of Information Science & Technology, Beijing University of Chemical Technology, Beijing 100029, China
Engineering Research Center of Intelligent PSE, Ministry of Education in China, Beijing 100029, China

A R T I C L E I N F O

Keywords:
Energy management
Energy optimization
Production prediction
Data fuzzification
Fuzzy extreme learning machine
Complex petrochemical industries

A B S T R A C T

Energy management and optimization play a key effect in the sustainable development. However, the uncertain data has a direct impact on the production prediction and energy optimization of complex petrochemical industries. Therefore, this paper proposes a novel energy management and optimization model based on the fuzzy extreme learning machine (FELM) method integrated the fuzzy set theory. The minimal, the median and the maximal values of the energy consumption data are obtained by data fuzzification to solve the problem of the fluctuation and uncertainty data. And the cross recombination of triangular fuzzy numbers (TFNs) is applied in the training of the FELM. Moreover, the upper and the lower limits of efficiency values are obtained on the basis of the network generalization to analyze the energy conservation and saving potentials. Furthermore, the FELM has better predictive performance and training speed than fuzzy error back propagation network (FBP) and fuzzy radical basis function network (FRBF) though University of California Irvine (UCI) standard datasets. Finally, the proposed method is applied to manage and optimize the energy status of China ethylene industry in complex petrochemical industries. The experimental results show that the proposed method is effective and applicable in the energy-saving potential, which is indicated up to about 15%.

1. Introduction

The sustainable development strategy has played an important guiding role in building a well-off society in an all-round way. However, China is a country with a large population, the shortage of natural resources and a relatively backward condition in economy and technology. Therefore, the social and economic virtuous circle can be achieved through the resource rational utilization and the environmental protection, especially complex petrochemical industries. The ethylene industry is the most crucial part of the complex petrochemical industries, in which both environmental and economic goals can be achieved by production prediction and energy management. However, in 2015, the ethylene production and the average fuel and power consumption of China Petrochemical Corporation was 11005.2 kt/a and 559.06 kg per ton of ethylene [1], respectively. And the ethylene production and the average fuel and power consumption of China National Petroleum Corporation in 2015 was 5032 kt/a and 594 kg per ton of ethylene [2], respectively. The energy efficiency of the ethylene production in China is far lower than international countries. Moreover, over 50% of operating costs in ethylene plants come from the cost of energy consumption of ethylene plants [3]. Therefore, energy management and optimization of the complex petrochemical industry can bring direct economic and environmental benefit.

In practical efficiency evaluation, there are two commonly used techniques named the mean method and the optimal index method [4]. However, these approaches can barely provide the energy efficiency benchmark of improvement factors, thus brings obstacles to the optimization decision. Meanwhile, some other drawbacks also reduce the usefulness of methods, such as the lack of predictive ability and the biased analysis caused by imprecise data.

In order to address these above issues, this paper proposes a novel fuzzy extreme learning machine (FELM) method integrated the fuzzy set theory to build the energy management and optimization model. The minimal, the median and the maximal values of the energy consumption data are obtained by data fuzzification to process the imprecise data. And then the triangular fuzzy numbers (TFNs) are integrated together with the FELM method to get the predictive production capacity and the efficiency evaluation. Though the University of California Irvine (UCI) standard datasets, the proposed method has advantages of higher forecasting accuracy, faster training speed than fuzzy error back propagation network (FBP) and fuzzy radical basis function network (FRBF), and can avoid from falling into the

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

https://doi.org/10.1016/j.enconman.2018.03.049
Received 2 January 2018; Received in revised form 9 March 2018; Accepted 17 March 2018
0196-8904/

local optimum. Meanwhile, with the input and output data fuzzified in triangular form, the robustness of evaluation method can be enhanced to overcome the noise and fluctuation in the crisp data. Moreover, the experimental results in managing and optimizing the energy status of China ethylene industries show that the proposed method is effective and applicable in the energy-saving potential, which is indicated up to about 15%.

The remainder of this paper has been arranged as follows. The development of energy prediction and management based on the fuzzy-based and ANN-based methods are introduced in Section 2. Section 3 describes the FELM analysis framework in detail. The validation test of the FELM based on the UCI standard dataset is laid out in Section 4. In Section 5, we present a case study about energy prediction and optimization of the ethylene production system based on the FELM. Finally, Section 6 gives the discussion about the experiment and Section 7 provides the conclusion.

2. Related work

Energy prediction and efficiency evaluation plays an important role in energy-saving and emission reduction. In the previous studies, different indexes and methods was proposed to analyze the energy efficiency and optimize the energy configuration of complex chemistry industries. The index decomposition analysis (IDA) is a typical method to manage the energy performance. Geng et al. obtained the activity, the structure and the intensity of three energy performance indicators (EPIs) that affect the energy consumption based on the IDA method [5]. Zhou et al. discussed and analyzed the macro energy efficiency evolution method based on the IDA [6]. However, the inherent shortcoming of not taking the energy-saving knowledge into account makes it difficult for decision-makers to optimize the energy usage.

Meanwhile, the stochastic frontier analysis (SFA) is a widely used method for estimating the efficiency by using stochastic frontier production functions. Jones et al. focused on the efficiency of electricity consumption in the Portuguese productive sector using the SFA [7]. Lin et al. adopt the SFA to study the average energy efficiency and energy-saving potentials of the Chinese chemical industry based on the assumption of the trans-log production function [8]. However, because the effects of stochastic factors are taken into consideration, the parametric frontier approach of the SFA may cause the disagreement among distributional assumptions and objective circumstances.

Compared with the SFA, data envelopment analysis (DEA) is a nonparametric frontier efficiency analysis method for multi-input and multi-output systems. The DEA was applied widely in the areas of energy and environmental analyses. Sueyoshi et al. proposed an improved DEA to assess the performance of energy industries [9]. Han et al. proposed an energy efficiency evaluation method based on the Malmquist production indexes integrated the improved DEA cross model [10]. However, the efficiency discrimination of DEA will be poor when more than a third of efficiency values are set to 1 [11]. Meanwhile, the uncertain and noisy characteristic of the statistical data may result in the erroneous evaluation [12]. Therefore, with the imprecise energy consumption data, the DEA is quite sensitive to outliers and just can handle crisp input and output data.

For solving problems exposed by the above methods, this paper proposes a novel fuzzy ANN method based on the fuzzy set theory. Zadeh proposed the fuzzy set theory and described fuzziness in precise mathematical languages [13]. Coppi et al. compared two kinds of clustering models for the left and right fuzzy data by using the empirical information affected by imprecision or vagueness [14]. Xiong et al. explored the application of the fuzzy set in association rules mining [15]. Chen et al. extended previous fuzzy mining approaches to handle the time-series data and the additional experiment showed favorable results [16]. In the energy efficiency evaluation of industry sectors, Kao et al. analyzed the fuzzy input and the output to the involved independent processes for parallel production systems [17]. Han et al. proposed an efficiency analysis method based on fuzzy DEA cross model with the fuzzy data for ethylene production systems in complex chemical industries [18]. Therefore, the fuzzy set theory is effective in predicting and optimizing the energy status of complex petrochemical industries.

The ANN is an important branch of artificial intelligence, with self-adaptive, self-organizing and self-learning characteristics. Meanwhile, the ANN is a dominant data-driven model in the machine learning area and has been widely applied in modeling and predicting. The BP neural network proposed by Rumelhart and McClelland is one of the most widely used methods currently [19]. Han et al. proposed an improved DEA integrating the BP neural network method for analyze the energy efficiency of complex petrochemical industries and get satisfactory results [20]. Other applications of the ANN covered a wide range of fields including the petrochemical industries [21,22], renewable energy field [23] and pharmaceutical industries [24]. The ELM is a single-hidden layer feedforward neural network proposed by Huang et al. [25]. With respect to the classical BP neural network, the ELM is more popular for enhancing capabilities in terms of generalization and training speed. At the same time, it can avoid the local minimization problem in traditional models. Geng et al. proposed an energy optimization and analysis modeling based on the ELM focusing on complex chemical processes [26]. Naji et al. used the ELM to estimate building energy consumption [27]. In addition, the ELM was also found a wide utilization in some comprehensive problems like classification in food industry [28], modeling for the batch process [29] and wind speed forecast [30], engine control [31] and nonlinear process optimization [32]. Extensive experiments on a standard UCI dataset show that the ELM is a simple neural network learning algorithm with high training speed and generalization accuracy. However, the inevitable error and the noise bring fluctuations to complex industrial data. The direct use of the ELM to predict the products and analyze the energy status of the petrochemical industry may lead to the poor modeling performance and the low forecast accuracy. Based on the above analysis, the advantages and disadvantages among various methods are shown in Table 1. Therefore, this paper introduces the FELM by using the TFNs to improve the fault tolerance of this method. Meanwhile, the FELM method is applied to manage and optimize the energy status of China ethylene industries.

Table 1
Advantages and disadvantages among various methods.

Method	Advantages	Disadvantages
IDA	A typical and simple method	Deep factors of energy intensity change are not taken into account
SFA	The effects of stochastic factors on outputs are considered	The disagreement among distributional assumptions and objective circumstances
DEA	A nonparametric frontier efficiency analysis method for multi-input and multi-output systems	Quite sensitive to outliers, and just can handle crisp input and output data
BP	A dominant data-driven method and fast modeling	May lead to the local minimization problem
ELM	Enhancing capabilities in terms of generalization and training speed, and can avoid the local minimization problem better	May lead to the poor modeling performance and the low forecast accuracy

3. FELM based on the fuzzy set theory

3.1. Fuzzy numbers and its operations

Precise rules are required when mathematical methods are used to handle the problems with realistic background, yet many descriptions or concepts in the real world are often difficult to be defined clearly. The fuzzy set theory excels in this aspect and fuzzy numbers can bridge the gap between rigorous mathematical descriptions and things with obscure boundary. A series of steps with the strict theoretical support help the method consider vagueness or ambiguity in decision-making models, hence fuzzy set theory is widely applied in various fields.

3.1.1. Triangular fuzzy number

Triangular and trapezoidal fuzzy numbers are two kinds of frequently-used shapes in fuzzification. For reasons that the 1-cut set of TFN includes only one point and membership functions on both sides are linear, the operations between TFNs are very convenient. Meanwhile, parameters of the TFN can represent the minimal, the median and the maximal levels of data samples. Accordingly, we use the TFN as a kind of data fusion method to deal with samples instead of using the crisp data directly.

Suppose that $A=(x_D,x_M,x_U),(x_D\leqslant x_M\leqslant x_U)$ is a TFN on R, then the membership function $\mu_A(x)$: $R\to[0,1]$ can be obtained as the following Eq. (1):

$$\mu_A(x)=\begin{cases}(x-x_D)/(x_M-x_D),x_D\leqslant x\leqslant x_M\\(x_U-x)/(x_U-x_M),x_M\leqslant x\leqslant x_U\\0,\text{otherwise}\end{cases}\tag{1}$$

Let $A_1=(x_{D1},x_{M1},x_{U1}),(x_{D1}\leqslant x_{M1}\leqslant x_{U1})$ and $A_2=(x_{D2},x_{M2},x_{U2}),(x_{D2}\leqslant x_{M2}\leqslant x_{U2})$ be any two TFNs, the operation laws are as follows:

$$A_1(\pm)A_2=(x_{D1},x_{M1},x_{U1})(\pm)(x_{D2},x_{M2},x_{U2})=(x_{D1}\pm x_{D2},x_{M1}\pm x_{M2},x_{U1}\pm x_{U2})\tag{2}$$

$$A_1(\cdot)A_2=(x_{D1},x_{M1},x_{U1})(\cdot)(x_{D2},x_{M2},x_{U2})=(x_{D1}x_{D2},x_{M1}x_{M2},x_{U1}x_{U2})\tag{3}$$

$$A_1(/)A_2=(x_{D1},x_{M1},x_{U1})(/)(x_{D2},x_{M2},x_{U2})=(x_{D1}/x_{U2},x_{M1}/x_{M2},x_{U1}/x_{D2})\tag{4}$$

$$A_1^{-1}=(x_{D1},x_{M1},x_{U1})^{-1}=(1/x_{U1},1/x_{M1},1/x_{D1})\tag{5}$$

$$k(\cdot)A_1=k(\cdot)(x_{D1},x_{M1},x_{U1})=\begin{cases}(kx_{D1},kx_{M1},kx_{U1}),k\geqslant 0\\(kx_{U1},kx_{M1},kx_{D1}),k<0\end{cases},\ k\text{ is constant}\tag{6}$$

3.1.2. Satisfaction with fuzzy numbers

Definition 1. For the supposed two TFNs $A_1=(x_{D1},x_{M1},x_{U1})$ and $A_2=(x_{D2},x_{M2},x_{U2})$, the intersections of the membership function can be calculated by the following equation:

$$\alpha_*=\begin{cases}1-\frac{x_{M2}-x_{M1}}{x_{U1}+x_{D2}},0\leqslant x_{M2}-x_{M1}\leqslant x_{U1}+x_{D2}\\1-\frac{x_{M1}-x_{M2}}{x_{D1}+x_{U2}},0\leqslant x_{M1}-x_{M2}\leqslant x_{D1}+x_{U2}\\0,\text{others}\end{cases}\tag{7}$$

If two TFNs intersect with each other, certainly they have only one intersection point and the point is the center of two TFNs. Eq. (7) can reflect the relationship between two TFNs. On the basis of intersection membership values, the satisfaction formula can be used in pairwise comparison of the fuzzy numbers.

Definition 2. The satisfaction P of TFNs in comparative relationship $A_1\leqslant A_2$ is defined as follows:

$$P_n(A_1\leqslant A_2)=\begin{cases}1-\alpha_*^n\frac{x_{U1}+x_{D2}}{x_{D1}+x_{U1}+x_{D2}+x_{U2}},A_1\leqslant A_2\\\alpha_*^n\frac{x_{U2}+x_{D1}}{x_{D1}+x_{U1}+x_{D2}+x_{U2}},A_1\geqslant A_2\end{cases}\tag{8}$$

For any two TFNs A_1 and A_2, there are some other noteworthy properties showed as follows:

(1) $0<P_n(A_1\leqslant A_2)\leqslant 1$
(2) When $A_1=A_2$, then $P_n(A_1\leqslant A_2)=1-((x_{U1}+x_{D2})/(x_{D1}+x_{U1}+x_{D2}+x_{U2}))$
(3) When $x_{M1}+x_{U1}\leqslant x_{M2}-x_{D2}$,$P_n(A_1\leqslant A_2)=1$
(4) $P_n(A_1\leqslant A_2)\leqslant 1-P_n(A_2\leqslant A_1)$

The parameter n in the calculation process is constant. The distinguish ability between TFNs is improved with n increases, but not the larger value the better. The practical problems should be taken into account to get a reasonable computation amount. In some applications it was generally considered as $n=3$ [33].

3.2. FELM with data fuzzification

3.2.1. Data fuzzification

Assume a data set with m samples: $x_1,x_2,x_3,...x_m$. Divide them into groups with certain length, and the interval length of each group is l. The group can be represented as $x_i,x_{i+1},x_{i+2},...,x_{i+l-1}$. The minimal value D, the maximal value U, and the median value M of the group are defined as follows for consequent calculations:

$$D=\min\{x_i,x_{i+1},...,x_{i+l-1}\}U=\max\{x_i,x_{i+1},...,x_{i+l-1}\}M=\frac{(x_i+x_{i+1}+\cdots+x_{i+l-1})}{l}\tag{9}$$

Two parameters D' and U' are needed to acquire TFNs:

$$D'=\begin{cases}M-2(M-D),M\geqslant 2(M-D)\\0,M\leqslant 2(M-D)\end{cases},\quad U'=M+2(U-M)\tag{10}$$

By definitions, the Gaussian membership function of TFNs for m data in each group can be calculated as:

$$\mu_D=\begin{cases}1,x\leqslant D'\\e^{-\frac{1}{2}\left(\frac{\frac{x-D'}{U'-D'}}{\sigma}\right)^2},D'<x<\frac{D'+U'}{2}\\0,x\geqslant\frac{D'+U'}{2}\end{cases}\tag{11}$$

$$\mu_M=\begin{cases}0,x\leqslant D'\\e^{-\frac{1}{2}\left(\frac{x-\frac{D'+U'}{2}}{\frac{U'-D'}{\sigma}}\right)^2},D'<x<U'\\0,x\geqslant U'\end{cases}\tag{12}$$

$$\mu_U=\begin{cases}0,x\leqslant\frac{D'+U'}{2}\\e^{-\frac{1}{2}\left(\frac{\frac{x-U'}{U'-D'}}{\sigma}\right)^2},\frac{D'+U'}{2}<x<U'\\0,x\geqslant U'\end{cases}\tag{13}$$

The l data, $x_i,x_{i+1},x_{i+2},...,x_{i+l-1}$, in each group can be substituted into Eqs. (11)–(13) to get μ_D, μ_M and μ_U. The maximal values of μ_D, μ_M and μ_U in this group correspond to a set consists of x_D, x_M and x_U. This set is the expected TFN: (x_D,x_M,x_U).

A data set with n samples can get k TFNs: $(x_{D_1},x_{M_1},x_{U_1})$, $(x_{D_2},x_{M_2},x_{U_2})$,…, $(x_{D_k},x_{M_k},x_{U_k})$, where $k=m/l$.

3.2.2. The FELM method

According to the previous algorithm description, the input and the output data are disposed respectively to generate TFNs as a data fusion process. Then TFNs can be used to build a FELM-based model that specific to deal with statistic noise and other imprecise information in the complex chemical process.

The judgment criterion for whether production configurations are

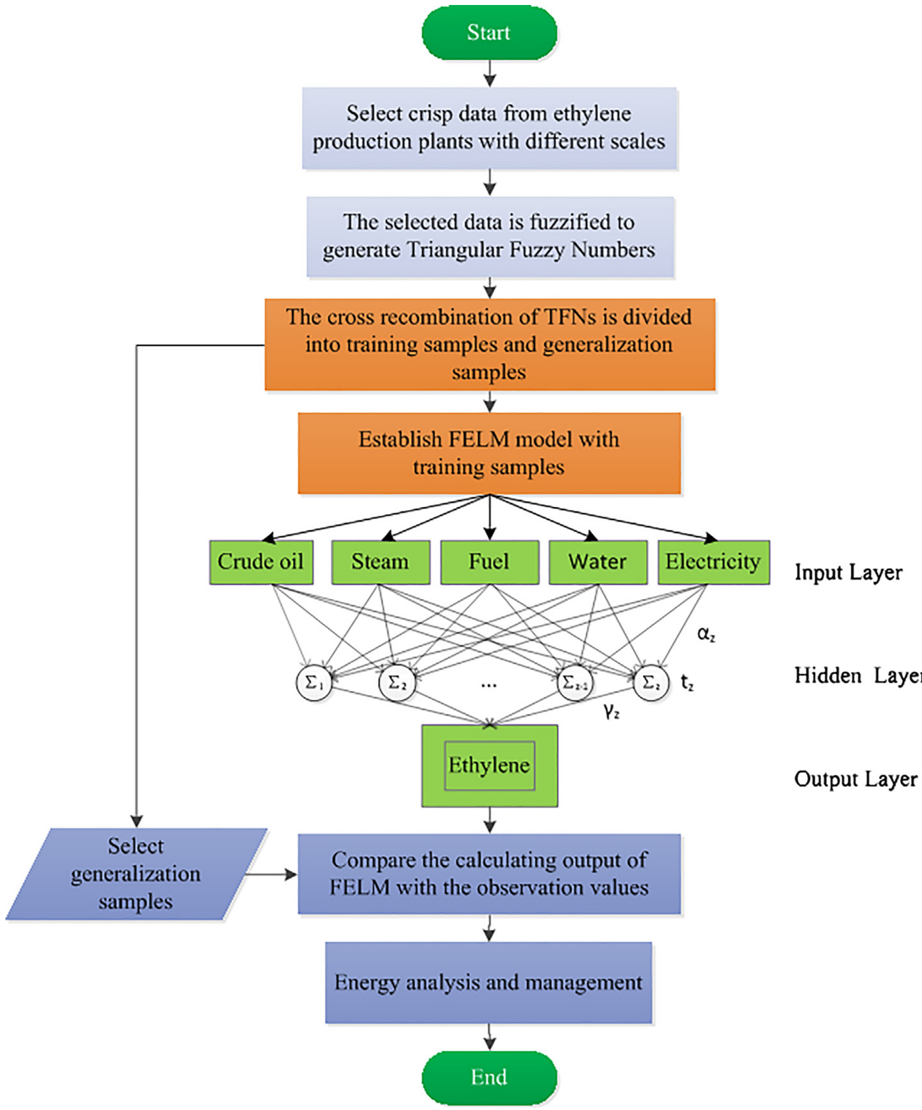


Fig. 1. The analysis and prediction framework based on the FELM.

efficient is ambiguous. As a matter of fact, an intuitive concept shows that when the output is larger and the input is fewer, the production is under a high efficiency condition. After data fuzzification, the input data and the output data are shown in Eq. (14).

$$(x_i)^t = ((x_i)^D,(x_i)^M,(x_i)^U),(y_i)^t = ((y_i)^D,(y_i)^M,(y_i)^U)$$
$$(i = 1,2,...,k;t = D,M,U;D \leqslant M \leqslant U) \tag{14}$$

Suppose that the training set includes m samples, the interval length is chosen as l, then the number of TFNs is k, $k = m/l$.

When plant produces with the minimal inputs $(x_i)^D$ and the maximal outputs $(y_i)^U$, the production is under the optimal configuration. When plant converts the maximal inputs $(x_i)^U$ into the minimal outputs $(y_i)^D$, it is a bad production point. And the daily operations transform the median inputs $(x_i)^M$ into the median outputs $(y_i)^M$. In order to express the above three situations, the obtained TFNs are substituted into the ELM to construct data-driven models that are on behalf of three allocations.

$$\begin{aligned} K_1 &= \{((x_i)^D,(y_i)^U)|i = 1,2,...,k;(x_i)^D \in R^J;(y_i)^U \in R^O\}K_2 \\ &= \{((x_i)^M,(y_i)^M)|i = 1,2,...,k;(x_i)^M \in R^J;(y_i)^M \in R^O\}K_3 \\ &= \{((x_i)^U,(y_i)^D)|i = 1,2,...,k;(x_i)^U \in R^J;(y_i)^D \in R^O\} \end{aligned} \tag{15}$$

The fuzzified data is recombined in the cross style. In each training sample, there are J attributes of the input data x_i and O attributes of the output data y_i.

$$\begin{aligned} (y_i)^U &= f_1((x_i)^D) = \sum_{z_1=1}^{Z_1} \gamma_{z1}^1 g_1(\alpha_{z1}^1\cdot(x_i)^D + t_{z1}^1)(y_i)^M = f_2((x_i)^M) \\ &= \sum_{z_2=1}^{Z_2} \gamma_{z2}^2 g_2(\alpha_{z2}^2\cdot(x_i)^M + t_{z2}^2)(y_i)^D = f_3((x_i)^U) \\ &= \sum_{z_3=1}^{Z_3} \gamma_{z3}^3 g_3(\alpha_{z3}^3\cdot(x_i)^U + t_{z3}^3) \end{aligned} \tag{16}$$

These models are derived from the classical ELM model. As the input and the output come from different locations of the data distribution, we distinguish them by the use of subscripts and superscripts 1,2,3. $g(\cdot)$ is the activation function. $\alpha_z = [\alpha_{z1},\alpha_{z2},...,\alpha_{zM}]$ are the weights between the m input nodes and the z^{th} hidden node. $\gamma_z = [\gamma_{z1},\gamma_{z2},...,\gamma_{zN}]$ are the weights from the z^{th} hidden node to the n output nodes. Variable t_z is the bias item of the z^{th} hidden node. Theoretical analysis of the ELM proves that the appropriate values of γ_z, α_z and t_z can make the output of the network approach the expectation with the error of zero. Take the first equation in Eq. (16) for instance, it can be reshaped as:

$$H_D\gamma^1 = Y_U \tag{17}$$

In which, H_D is the hidden layer output:

Table 2
UCI data sets.

Data Sets	Samples		Attributes	
	Training	Testing	Inputs	Outputs
Combined cycle power plant	700	256	3	1
Airfoil self-noise	70	36	5	1

$$H_D = \begin{bmatrix} g_1(\alpha_1^1 \cdot (x_1)^D + t_1^1) & \cdots & g_1(\alpha_{z_1}^1 \cdot (x_1)^D + t_{z_1}^1) \\ \vdots & \vdots & \vdots \\ g_1(\alpha_1^1 \cdot (x_k)^D + t_1^1) & \cdots & g_1(\alpha_{z_1}^1 \cdot (x_k)^D + t_{z_1}^1) \end{bmatrix}_{k \times Z_1} \tag{18}$$

In Eqs. (17)–(18), the connection weight matrix α and the threshold vector t are generated randomly. The unique unknown parameter of the model is the connection weight matrix γ, and we can calculate it as:

$$\hat{\gamma}^1 = H_D^+ Y_U \tag{19}$$

In which H_D^+ is the Moore-Penrose generalized inverse of H_D. The parameters of the other two models can be obtained by the same way.

After training with data sets K_1, K_2 and K_3 in accordance with learning algorithm, some parameters need to be recorded for generalization, which include the input weight matrices α^1, α^2, α^3, the output weight matrices γ^1, γ^2, γ^3, and the threshold vectors of the hidden layer t^1, t^2, t^3.

A set of samples different from what we used in the training process is applied to generalization. The crisp data in the new set is also fuzzified to acquire TFNs in the same way. The generalization results of the trained FELM are compared with the expected outputs to calculate their respective relative errors as performance indices.

3.2.3. The analysis and prediction framework of the FELM

As is shown in the flowchart in Fig. 1, the framework of energy prediction and optimization based on the FELM method can be established as follows:

Step 1: Collecting available data from the chemical processes. The typical data should reflect the relevant information about production status. The required data is divided in a proper way for training and generalization. The crisp data is fuzzified to generate new samples, and then they can be used in the ELM. Two sets for training and generalization have the same attributes, whereas their difference lies in diverse time periods. It should be noted that the generalization set covers the time periods that to be analyzed.

Step 2: Generating triangular fuzzy numbers. The data selected in previous step is divided into groups according to the fixed time period interval, and each group is fused to obtain a TFN. The TFN contains three parameters, which respectively represent the minimal, the median and the maximal levels of the group. The fuzzification process reduces the impact of fluctuations in data. Meanwhile, the number of samples can be decreased. The unchanged attributes will be regarded as indicators to analyze the energy efficiency.

Step 3: Training of the ELM. During the training process, three models are established simultaneously. The model which includes the maximal values of inputs and the minimal values of outputs is regarded as the most inefficient production situation. The model which involves the median values of inputs and the median values of outputs represents the median production situation. The model constructed with the minimal values of inputs and the maximal values of outputs is the benchmark of the most efficient production situation, and it indicates the improvement direction for future production. These ELM models can simulate the actual production situation reflected by the data from complex chemical plants.

Step 4: Energy analysis and optimization for the complex chemical process. The results of generalization are compared with the expected outputs to verify the prediction accuracy and get ready for further analysis. The upper and the lower limits of the output are identified by generalization curves, together with the predictive daily production. The interval belt of curve variation reflects the efficiency changing among different plants. In dealing with the optimization problem, the distance between the upper and the median curve is considered an efficiency indicator, which is used to conduct concrete calculation for energy conservation.

4. The validation test of the FELM based on UCI standard datasets

In order to verify the feasibility and the effectiveness of the FELM, the Airfoil Self-Noise and the Combined Cycle Power Plant data sets in the UCI database (More detailed information is provided in the website: http://archive.ics.uci.edu/ml /datasets.html) are selected to implement the verification. The description of attributes and the partition of samples are shown in Table 2.

The original number of samples for the Combined Cycle Power Plant data set is 9568, and the Airfoil Self-Noise data set is 1503. Based on data fuzzification, the number of samples in the Combined Cycle Power Plant and the Airfoil Self-Noise data sets are reduced to 956 and 106, respectively. And then we use the FELM, the FBP and the FRBF to carry on the data modeling with the recombined TFNs. Through the prediction performance and the convergence speed of them, for the FBP network, the nodes number of hidden layer is 30, the learning rate is 0.01 and the goal of training error is less than 0.001. The hidden-node number of FELM is set as 30. While the parameter of the FRBF is special and it depends on the smooth degree of data. For this reason, the spread speed of the Combined Cycle Power Plant data is set as 0.5, which of the Airfoil Self-Noise data is 9. The prediction performance of networks is evaluated by the average relative generalization error (ARGE) and the convergence speed is compared by the training time. The test results of the three neural networks for two UCI standard data sets are recorded in Tables 3 and 4.

From Tables 3 and 4, the FELM with data fuzzification provides the best performance with relatively simple structure (The hidden-node number of the FELM and the FBP are the same. This parameter of the FRBF is equal to the number of samples when the spread speed is determined, which is much larger than the ones of the previous two.) and faster training speed, which reveals the feasibility of the FELM.

5. Case study: energy management and optimization modeling of the complex petrochemical industry based on the FELM

In order to verify the actual value of the proposed framework, the FELM is applied to build the energy management and optimization model of the ethylene industry. The output of ethylene plants can be predicted, and then the energy utilization status as well as the merits and demerits of the production allocation can be analyzed.

5.1. Data analysis

The ethylene industry is a typical representative of the complex process industries. The ethylene production contains two main parts: cracking and separation. The production and operation of the cracking

Table 3
Test results of the Combined Cycle Power Plant standard data set.

	Median inputs to median outputs		Maximal inputs to minimal outputs		Minimal inputs to maximal outputs	
	ARGE	Training time	ARGE	Training time	ARGE	Training time
FELM	0.0146	0.0213	0.0064	0.0163	0.0068	0.0178
FBP	0.0152	0.9683	0.0065	0.9323	0.0067	0.8426
FRBF	0.0164	0.7819	0.0090	0.7901	0.0076	0.8101

Table 4
Test results of the Airfoil Self-Noise standard data set.

	Median inputs to median outputs		Maximal inputs to minimal outputs		Minimal inputs to maximal outputs	
	ARGE	Training time	ARGE	Training time	ARGE	Training time
FELM	0.0187	0.0199	0.0196	0.0206	0.0346	0.0202
FBP	0.0323	0.7666	0.0251	0.7591	0.0396	0.7300
FRBF	0.0434	0.6774	0.0310	0.6559	0.0369	0.6358

furnace consumes a lot of fuel to provide heat in the tube of the cracking reactions, which results in huge energy usage in ethylene production together with other high-energy-consumption taches. Meanwhile, the energy consumption in the preheat of the mixture of the raw materials and the stream and the waste heat released to the environment like afterheat in the flue gas are also high. And the separation section can be described as a three-step process: a rapid cooling, compression and a separation part [34], which have the main energy consumption including the reboiler consumption in the distillation column, the power consumption in the compressor, the cooling energy consumption in the compressor and the cold box, and the steam consumption. A typical schematic flow diagram of an ethylene plant is shown in Fig. 2.

In practice, many different battery limits and computational methods have been applied to analyze the energy efficiency [35]. Referring to the standard GB/T 2589-2008 and DB 37/751-2007 [36,37], we calculate the relevant energy consumption indices and divide the ethylene plants into different techniques and scales.

Known from the mechanism of the ethylene production industry, the raw material, the consumption of fuel, water, steam and electricity and the yield are the main factors which affect the energy efficiency. As the raw material, crude oil is made up of many components which include C_{345}, hydrogenation tail oil, light diesel fuel, naphtha and others.

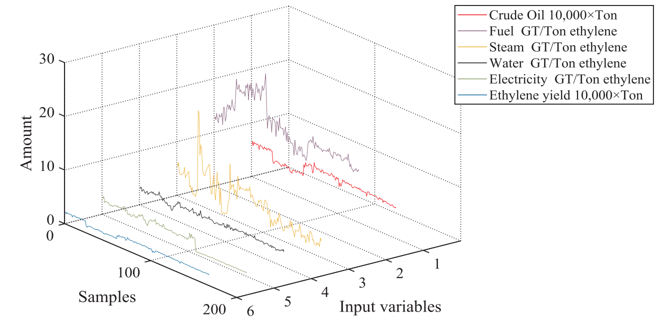


Fig. 3. The required crisp data from one ethylene plant.

The water includes circulating water, industrial water, boiler water and other water. The fuel consists of light weight oil, heavy weight oil and fuel gas. The steam is divided into super-high-pressure steam, high-pressure steam, medium-pressure steam and low-pressure steam. In this paper, the yield of ethylene is set as the output of the FELM. The amount of fuel, steam, water, electricity and crude oil are regarded as the inputs of the FELM. For convenience of calculation, the measure units of energy consumption indicators like fuel, steam, water and electricity are converted to GJ/Ton ethylene [38], while crude oil and ethylene yield are still measured in tons.

In China, there are about 20 ethylene production plants under 7 technologies. The monthly data from 2003 to 2013 is extracted to illustrate the production prediction and energy optimization of ethylene production processes by the FELM. The input variables of the algorithm (the required fuel, steam, water, electricity and crude oil) and the output variable (ethylene yield) of one actual ethylene plant are described in Fig. 3.

First, the data is fuzzified by Eqs. (9–13). The interval length of the fuzzified group is set as 12, which is the number of available monthly data in a year. After data fuzzification, we can get the maximal, the median and the minimal values of each working substance and the yield

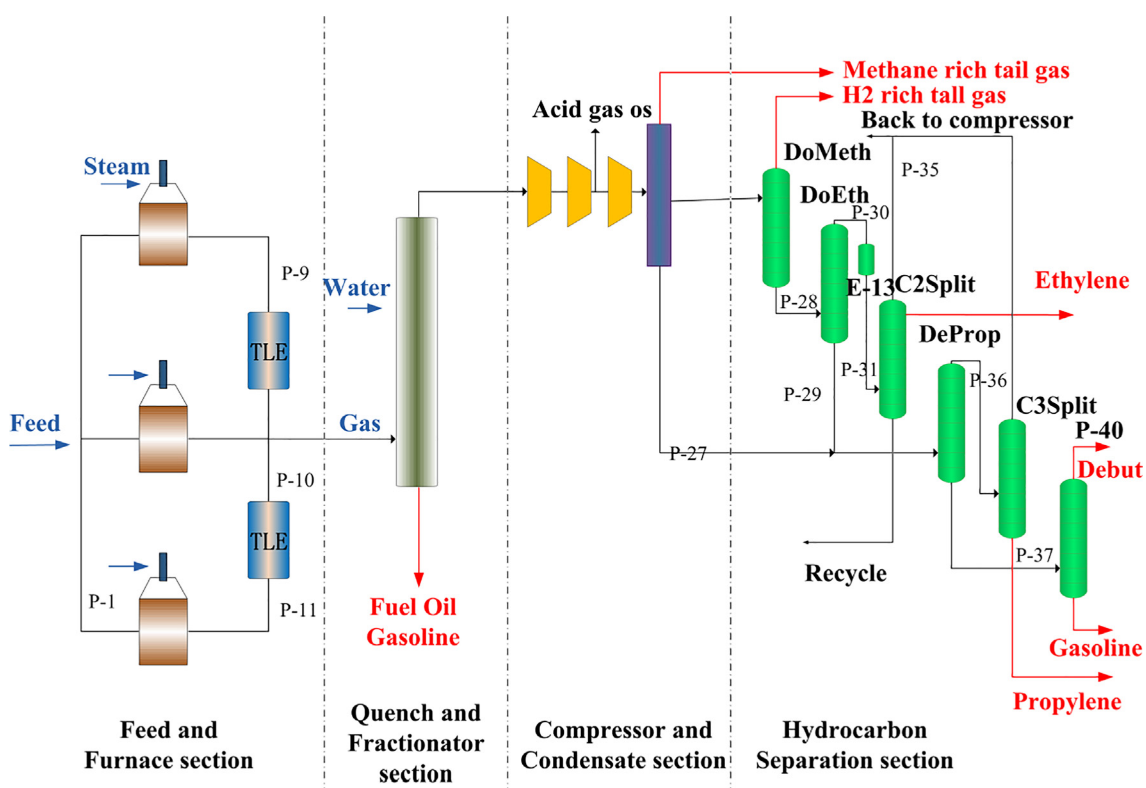
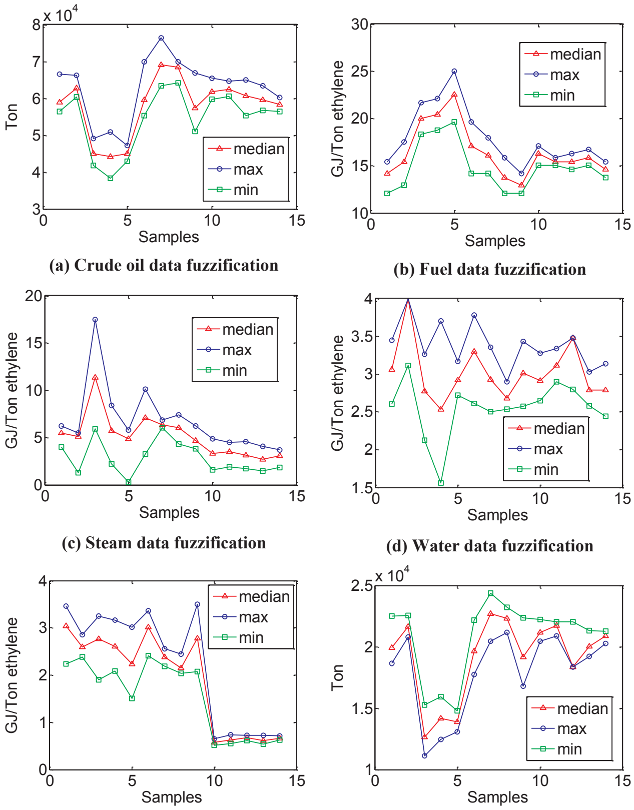


Fig. 2. Ethylene production plant diagram.



(a) Crude oil data fuzzification (b) Fuel data fuzzification (c) Steam data fuzzification (d) Water data fuzzification (e) Electricity data fuzzification (f) Ethylene data fuzzification

Fig. 4. The obtained TFN samples of one plant.

of ethylene plants. On account that the fuzzification is used like a kind of data fusion method, here the number of samples are decreased after processing. The obtained TFN samples of each input and output variables are shown in Fig. 4.

For the record, the crisp data from the practical production has complicated nonlinear timing characteristics, and the overall application performance is seriously affected by abnormal data. Therefore, data preprocessing is of utmost importance to further analysis. The consistency test is taken before data fuzzification. The data is checked by Grubbs Test: if $G \geqslant G(n,\alpha)$, then x_i is discarded (x_i boundary value of ordered sequence). n denotes the number of values and α is the significant level.

$$G = \frac{|x_i - \bar{x}|}{S} \tag{20}$$

where $\bar{x} = (1/N)\sum_{i=1}^{N} x_i$, $S = \sqrt{(1/N)\sum_{i=1}^{N}(x_i-\bar{x})^2}$. The critical values $G(n,\alpha)$ under different parameters can refer to the literature [39].

5.2. Energy management and optimization modeling of ethylene plants

The all fuzzified data is divided into two sets to complete the tasks of training and generalization. According to the previous discussion, we have the following definitions of three models:

Model 1: When the maximal values of each energy-relevant feed stocks like crude oil, fuel, steam, water, electricity are taken as the input and the minimal ethylene yield as the output, the production allocation is a kind of bad practice (bad).

Model 2: When the minimal values of consumption and the maximal values of the yield are used to construct the model, the network output is the expected ideal production under the most efficient situation and can be regarded as a reference for future production adjustment (well).

Model 3: When the median values of the feed stocks are taken as the input of the network and the median values of the ethylene yield are taken as the output, the model represents the daily production situation (median).

In our experiment, compared with three artificial neural networks with the FELM, the FBP and the FRBF, the prediction accuracy and the training time are shown in Table 5. It can be seen from Table 5 that the FELM has the minimum ARGE. Take the bad production practice as an example, the average relative generalization errors of the predictive results for model 1 are 4.61% (FELM), 12.95% (FBP) and 9.32% (FRBF), respectively. Meanwhile, the FELM is significantly faster than the other two compared networks with regard to the required training process, the training time of which is almost 3% of the FBP and the FRBF. The comparisons of predictive results among different networks are demonstrated in Fig. 5, which show that the FELM provides a more detailed simulation for current plants than others.

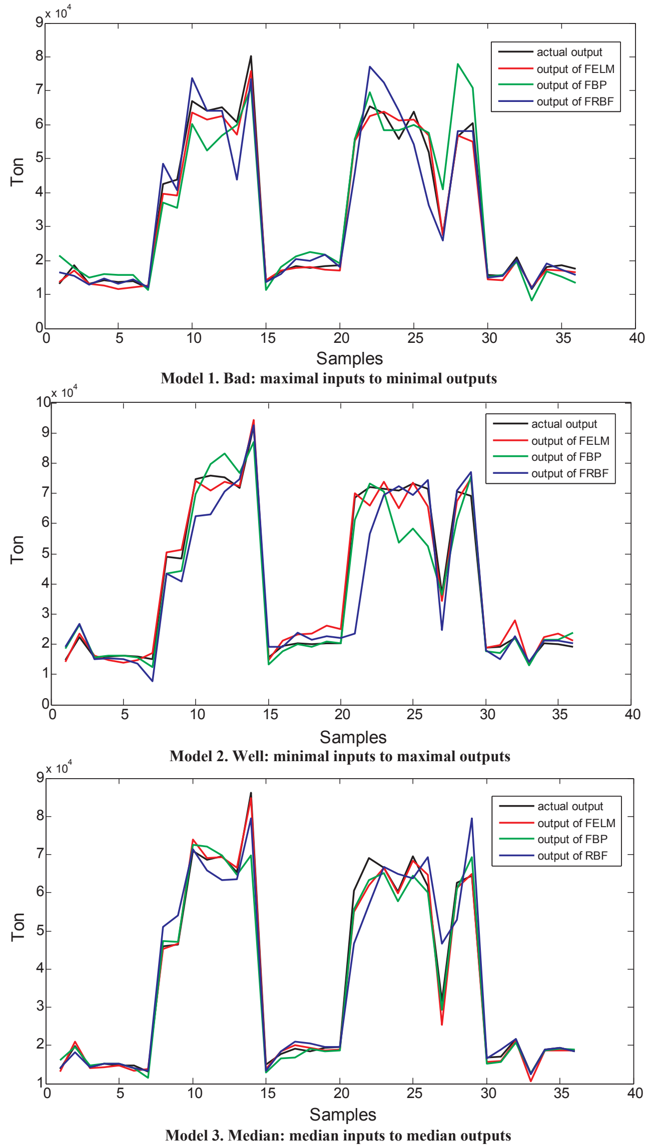
The FELM prediction results for the well, the median and the bad production practices are shown in Fig. 6.

According to what we have discussed, the curves 'well' and 'bad' represent the upper and lower limits of ethylene outputs under specific production inputs, and the curve 'median' is the daily predictive yield under the same production conditions. When the median point is closer to the upper limit instead of the lower limit, the energy efficiency of this ethylene plant is high, and other plants should be adjusted to match the condition and allocation of this plant. For instance, the 10th sample point has a relatively high efficiency from an intuitive view. The median inputs in production are 232,358 tons of crude oil, 21.69 GJ fuel, 1.57 GJ steam, 2.79 GJ water and 0.73 GJ electricity per ton of ethylene. Meanwhile, the median output is 73,982 tons of ethylene and the predictive maximal output is 74,247 tons of ethylene. These two values are very close and the point indicates a direction for future production improvement. On the contrary, the median output of the 21th sample is obviously close to the minimal value, where the median inputs are 187,543 tons of crude oil, 19.30 GJ fuel, 7.93 GJ steam, 1.38 GJ water and 0.82 GJ electricity per ton of ethylene, repectively. The median output is 54,959 tons of ethylene and is far from the predictive maximal output which is 70,084 tons of ethylene. From the previous analysis, it indicates a benchmark of low efficiency and means a huge space for improving the energy efficiency.

A visualized presentation of the proportional relation of the room for efficiency ascending and descending and energy-saving potentials among different production points is shown in Fig. 7. The lower the proportion of the green bar (which represents the ascending room of efficiency), the higher the actual efficiency is. The curve of the energy-saving potential reveals a considerable latent capacity and we can adjust production-related conditions to achieve the optimal production.

Table 5
Test results of the ethylene production data set.

	Model 1		Model 2		Model 3	
	ARGE	Training time	ARGE	Training time	ARGE	Training time
FELM	0.0461	0.0195	0.0547	0.0167	0.0408	0.0179
FBP	0.1295	0.6474	0.1323	0.6533	0.0615	0.6732
FRBF	0.0932	0.6460	0.1267	0.6449	0.0776	0.6889



Model 1. Bad: maximal inputs to minimal outputs

Model 2. Well: minimal inputs to maximal outputs

Model 3. Median: median inputs to median outputs

Fig. 5. The comparison of predictive results among different networks.

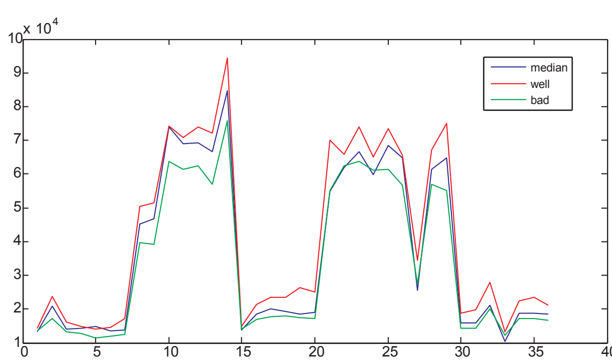


Fig. 6. The FELM prediction results for well, median and bad practices.

6. Discussion

First, the improved FELM can help to reduce the influence of imprecise information with assistance of the fuzzy set theory and get the predictive production model simultaneously. The validation test based on UCI standard data sets proves the superiority of the FELM than that of the FBP and the FRBF.

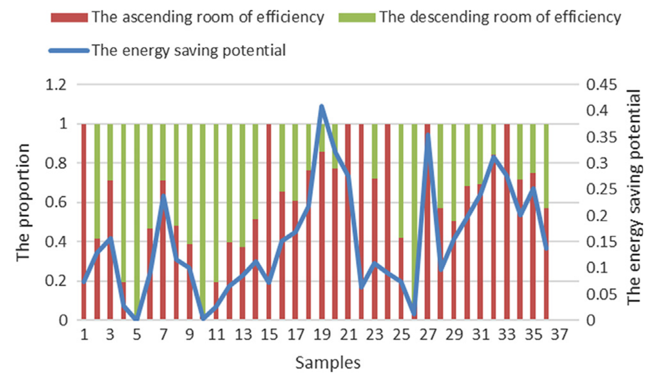


Fig. 7. The proportion between the ascending room and the descending room of energy-saving potentials.

Second, in the case study of the ethylene industry, the FELM establishes the input-output model of the ethylene production and demonstrates the predictive upper, lower limits and the median yield of ethylene. Three models are defined to represent the well, the bad and the daily production practices. The interval belt divided by the median curve shows the ascending and descending room of the energy efficiency. The analysis results provide the theoretical output reference for ethylene yields and explore the optimization direction of the ineffective plants. The experiment shows a considerable energy-saving potential. And the calculated improvement potential is indicated as about 15%.

Third, the proposed method is designed for building the energy management and optimization model in complex petrochemical industries. The case study in this paper is conducted on the basis of the monthly data. The sparsity of statistical data results in obstacles for real time analysis and optimization. Therefore, we will improve our model with data on more detailed time density to adapt the current production in the future. Moreover, we will take some non-technical factors into account, such as profit-related price and environmentally relevant emission index. Furthermore, we will combine some energy analysis and prediction methods, such as entropy-weight, Bayesian evaluation models etc., to compare with the current result of the work.

7. Conclusion

This paper proposes an improved FELM method for build the energy management and optimization model of complex petrochemical industries. The triangular fuzzification is good at processing the imprecise data and the FELM can do well in comprehensive modeling and prediction. The fuzzy set theory is used to deal with energy consumption data from the complex petrochemical process. Instead of the crisp data, TFNs are used as the input and output of the FELM to enhance the fault tolerance of impact caused by noise and data fluctuation. Meanwhile, compared with the FBP and the FRBF, the FELM can produce predictive results with the best accuracy and the fastest convergence speed by using UCI standard data sets. Moreover, in the case study of China ethylene industry, the production data, including crude oil, fuel, steam, water, electricity and yield of ethylene, is fuzzified to obtain TFNs after pretreatment. The cross recombination of the TFNs are then applied in training and generalization of the FELM. With the predefined three models that represent the bad, the well and the median production situations, the upper and the lower limits of energy efficiency are demonstrated, and the analysis results can act as the guidance for decision-makers to improve the production and save the energy. The experiment shows a considerable energy-saving potential existing in the analyzed plants. The latent capacity of production improvement is calculated as about 15%.

Acknowledgments

This research was partly funded by National Natural Science Foundation of China (61533003, 61603025, 61673046), the Natural Science Foundation of Beijing, China (4162045), National Key Research and Development Program of China (2017YFC1601800) and the Fundamental Research Funds for the Central Universities (ZY1703).

Reference

[1] Yu RW, Ma GF, Xu YH. Review of Sinopec's ethylene production in 2015. Ethylene Ind 2016;28(1):1–6.
[2] Zhang LJ, Wang ZY, Xing YC. Review of Petrochina's ethylene production in 2015. Ethylene Ind 2016;28(1):7–11.
[3] Tao R, Patel M, Blok K. Olefins from conventional and heavy feed stocks: energy use in steam cracking and alternative processes. Energy 2006;31(4):425–51.
[4] Bailey JA, Gordon R, Burton D, Yiridoe EK. Energy conservation on Nova Scotia farms: baseline energy data. Energy 2008;33(7):1144–54.
[5] 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.
[6] Zhou P, Ang BW, Zhou DQ. Macro energy efficiency evaluation based on the index decomposition analysis approach. Energy Technol Manage 2007;5:5–8.
[7] Jones WT, Boucinha JM, Inacio CF, Martins PM. Measuring non-residential electric energy efficiency in the Portuguese economy. Energy Procedia 2016;106:2–13.
[8] Lin BQ, Long HY. A stochastic frontier analysis of energy efficiency of China's chemical industry. J Cleaner Prod 2015;87(1):235–44.
[9] Sueyoshi T, Goto M. World trend in energy: an extension to DEA applied to energy and environment. Econom Struct 2017;6(1):1–13.
[10] Han YM, Geng ZQ, Gu XB, Wang Z. Performance analysis of china ethylene plants by measuring malmquist production efficiency based on an improved data envelopment analysis cross-model. Ind Eng Chem Res 2014;54(1):272–84.
[11] Ruggiero J. Impact assessment of input omission on DEA. Int J Inform Technol Decis Making 2005;4(3):359–68.
[12] Geng ZQ, Han YM, Yu CP. Energy efficiency evaluation of ethylene product system based on density clustering data envelopment analysis model. Adv Sci Lett 2012;9(1):735–41.
[13] Zadeh LA. Fuzzy sets. Inform. Control 1965;8:338–53.
[14] Coppi R, D'Urso P, Giordani P. Fuzzy and possibility clustering for fuzzy data. Comput Stat Data Anal 2012;56(4):915–27.
[15] Xiong XH, Yao JC. Research and application of data mining based on fuzzy set. Comput Eng Appl 2002;1:203–5.
[16] Chen CH, Hong TP, Vincent ST. Fuzzy data mining for time-series data. Appl Soft Comput 2012;12(1):536–42.
[17] Kao C, Lin PH. Efficiency of parallel production systems with fuzzy data. Fuzzy Sets Syst 2012;198:83–98.
[18] Han YM, Geng ZQ, Zhu QX, Qu YX. Energy efficiency analysis method based on fuzzy DEA cross-model for ethylene production systems in chemical industry. Energy 2015;83:685–95.
[19] Rumelhart DE, McCelland JL. Parallel distributed processing. Volsland 2: MIT Press; 1986, vol. 1, no. 2. p. 125–87.
[20] 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.
[21] 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(8):1763–73.
[22] Li H, Sanchez R, Qin SJ, et al. Computer simulation of gas generation and transport in landfills. V: use of artificial neural network and the genetic algorithm for short- and long-term forecasting and planning. Chem Eng Sci 2011;66(12):2646–59.
[23] Jiang H, Dong Y, Wang J, Li YQ. Intelligent optimization models based on hard-ridge penalty and RBF for forecasting global solar radiation. Energy Convers Manage 2015;95:42–58.
[24] Barrasso D, Tamrakar A, Ramachandran R. A reduced order PBM-ANN model of a multi-scale PBM-DEM description of a wet granulation process. Chem Eng Sci 2014;119:319–29.
[25] Huang GB, Zhu QY, Siew CK. Extreme learning machine: theory and applications. Neurocomputing 2006;70(1–3):489–501.
[26] 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.
[27] Naji S, Keivani A, Shamshirband S, Alengaram UJ, Jumaat MZ, Mansor Z, et al. Estimating building energy consumption using extreme learning machine method. Energy 2016;97:506–16.
[28] Zheng W, Fu X, Ying Y. Spectroscopy-based food classification with extreme learning machine. Chemometr Intell Lab Syst 2014;139:42–7.
[29] Xiao D, Wang J, Mao Z. The research on the modeling method of batch process based on OS-ELM-RMPLS. Chemometr Intell Lab Syst 2014;134(5):118–22.
[30] Wang J, Hu J. A robust combination approach for short-term wind speed forecasting and analysis combination of the ARIMA (Autoregressive Integrated Moving Average), ELM (Extreme Learning Machine), SVM (Support Vector Machine) and LSSVM (Least Square SVM) forecasts using a GPR (Gaussian Process Regression) model. Energy 2015;93:41–56.
[31] Janakiraman VM, Nguyen XL, Assanis D. An ELM based predictive control method for HCCI engines. Eng Appl Artific Intell 2016; 48(C): 106–18.
[32] Zhang Y, Zhang P. Optimization of nonlinear process based on sequential extreme learning machine. Chem Eng Sci 2011;66(20):4702–10.
[33] Han YM, Geng ZQ, Zhu QX. Energy efficiency analysis method based on fuzzy DEA cross-model for ethylene production systems in chemical industry. Energy 2015;83:685–95.
[34] Geng ZQ, Zhu QX, Gu XB. Study on dependent function analytic hierarchy process model for energy efficiency virtual benchmark and its applications in ethylene equipments. J Chem Ind Eng (China) 2011;62(8):2372–7.
[35] 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.
[36] China standards: the general computing guide of special energy consumption (GB/T2589-2008); 2008.
[37] China standards: the limitation of energy consumption for ethylene product (DB37/751-2007); 2008.
[38] Calculation method for energy consumption in petrochemical engineering design (SH/T3110-2001); 2002.
[39] Geng ZQ, Han YM, Gu XB, Zhu QX. Energy efficiency estimation based on data fusion strategy: case study of ethylene product industry. Ind Eng Chem Res 2012;51(25):8526–34.