Energy Conversion and Management 124 (2016) 73–83

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Energy optimization and prediction of complex petrochemical industries using an improved artificial neural network approach integrating data envelopment analysis

Yong-Ming Han, Zhi-Qiang Geng *, Qun-Xiong 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

ARTICLE INFO

Article history:
Received 27 March 2016
Received in revised form 5 June 2016
Accepted 2 July 2016

Keywords:
Data envelopment analysis
Artificial neural network
Energy optimization
Energy prediction
Ethylene production systems

ABSTRACT

Since the complex petrochemical data have characteristics of multi-dimension, uncertainty and noise, it is difficult to accurately optimize and predict the energy usage of complex petrochemical systems. Therefore, this paper proposes a data envelopment analysis (DEA) integrated artificial neural network (ANN) approach (DEA-ANN). The proposed approach utilizes the DEA model with slack variables for sensitivity analysis to determine the effective decision making units (DMUs) and indicate the optimized direction of the ineffective DMUs. Compared with the traditional ANN approach, the DEA-ANN prediction model is effectively verified by executing a linear comparison between all DMUs and the effective DMUs through the standard data source from the UCI (University of California at Irvine) repository. Finally, the proposed model is validated through an application in a complex ethylene production system of China petrochemical industry. The optimization result and the prediction value are obtained to reduce energy consumption of the ethylene production system, guide ethylene production and improve energy efficiency.

1. Introduction

Nowadays, energy conservation and emission reduction of the industry is required to meet the challenges of the sustainable development. And in complex petrochemical industries, energy optimization and prediction are effective means to achieve both environmental and economic goals. Wherein, the ethylene industry is one of the most key parts of the petrochemical industry. According to statistics data, China Petrochemical Corporation's ethylene production was 10,420 kt/a, and the average fuel and power consumption (converted to the standard oil) was 571.39 kg per ton of ethylene in 2014 [1]. The ethylene production capacity and the average fuel and power consumption of China National Petroleum Corporation was 4976 kt/a and 616.7 kg per ton of ethylene in 2014, respectively [2]. The energy efficiency of ethylene industries in China is significantly lower than that in the advanced countries, and there is great improvement space for energy efficiency in China ethylene industries. Moreover, the energy consumption cost of ethylene production plants took up over half of the operation cost [3]. Therefore, studying about energy optimization and prediction of ethylene industries is beneficial for the sustainable development of Chinese economy.

Currently, the mean method and the optimal index method are commonly used by enterprises to analyze the energy efficiency [4]. Because energy efficiency values contain indexes and influencing factors, and various indexes make different effects to energy efficiency factors, these methods cannot provide the energy efficiency benchmark of indexes or excellent factors to optimize the practical state of the energy usage. Han et al. proposed the energy efficiency evaluation method of ethylene production plants based on Analytic Hierarchy Process (AHP) and Data Envelopment Analysis (DEA), with better results [5]. Meanwhile, Han et al. multi-level analyzed the month and year energy efficiency of the petrochemical process under different technologies and scales [6]. However, all these methods did not take the impact factor of crude oils into consideration. In addition, Dias et al. used chemical and biological alternated step-size control to analyze the Dengue vector control problem in a multiobjective optimization approach [7]. Grom et al. used a generalized kinetic model and an appropriate algorithm to evaluate the rate parameters and subsequent regression of the chemical engineering [8]. However, the economic cost of

* 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.-Q. Geng), zhuqx@mail.buct.edu.cn (Q.-X. Zhu).

http://dx.doi.org/10.1016/j.enconman.2016.07.002
0196-8904/

Nomenclature

DEA	data envelopment analysis
ANN	artificial neural network
DMU	decision making units
UCI	University of California at Irvine
BP	back propagation
C4	Carbon 4
C^2WH	Charnes, Cooper, Wei, Huang
x_{ij}	ith input of DMU_{-j}
y_{rj}	rth output of DMU_{-j}
α	weight coefficients of m inputs
β	weight coefficients of s outputs
ω_j	the weight vector of inputs and outputs
ε	non-Archimedean infinitesimal
τ_t^- and τ_t^+	the slack variables
λ	the optimum values of the DMU_{-j0}
$X=(x_1,x_2,\ldots,x_i,\ldots,x_m)^T$	the input vector
$V=(v_1,v_2,\ldots,v_j,\ldots,v_l)^T$	the weight vector of the input-layer to the hidden-layer
$W=(w_1,w_2,\ldots,w_k,\ldots,w_n)^T$	the weight vector of the hidden-layer to the output-layer
$g(x)$	The sigmoid function
$Y=(y_1,y_2,\ldots,y_j,\ldots,y_l)^T$	the output vector of the hidden-layer
$O=(o_1,o_2,\ldots,o_k,\ldots,o_n)^T$	the output vector of the output-layer
$d=(d_1,d_2,\ldots,d_k,\ldots,d_n)^T$	the desired output vector

transforming petrochemical plants was not included. Therefore, this paper proposes the energy optimization and prediction method of China petrochemical industries based on the DEA-integrated artificial neural network (ANN) (DEA-ANN).

The DEA method is firstly proposed by the famous operational researchers Charnes, Cooper and Rhodes (CCR) in 1978 [9]. When it was applied in production apartment of multi-inputs and multi-outputs, the DEA was proved both sizeable effective and technological effective. As a result, the DEA method is satisfactory and effective with application in practical engineering fields [10], especially in the petrochemical industry. Sueyashi et al. studied the DEA method and the DEA window analysis to analyze the operational and environmental performance of the coal-fired power plant, respectively [11,12]. Han et al. evaluated the energy efficiency of ethylene industries based on the combination of interpretative structural model (ISM) and the traditional DEA [13]. Sueyoshi et al. studied the DEA radial measurement approach to optimize energy utilizations and environmental protections of Japanese chemical and pharmaceutical firms [14]. Bi et al. investigated the relationship between the fossil fuel consumption and the environmental regulation of China's thermal power generation by the DEA model [15]. Zhu et al. analyzed many factors influencing the energy use of an ethylene plant and evaluated the efficiency of the ethylene plants by principal component analysis (PCA) integrated DEA models [16]. Han et al. utilized the fuzzy DEA cross-model to analyze and evaluate the energy efficiency of Chinese ethylene production systems [17]. However, the results were easily affected by the quantities of samples and the numbers of input and output factors [18,19], and the energy saving is not predicted. When the quantities of samples are too small or the numbers of input and output factors are too large, more than one-third of efficiency values, which are obtained by the DEA model, are then set to 1 [20,21]. Moreover, the weight distribution of the input and output factors is generally unreasonable, so the maximum efficiency DMU of the DEA model was not obtained [22]. Therefore, the ANN is applied to analyze and predict the energy usage of the petrochemical industry.

The ANN algorithm is a learning algorithm used for the multi-layer feed-forward neural network (FNN). And the ANN with error back propagation (BP neural network) was designed by the research group led by Rumelhart and McCelland in 1986, based on the error back propagation algorithm [23]. The ANN is widely applied in prediction and optimization of different fields. Avramovic et al. used the ANN integrated genetic algorithm (GA) method to optimize the sunflower oil ethanolysis catalyzed by calcium oxide [24]. Jiang et al. studied an intelligent optimization models based on hard-ridge penalty and RBF for forecasting global solar radiation [25]. Meo et al. studied a new hybrid approach obtained combining a multi-objective particle swarm optimization (PSO) and ANN for the design optimization of a direct-drive permanent magnet flux switching generators for low power wind applications. [26]. Although the BP algorithm is a powerful supervised learning algorithm for training multi-layer FNNs [27,28], this gradient descent-based algorithm often suffers from problems of the local minima and slow convergence rate. Aiming at ameliorating the convergence properties of the BP neural network [29], many proposals have been made by researchers in the last twenty years. Dai et al. studied a two-phased and ensemble-scheme integrated BP (TP-ES-BP) algorithm, which could dispose the local minima of the standard BP (SBP) algorithm better, and overcome the limitations of individual component BP algorithm in classification performance [30]. In order to improve the learning convergence of the BP, some researchers have tried to integrate the BP with the PSO method or with the GA to overcome its shortcomings [31,32].

In order to overcome the insufficient of the DEA and the BP neural network and predict the energy usage better, Wu get the Travel & Tourism competitiveness rankings reasonably using the integrated method with the grey system theory (GST), the DEA, the ANN and the Borda count [33]. Olanrewaju et al. proposed an integrated approach, including logarithmic mean divisia index (LMDI), a type of index decomposition analysis (IDA), the ANN, and the DEA methods, to evaluate and optimize the energy efficiency of the industrial sector in Canada [34,35]. This paper proposed an integrated method based on the BP with the momentum factor and the DEA model to optimize and predict the energy usage of ethylene production plants in the petrochemical industry. First, in terms of the model accuracy and training time, the robustness and effectiveness of the DEA-BP model is better than the BP. And then the proposed approach utilizes the DEA model with slack variables to find the effective DMUs and obtain the optimized direction of the ineffective DMUs. Based on the energy efficiency analysis and optimization results of the DEA, we utilize the energy consumption data of crude oils, fuel, steams, water, electricity, and the productions of ethylene, propylene and carbon 4 (C4), as inputs and outputs of the BP, respectively. In the BP model based on the effective DMUs, we introduce the momentum factor to avoid the problems of the local minima and slow convergence. The proposed method is validated by applying in energy optimization of the production data, and prediction of the ethylene production system in China petrochemical industry. Meanwhile, the prediction values are obtained to reduce energy consumption of the ethylene production system, guide ethylene production and improve energy efficiency.

2. Experimental

The DEA method is a classic performance or efficiency evaluation method. The discrimination of analysis results are lowered by a large number of input and output factors. When input and

output indicators remain constant, if the quantity of samples increase, the ratios of effective units and the efficiency values of the DEA will decrease. In the DEA model, the reasonably selected input and output indicators will produce better results. The ANN model often suffers from the local minima and slow convergence. However, the ANN model based on effective DMUs from the DEA model can reach rapid convergence rate. On the other hand, the introduced momentum factor could help to avoid the problem of the local minima. Meanwhile, this proposed method improves the prediction accuracy of the production multi-criteria DMUs.

2.1. The DEA model

The *CCR* model is the first model used in DEA analysis. Moreover, in order to reflect the preferences of DMUs and optimized results of the technology and the scale, Charnes et al. established a new C^2WH-*DEA* model (Charnes, Cooper, Wei, Huang) in 1986 [36].

There are n supposed plants or departments (DMUs). Each DMU has m inputs and s outputs. x_{ij} is the ith input of the DMU_{-j}, where y_{rj} is the rth output of the DMU_{-j} $(i = 1,2,\ldots,m;\ j = 1,2,\ldots,n;\ r = 1,2,\ldots,s)$. And the DMU_{-j_0} has the corresponding input data as $x_0 = x_{j_0}$ $(1 \leqslant j_0 \leqslant n)$ and output data as $y_0 = y_{j_0}$. The C^2WH-*DEA* model for evaluating the DMU_{-j_0} is shown in Eq. (1) as the fractional programming.

$$\begin{cases} \max \gamma^T y_0 \\ \alpha^T x_j - \beta^T y_j \in K,\ j = 1,2,\ldots,n \\ \alpha \in A - \{0\} \\ \beta \in B - \{0\} \end{cases} \tag{1}$$

Wherein, $x_j = (x_{1j}, x_{2j}, \ldots, x_{mj})^T > 0$, $y_j = (y_{1j}, y_{2j}, \ldots, y_{sj})^T > 0$. And $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)^T$, $\beta = (\beta_1, \beta_2, \ldots, \beta_s)^T$ denote the weight vector of inputs and outputs, respectively. $A \subset R^m$, $B \subset R^s$, $K \subset R^n$, and $\ln A \neq \varnothing$, $\ln B \neq \varnothing$. The C^2WH-*DEA* model transformation equation for the fractional programming is shown in Eq. (2).

$$\begin{cases} \max \gamma^T y_0 = \alpha^0, \\ \alpha^T x_j - \beta^T y_j \geqslant 0,\ j = 1,2,\ldots,n, \\ \alpha^T x_0 = 1, \\ \alpha \geqslant 0,\ \beta \geqslant 0 \end{cases} \tag{2}$$

Meanwhile, we use the equivalent linear programming to transfer the C^2WH-*DEA* model into the form as in Eq. (3).

$$\begin{cases} \min \lambda, \\ \sum_{j=1}^{n} x_j \omega_j \leqslant \lambda x_0, \\ \sum_{j=1}^{n} y_j \omega_j \geqslant y_0, \\ \omega_j \geqslant 0,\ j = 1,2,\ldots,n, \lambda \in O_1. \end{cases} \tag{3}$$

Wherein, O_1 is a unit vector and ω_j denote the weight vector of inputs and outputs. We use Eq. (3) to obtain Eq. (4) (including the non-Archimedean infinitesimal ε and slack variables):

$$\begin{cases} \min \left[\lambda - \varepsilon(e_1^T \tau^- + e_2^T \tau^+)\right], \\ \sum_{i=1}^{n} \omega_i x_{ji} + \tau^- = \theta x_{jN},\ j = 1,2,\ldots,m \\ \sum_{i=1}^{n} \omega_i y_{ri} - \tau^+ = y_{rN},\ r = 1,2,\ldots,s \\ \omega_i \geqslant 0,\ i = 1,2,\ldots,n, \\ \tau^- \geqslant 0, \tau^+ \geqslant 0,\ N = 1,2,\ldots,n \end{cases} \tag{4}$$

Wherein, the non-Archimedean value ε is used to guarantee the strict positivity of variables. ε is also an "abstract number", which is greater than 0 and less than any positive number. Generally, in order to compute the convergence, the value ε is set as large as possible, or as small as possible to ensure that it is near to the infinitesimal. In this paper, we set ε as 10^{-6} to separate effective and weakly effective DMUs [37–39]. $e_1^T = (1,1,\ldots,1) \in R^m$, $e_2^T = (1,1,\ldots,1)^T \in R^s$. And τ_t^- and τ_t^+ are the slack variables. $\tau_t^- = (tau_t^{1-}, \tau_t^{2-}, \ldots, \tau_t^{m-})^T$ and $\tau_t^+ = (\tau_t^{1+}, \tau_t^{2+}, \ldots, \tau_t^{r+})^T$ are the redundant amount of input factors and the shortfall of output factors, respectively. λ is the optimum value of the DMU_{-j_0}, which means the relative efficiency value of the input to the output. Therefore, when $\lambda = 1$, if the two constrained inequalities are established, the DMU_{-j0} is effective, or it is weakly effective. When $\lambda < 1$, the DMU_{-j0} is ineffective. All effective DMUs make up the production effective frontier. Meanwhile, the lower the value of λ is, the further relative efficiency is from the production frontier.

2.2. The artificial neural network with error back propagation (BP)

The three-layer neural network can realize any kinds of nonlinear projection, while the four-layer network is more liable to be trapped in the local minima [40]. Meanwhile, too many network nodes might weaken the generalization and prediction abilities. Further, as Hecht-Nielsen pointed out that, for the neural network with only one hidden layer, as long as its nodes are enough, it could approximate the nonlinear function with any precision [41]. Therefore, this paper adopted the three-layer FNN, whose classic structure was shown in Fig. 1. The network is consisted of one input layer, one hidden layer and one output layer. The input signals are successively transmitted from the input-layer node through the hidden-layer node, to the output-layer node. The nodes of each layer can only affect that of the next layer.

In Fig. 1, $X = (x_1, x_2, \ldots, x_i, \ldots, x_m)^T$ is the input vector.

$V = (v_1, v_2, \ldots, v_j, \ldots, v_l)^T$ is the weight vector of the input-layer to the hidden-layer.

$W = (w_1, w_2, \ldots, w_k, \ldots, w_n)^T$ is the weight vector of the hidden-layer to the output-layer.

The sigmoid function is designed as $g(x)$:

$$g(x) = \frac{1}{1 + e^{-x}} \tag{5}$$

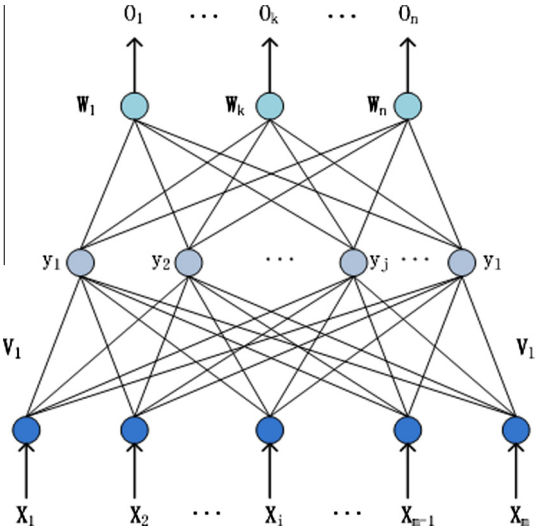

Fig. 1. The three-layer BP neural network model.

The input parameter of the sigmoid function could be adopted as any value, but the output value is limited between 0 and 1. For the advantage of differentiability of the function, it is usually used in the BP network. However, the output layer may often make use of the linear function for no limitation of output values.

$Y = (y_1, y_2, \ldots, y_j, \ldots, y_l)^T$ is the output vector of the hidden-layer.

Wherein:

$$y_j = g(net_j)(j = 1, 2, \ldots, l) \tag{6}$$

$$net_j = \sum_{i=0}^{m} v_{ij} x_i (j = 1, 2, \ldots, l) \tag{7}$$

$O = (o_1, o_2, \ldots, o_k, \ldots, o_n)^T$ is the output vector of the output-layer.

Wherein:

$$o_k = g(net_k)(k = 1, 2, \ldots, n) \tag{8}$$

$$net_k = \sum_{j=0}^{l} w_{jk} y_j (k = 1, 2, \ldots, n) \tag{9}$$

$d = (d_1, d_2, \ldots, d_k, \ldots, d_n)^T$ is the desired output vector. And we set values of the desired error *xi* and the training times *T*.

The BP algorithm often suffers from problems of the local minima and slow convergence. Thus an improved BP network algorithm is adopted in this paper. The algorithm tries to reduce the vibration trend of the learning process through the introduction of the momentum factor, as well as to improve the convergence rate. Since the momentum factor could reduce the network sensibility to local details of error curved surface, and escape the local minima, the BP algorithm could obtain better solutions.At the same time, the initial value of the l_{th} node in the hidden-layer could be determined through the following equation [42].

$$l = \sqrt{0.12n^2 + 0.43mn + 0.77n + 2.54m + 0.35} + 0.51 \tag{10}$$

Then other nodes of the hidden layer could be selected by using the gradual increasing method, in order to shorten the training time. The method starts from a simple network, and gradually increases units of the hidden layer to the proper amount till the requirement is met.

The execution steps of the improved BP algorithm are described in the following:

(1) Calculate errors between the desired output values and the network output values by Eq. (11).

$$E = \frac{1}{2}(d - O)^2 = \frac{1}{2}\sum_{k=1}^{n}(d_k - o_k)^2 \tag{11}$$

(2) Replace the output values in the above equation with the expanded calculation results from the hidden layer.

$$E = \frac{1}{2}\sum_{k=1}^{n}[d_k - g(net_k)]^2 = \frac{1}{2}\sum_{k=1}^{n}\left[d_k - g\left(\sum_{j=0}^{l} w_{jk} y_j\right)\right]^2 \tag{12}$$

(3) Replace the value of y_j with the expanded calculation results from the input layer.

$$E = \frac{1}{2}\sum_{k=1}^{n}\left[d_k - g\left(\sum_{j=0}^{l} w_{jk} y_j\right)\right]^2 = \frac{1}{2}\sum_{k=1}^{n}\left\{d_k - g\left[\sum_{j=0}^{l} w_{jk} g\left(\sum_{i=0}^{m} vijx_i\right)\right]\right\}^2 \tag{13}$$

(4) Adjust the weight values to decrease the error gradually.

$$\Delta v_{ij} = -\eta \frac{\partial E}{\partial v_{ij}} (i = 0, 1, 2, \ldots, m;\ j = 1, 2, \ldots, l) \tag{14}$$

$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}} (j = 0, 1, 2, \ldots, l;\ k = 1, 2, \ldots, n) \tag{15}$$

Wherein, η denotes the learning rate, i.e. the step length, valued between 0 and 1.

$$\Delta v_{ij} = -\eta\left[(1 - \alpha)\frac{\partial E}{\partial v_{ij}} + \alpha \frac{\partial E}{\partial v_{i(j-1)}}\right]; \quad (i = 0, 1, 2, \ldots, m;\ j = 1, 2, \ldots, n) \tag{16}$$

$$\Delta w_{jk} = -\eta\left[(1 - \alpha)\frac{\partial E}{\partial w_{jk}} + \alpha \frac{\partial E}{\partial w_{j(k-1)}}\right]; \quad (j = 0, 1, 2, \ldots, l;\ k = 1, 2, \ldots, n) \tag{17}$$

Wherein, α ($0 \leqslant \alpha < 1$) means the momentum factor [43].

$$V(t + 1) = V(t) + \Delta V \tag{18}$$

$$W(t + 1) = W(t) + \Delta W \tag{19}$$

(5) If $E < \xi$, or the training times t = T, or the weight values meet the requirement, stop the calculation. Else, return back to step 2.

2.3. Energy optimization and prediction framework based on DEA-integrated ANN (BP)

Step 1: Select the input data. In the paper, the input and output variables are normalized to meet input requirements of the neuron transfer function. Thus, each variable is offered the same importance and would vary in not too wide range. Meanwhile, the training difficulty would be diminished. Moreover, in order to avoid difficulties in adjusting weight values of the hidden-layer to the input-layer, and to guarantee that the convergence rate and precision of the network be free from influence, the normalization method is illustrated as Eq. (20). Each weight variable would be adjusted in the range of 0–1 [44].

$$x'_{ij} = \frac{x_{ij} - x_j^{\min}}{x_j^{\max} - x_j^{\min}} \times 0.8 + 0.1 \tag{20}$$

The inverse normalization of variables in the network is taken in Eq. (21).

$$x'_{ij} = (x_{ij} - 0.1) \times 1.25 \times \left(x_j^{\max} - x_j^{\min}\right) + x_j^{\min} \tag{21}$$

Wherein, x_{ij} indicates the variable of the jth sample in the ith network, $x_j^{\max} = \max\{x_{1j}, x_{2j}, \ldots, x_{nj}\}$, $x_j^{\min} = \min\{x_{1j}, x_{2j}, \ldots, x_{nj}\}$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$.

And then carry on data normalization using Eqs. (20) and (21).

Step 2: Analyze energy efficiency by Eqs. (1)–(4) of the DEA and find the effective improvements for ineffective DMUs by slack variables.

Step 3: Establish a BP based prediction model using the data of all effective DMUs and the original data to by Eqs. (5)–(19), respectively.

Step 4: Energy prediction and analysis based on the prediction model.

Diagrammatically, the energy optimization and prediction framework based on DEA-BP is described in Fig. 2.

2.4. Algorithm comparison based on standard data sets

In order to validate the effectiveness and robustness of the DEA-BP model, this paper uses some standard data sets (Abalone

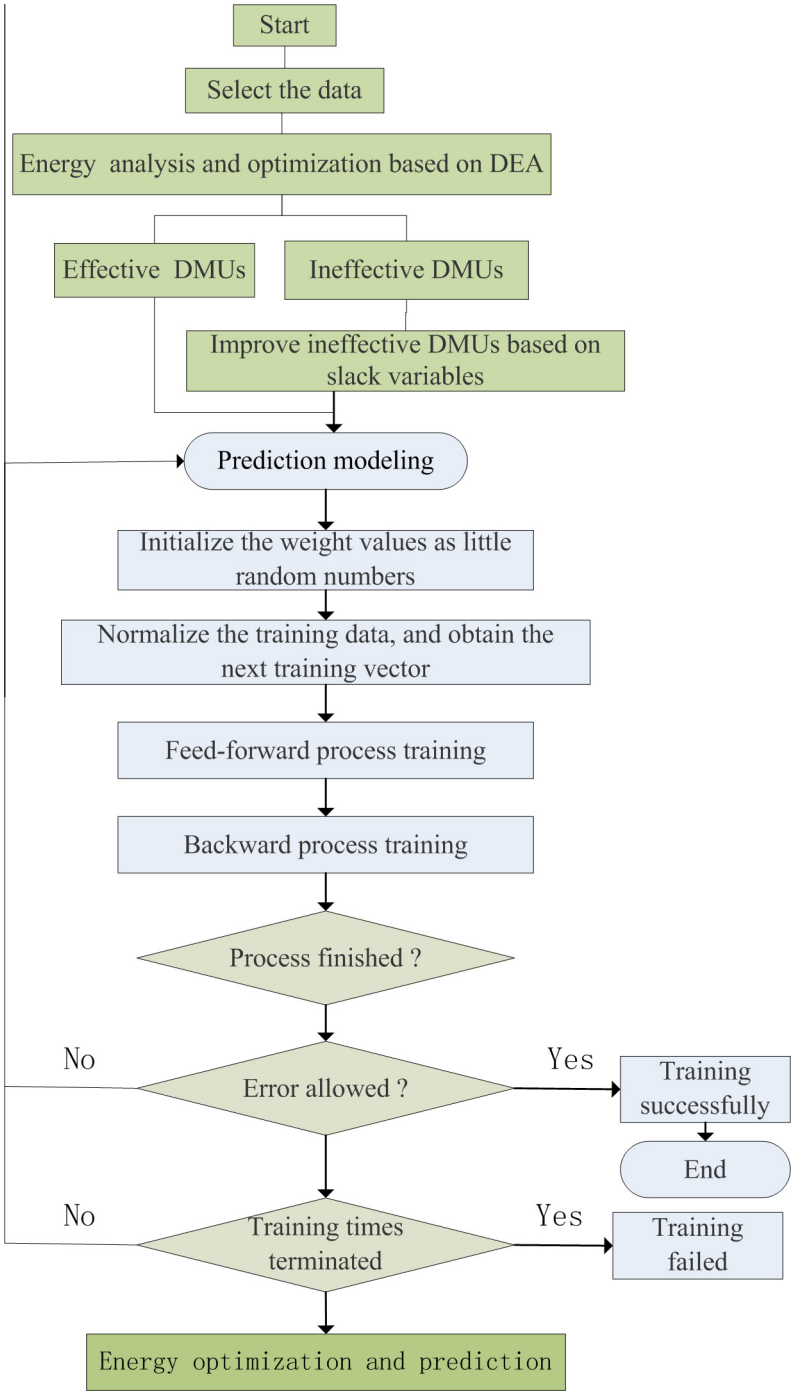

Fig. 2. Energy optimization and prediction framework based on DEA-BP.

Table 1
Specification of standard data sets.

Data sets	#Samples		#Attributes	
	Training	Testing	Inputs	Output
Abalone	274	137	7	1
Parkinsons	592	295	18	1

and Parkinsons telemonitoring (Parkinsons)) from UCI repository (For the details of the dataset, one can refer to the following website: http://archive.ics.uci.edu/ml/datasets.html), whose specification is listed in Table 1.

We build the BP and DEA-BP network model by setting the step length as 0.7, the momentum factor as 0.3, the training times as 1000 and the maximum allowed error as 0.001 in error back-propagation procedure. The average relative generalization error (ARGE) and root mean square error (RMSE) are obtained by Eqs. (22) and (23), respectively.

$$\begin{cases} \begin{cases} RGE_i = Abs\left(\frac{Nout_i^{inver} - Eout_i^{inver}}{EOut_i^{inver}}\right), & if\ Eout_i^{inver} \neq 0 \\ RGE_i = Abs\left(Nout_i^{inver} - Eout_i^{inver}\right), & else \end{cases} \\ ARGE = \frac{\sum_{i=1}^{n} RGE_i}{n} * 100 \end{cases} \tag{22}$$

Table 2
Comparison of the ARGE and RMSE of BP and DEA-BP.

Data sets	Nodes	BP			DEA-BP		
		ARGE	RMSE	Time	ARGE	RMSE	Time
Abalone	6	0.180	0.157	0.071	0.178	0.156	0.079
	15	0.177	0.156	0.083	0.171	0.151	0.091
	30	0.163	0.149	0.097	0.160	0.147	0.099
Parkinsons	8	0.0627	0.0617	0.102	0.0414	0.0416	0.110
	15	0.0573	0.0568	0.145	0.0394	0.0397	0.157
	25	0.0794	0.0790	0.153	0.0615	0.0609	0.171

$$\begin{cases} MSE_i = (Nout_i^{norm} - Eout_i^{norm})^2 \\ RMSE = \sqrt{\frac{\sum_{i=1}^{n} MSE_i}{n}} \end{cases} \quad (23)$$

where $Nout_i^{inver}$ is the inverse normalized network output for all samples, $Eout_i^{inver}$ is the expected output, and $Nout_i^{norm}$ is the normalized network output for all samples, $i = 1, 2, \ldots\ldots, n$.

Based on the DEA model, the number of the effective DMUs in the Abalone and Parkinsons datasets is 7 and 31, respectively. Due to the less number of samples, the BP based prediction model cannot be established. However, the effective improvements for ineffective plants are executed by slack variables of the DEA, so the effective DMUs is enough to build the BP based prediction model. Meanwhile, we assign the initial hidden layer node as 6 and 8 of Abalone and Parkinsons datasets by Eq. (10) and increase the hidden layer node of the BP network and make the simulation again. The ARGE, the RMSE and the training time (Time, Unit: second) are listed in Table 2.

It can be seen from Table 2 that both ARGE and RMSE of the DEA-BP model are better than that of BP when nodes of hidden layer increase for Abalone and Parkinsons data sets. However, the computing time is longer along with increasing hidden nodes. According to the Table 2, the effectiveness of the DEA-BP model is validated.

3. Results

3.1. Ethylene production system

There are about seven ethylene production technologies in the ethylene production industries in China [45]. In this paper, we take the Lummus order separation technology as a case study to show the availability of the proposed method.

The ethylene production system mainly contains processes of cracking and separation [13]. When the cracking furnace is at work, a great number of fuels are used to provide heat for cracking reactions in the tube. And the Transfer Line Exchanger (TLE) produces a large number of steams by recovering the waste heat. With the purpose of making the raw material hydrocarbon finish the optimal cleavage reactions in a short while, and reducing the coke, the steam should be infused while the hydrocarbon is put into the cracking furnace.

The separation section is divided into the rapid cooling part, the compression part and the separation part. The main consumed energy concludes the power consumption in the compressor, the cooling energy consumption in the compressor and the cold box, the reboiler consumption in the distillation column and the steam consumption in the heat separation part. The main part of the energy consumption in cracking section contains the preheat of the mixture of the feedstock and the stream, the reaction heat consumption in the cracking reactions, and the waste heat released to the environment like afterheat in the flue gas. A typical framework of the ethylene plant flowchart is shown in Fig. 3 as follows.

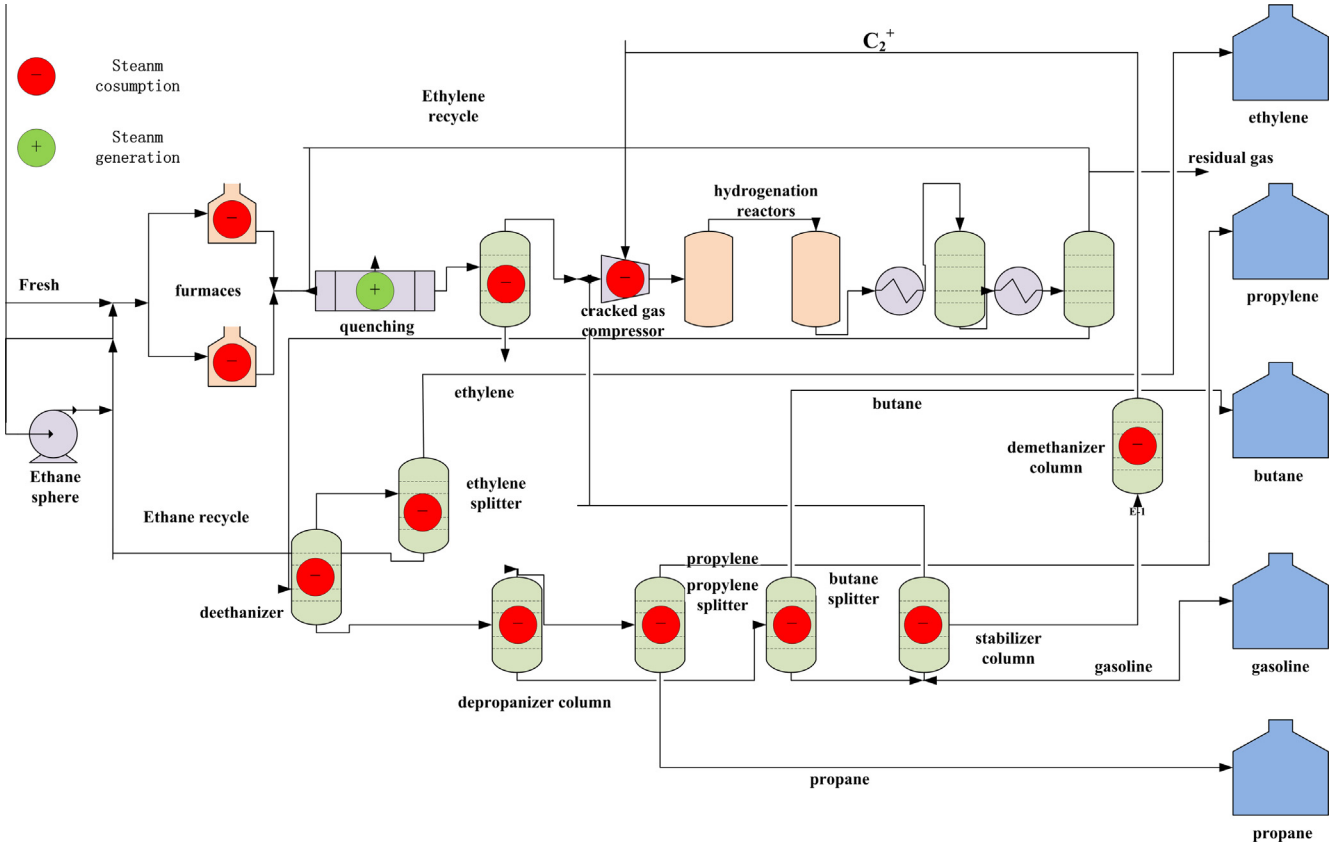

Fig. 3. A typical flowchart of the ethylene plant.

3.2. Energy utilization boundary of ethylene production system

In the ethylene production system, different ethylene production plants usually have different energy utilization boundaries. In order to make a unanimous criterion of computing energy efficiency objectively, the national standards for energy consumption (GB/T2589-2008) and the ethylene industrial standards (DB37/751-2007) are adopted in China [46,47]. We can thus get the energy utilization boundaries of the above mentioned ethylene production process [45].

For the ethylene utilization boundaries, the main energy types are shown in the following: electricity; steams, including low pressure steam, middle pressure steam, high and super-high pressure steam; water, including recycle water, boiler water and industrial water; fuel, including light oil, fuel gas and heavy oil; compressing air and N_2. Since the consumption of compressing air and N_2 is very little, they are not taken into consideration in the energy efficiency analysis and prediction of the ethylene production process.

In the ethylene system, the ethylene, propylene and C4 are the main productions of the ethylene plants. Therefore, the consumption of crude oils, fuel, water, steams and electricity are taken as the input of ethylene systems, and the productions of propylene, ethylene and C4 are taken as the output.

3.3. Data preprocessing

Ethylene production data have complex nonlinear timing relationship, and may include abnormal and noisy data. And the dimension of variables in the timing data is usually different, which makes values of variables incomparable [45]. Meanwhile, the efficiency value of the DEA is relative. The multi-input and multi-output data of ethylene plants and data precision easily affect the accuracy of analysis results. Therefore, the data would be reprocessed by consistency tests, uniform dimension disposal of units and the normalized. According to Grubbs criterion [48], the lengthy data could be cleared by Eq. (24): if $T \geqslant T(Q,\alpha)$ (Q is the quantity of data and α expresses the significant level), then x_i is eliminated [45].

$$T = \frac{|V|}{S} = \frac{|x_i - \bar{x}|}{S} \tag{24}$$

Wherein, $\bar{x} = \frac{1}{Q}\sum_{i=1}^{Q} x_i$, $S = \sqrt{\frac{1}{Q}\sum_{i=1}^{Q}(x_i - \bar{x})^2}$.

Based on characteristics of the energy data of ethylene production plants, and Tables 3.0.2 and 3.0.3 from Energy Consumption Calculation Method of Petroleum Chemical Design (SH/T3110-2001) [49], all measurement units of energy consumption parameters of steams, electricity, fuel and water are translated into the uniform GJ/Ton ethylene (GJ per ton of ethylene).

3.4. The energy optimization and prediction process of ethylene plants

The ethylene production process is deemed to be a multi-input and multi-output process. After the data reprocessing work, the multi-dimensional data can be processed more objectively and accurately. Then, the energy optimization and prediction results of ethylene production plants could be improved significantly by the proposed method. Through these measures, we are able to demonstrate how to improve the production design and the energy efficiency of ethylene production plants. And we could also find the ways to increase productions of plants. The energy optimization and prediction flowchart of ethylene plants based on the DEA-BP method is shown in Fig. 4 as follows.

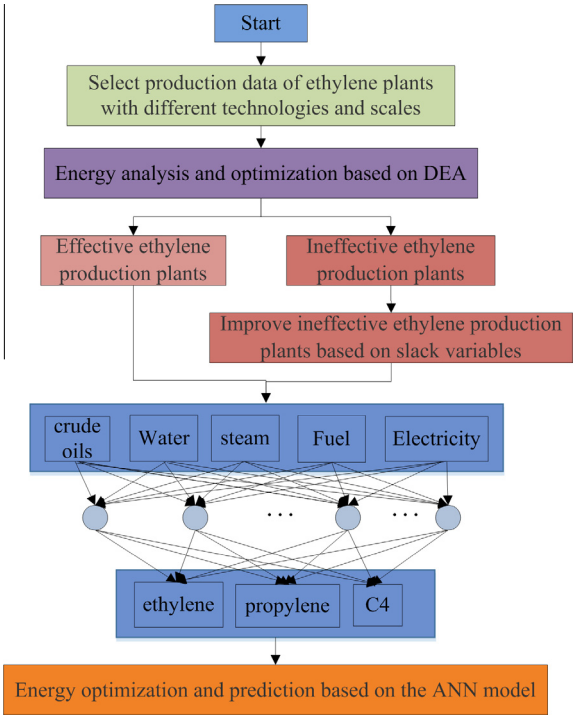

Fig. 4. Ethylene energy optimization and prediction flowchart based on the DEA-BP method.

3.5. Energy optimization and prediction of ethylene systems

With the purpose of illustrating the energy optimization of ethylene production industries, it is important to take the order separation technology, with the scale of 800,000 tons of ethylene per year, as the case study. Correspondingly in China, there are Qilu, Saike, Maoming, Yangtze, Yanshan, and Tianjin ethylene plants. And there are a lot of monthly data from these ethylene plants in the period of 2011–2013. In Fig. 5 shown are curves of all required crude oils, fuel, water, steams and electricity of one ethylene plant in 2011. And in Fig. 6 are monthly productions of ethylene, propylene and C4 of the same ethylene plant in 2011.

Firstly, we reprocess data of crude oils, water, electricity, steams and fuel from the order separation technology, as well as productions of ethylene, propylene and C4 by Eq. (24). And then resources of crude oils, fuel, water, steams, and electricity are taken as inputs of the DEA, and productions of propylene, ethylene and C4 as outputs. Based on the DEA model, monthly energy efficiency values of the order separation technology could be produced by Eq. (4) in Fig. 7. And slack variables of inefficient ethylene plants are shown in Table 3, respectively.

It can be seen from Fig. 7 that efficiency values of the plant 3 and plant 4 in most months are 1, which indicate that their productions of this plant in these months are effective. However, efficiency values of the plants 1, 2, 5 and 6 in most months are less than 1, which indicate that productions of them are ineffective. The whole production status of the plant 3 and plant 4 is better than that of the plants 1, 2, 5 and 6. Therefore, from Table 3, we could conclude that if the efficiency value is less than 1, the improvement quantity of ineffective plants can be calculated by slack variables. For instance, the efficiency value of the plant 1 is 0.995 in February 2011. If it has outputs increase 2410.79 tons of propylene and 505.71 tons of C4, and inputs reduce 0.18 GJ steams, 1.17 GJ fuel and 0.07 GJ electricity of per ton of ethylene, then the efficiency value can improve the ratio of 0.5% and reach the effective level. Take another instance, the efficiency value of the plant 1 is 0.98 in July 2011. If its outputs ethylene and propylene increase

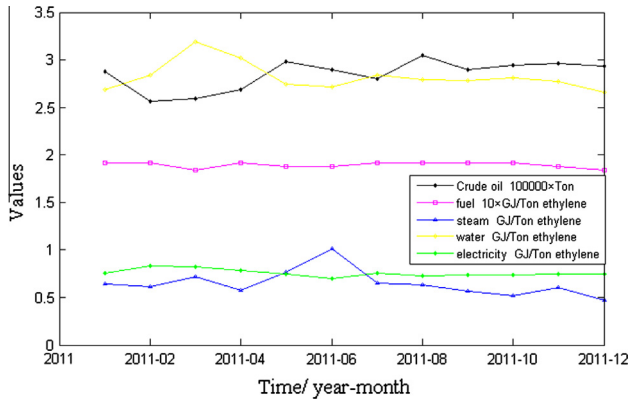

Fig. 5. The required crude oils, fuel, steams, water and electricity of one ethylene plant in 2011.

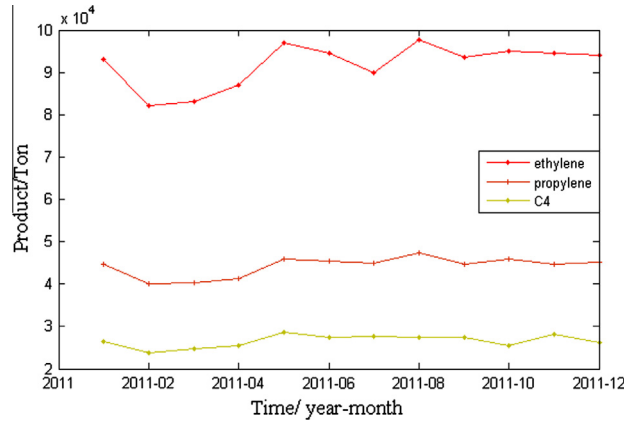

Fig. 6. The monthly productions of ethylene, propylene and C4 of one ethylene plant in 2011.

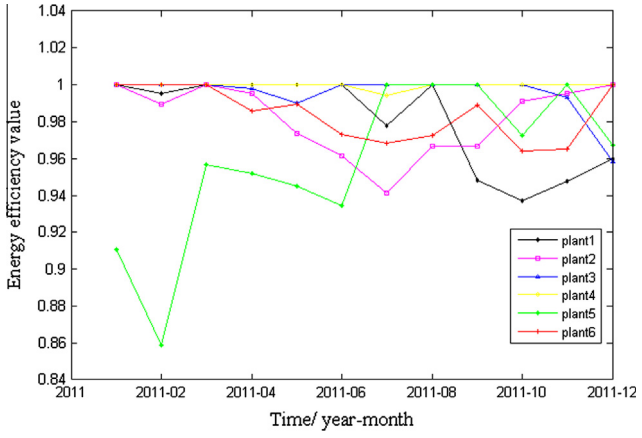

Fig. 7. Energy efficiency values in 2011.

3159.19 tons and 3119.34 tons, and inputs fuel, steams, electricity and water reduce respectively 1.42 GJ 0.02 GJ 0.04 GJ and 0.04 GJ per ton of ethylene, then the efficiency value of the plant 1 can improve the energy efficiency ratio of 2% and reach the effective level. The same analysis of other ineffective ethylene plants can also be obtained by slack variables in Table 3.

Similarly, energy efficiency values of all plants under the order separation technology in 2012 are shown in Fig. 8.

Fig. 8 shows that in 2012, the energy efficiency values of all plants fluctuated, but are all above 0.9. The improved direction for ineffective plants could be found based on slack variables. Compared with energy efficiency values of 2011 in Fig. 7, the DEA efficiency of most plants have increased except for the plant 1, which indicates that the ethylene production performance of each plant had been improved.

The effective improvements for ineffective plants are executed by slack variables. And then we utilize effective units of 2011 as the training data, and those of 2012 as the generalization data. Resources of crude oils, fuel, steams, electricity, and water serve as inputs of the BP model, and productions of ethylene, propylene and C4 are as outputs of the BP model. Then, the BP network is established by setting the step length as 0.7, the momentum factor as 0.3, the training times as 1000 and the maximum allowed error as 0.001. At the same time, initialize the hidden layer nodes as 5 based on Eq. (10). And then the node numbers are increased step by step to get results in Table 4.

It can be shown in Table 4 that the BP model reaches the optimum result when the node is 10. If the node is more or less than 10, the error would both increase. Therefore, the nodes of the hidden layer in the BP network are chosen as 10.

Further, we select the original data (Y-data), the plant data with efficiency valued 1 (O-data), and all optimized effective plant data of 2011 (A-data) as the training data, and those of 2012 as the generalization data. And based on Eq. (10) and Table 4, the hidden layer node number is chosen as 10 and the model training times is 10. The ARGE and the RMSE are demonstrated in Figs. 9 and 10, respectively.

It can be shown in Figs. 9 and 10 that, the ARGE of the Y-data and the O-data are smaller than the ARGE of the A-data. However, the RMSE of the former two are relatively above 7.0% and 6.0%, far bigger than their own ARGE, which would result in the inaccuracy of the model prediction ability. Meanwhile, the RMSE of the A-data is 5.0%, smaller than errors of the other two kinds of data. It indicates the relatively accurate model prediction ability of the A-data, thus proving the effectiveness and accuracy of the proposed method.

On the basis of the above mentioned BP model, the optimized data of 2011 and 2012 is used as the training and generalization data, respectively. And the derived trends of the training errors and the expected errors are shown in Fig. 11.

At the same time, the ratios of predictive outputs and practical outputs of all plants in 2012 are also calculated. For instance, the predictive and practical values of ethylene, propylene and C4 of the plant 1 and the plant 2 are shown in Figs. 12 and 13, respectively.

Seen from the Fig. 12, the predictive values of the DEA-BP model are approximate with the practical values, especially for propylene values. As a whole, the predictive values of each month in 2012 fluctuated little around the practical values, which indicate the steady output level and good production conditions of the plant 1.

In Fig. 13, the predictive values of propylene and C4 of the plant 2 are the same with the practical values, while the predictive values of ethylene have a little discrepancy with its practical ones. And the outputs of ethylene, propylene and C4 in 2012 go down to the lowest till April. Based on the predictive results of the former four months, the input data of May has been adjusted, and the plant production would return back to the normal level since this month. Obviously, the production conditions of the plant 2 are a little worse and less steady than those of the plant 1.

4. Discussion

First, the energy optimization and prediction approach based on the improved DEA-ANN is proposed. Compared with the traditional

Table 3
Energy efficiency values and slack variables of inefficiency plants in 2011.

s^{1-}	s^{2-}	s^{3-}	s^{4-}	s^{5-}	s^{1+}	s^{2+}	s^{3+}	θ	Plant	Time
0.00	1.17	0.18	0.00	0.07	0.00	2410.79	505.71	0.995	Plant1	2011-02
0.00	1.42	0.02	0.04	0.04	3159.19	3119.34	0.00	0.978	Plant1	2011-07
0.00	1.36	0.10	0.00	0.17	0.00	926.49	821.80	0.948	Plant1	2011-09
0.00	0.95	0.13	0.00	0.13	0.00	1317.05	586.66	0.937	Plant1	2011-10
0.00	1.33	0.00	0.00	0.15	0.00	606.60	401.62	0.948	Plant1	2011-11
0.00	1.28	0.00	0.00	0.12	510.16	1515.87	0.00	0.960	Plant1	2011-12
0.00	0.00	0.35	0.09	0.29	0.00	0.00	0.00	0.990	Plant2	2011-02
0.00	0.32	0.57	0.00	0.47	0.00	0.00	0.00	0.974	Plant2	2011-05
0.00	0.84	0.68	0.00	0.39	0.00	0.00	0.00	0.962	Plant2	2011-06
0.00	1.18	0.84	0.00	0.68	0.00	0.00	0.00	0.941	Plant2	2011-07
0.00	0.00	0.33	0.00	0.16	0.00	0.00	0.00	0.967	Plant2	2011-08
0.00	0.00	0.42	0.00	0.39	0.00	0.00	0.00	0.967	Plant2	2011-09
0.00	0.00	0.21	0.00	0.24	1173.08	0.00	0.00	0.992	Plant2	2011-10
0.00	0.00	0.09	0.00	0.13	0.00	0.00	57.22	0.995	Plant2	2011-11
0.00	0.68	0.00	0.00	0.00	0.00	1774.84	1822.84	0.998	Plant3	2011-04
0.00	0.16	0.00	0.00	0.00	0.00	0.00	57.22	0.990	Plant3	2011-05
0.00	0.08	0.00	0.00	0.00	0.00	1649.79	1469.20	0.993	Plant3	2011-11
0.00	0.00	0.00	0.00	0.00	529.14	0.00	170.26	0.958	Plant3	2011-12
0.00	0.00	0.00	0.00	0.00	0.00	36.48	0.00	0.994	Plant4	2011-07
0.00	2.67	0.22	0.00	0.12	0.00	420.09	2075.21	0.911	Plant5	2011-01
0.00	0.42	1.18	0.18	0.13	0.00	373.82	841.07	0.859	Plant5	2011-02
0.00	0.10	0.31	0.00	0.09	0.00	1277.92	1929.24	0.956	Plant5	2011-03
0.00	0.23	0.99	0.00	0.11	0.00	1156.96	2270.94	0.952	Plant5	2011-04
0.00	0.00	0.53	0.07	0.02	0.00	620.68	1060.81	0.945	Plant5	2011-05
0.00	0.27	0.15	0.01	0.25	0.00	0.00	780.69	0.934	Plant5	2011-06
0.00	0.00	1.19	0.00	0.00	0.00	1266.59	1763.49	0.972	Plant5	2011-10
0.00	0.40	1.54	0.00	0.07	0.00	2027.84	3748.18	0.967	Plant5	2011-12
0.00	0.53	0.00	0.00	0.13	0.00	574.82	1067.15	0.986	Plant6	2011-04
0.00	0.13	0.00	0.00	0.21	0.00	564.05	1493.10	0.990	Plant6	2011-05
0.00	1.28	0.00	0.00	0.21	0.00	552.51	1027.07	0.973	Plant6	2011-06
0.00	1.41	0.25	0.00	0.28	0.00	210.72	716.94	0.968	Plant6	2011-07
0.00	1.63	0.00	0.00	0.25	0.00	185.11	660.26	0.972	Plant6	2011-08
0.00	1.28	0.00	0.00	0.21	0.00	944.69	1763.05	0.989	Plant6	2011-09
0.00	1.79	0.22	0.00	0.30	0.00	0.00	1358.73	0.964	Plant6	2011-10
0.00	2.25	0.27	0.00	0.36	0.00	53.14	1224.29	0.965	Plant6	2011-11

Note: the unit of s^{1-}, s^{2-}, s^{3-}, s^{4-}, s^{5-} is GJ/ton ethylene, and the unit of s^{1+}, s^{2+}, s^{3+} is Ton. θ denotes the efficiency value for the ethylene plant.

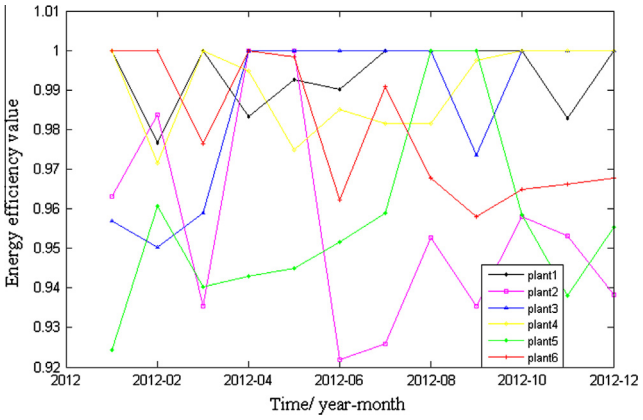

Fig. 8. Energy efficiency values in 2012.

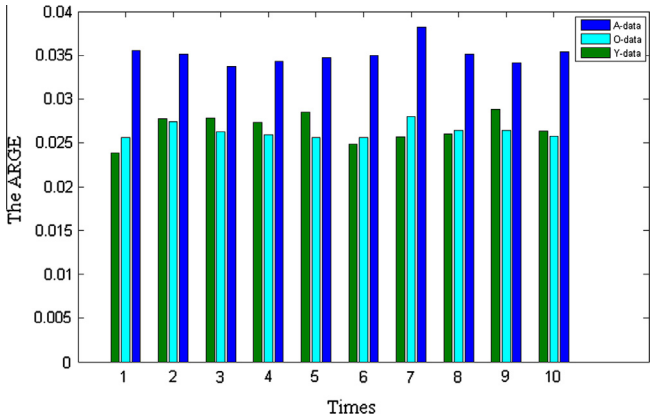

Fig. 9. The ARGE of different data for 10 times.

Table 4
The ARGE and the RMSE of different hidden layer nodes.

Hidden layer nodes	ARGE	RMSE
5	0.03712	0.06369
10	0.03375	0.05184
15	0.03795	0.06929
20	0.03708	0.09808

ANN approach, the robustness and effectiveness of DEA-ANN model is validated through the standard data source from the UCI repository.

Second, the DEA method could help to find the optimized direction of the ineffective production plants, leading plants production to optimal conditions. And making use of the BP network based on the optimized production data, the input-output model of the ethylene plants could be established. The model would serve as an operation guide for the ethylene production process. And it could also be used for predicting the optimal output related to the input data. This proposed optimization modeling method could solve problems like, when using the DEA to evaluate the ethylene production, the data samples are less than three times the number of input and output indexes [18,19], and productions of ethylene, propylene and C4 could not be predicted.

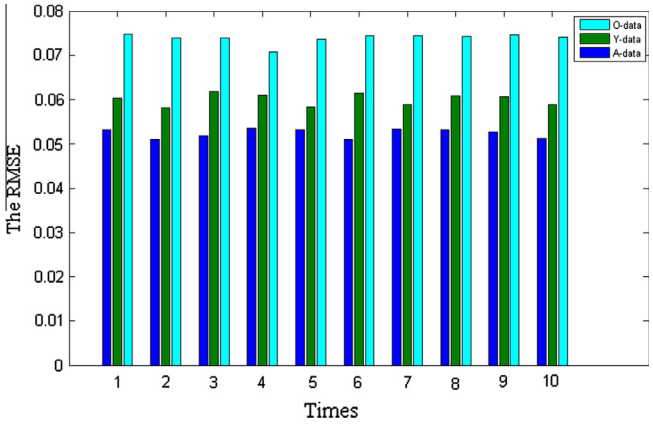

Fig. 10. The RMSE of different data for 10 times.

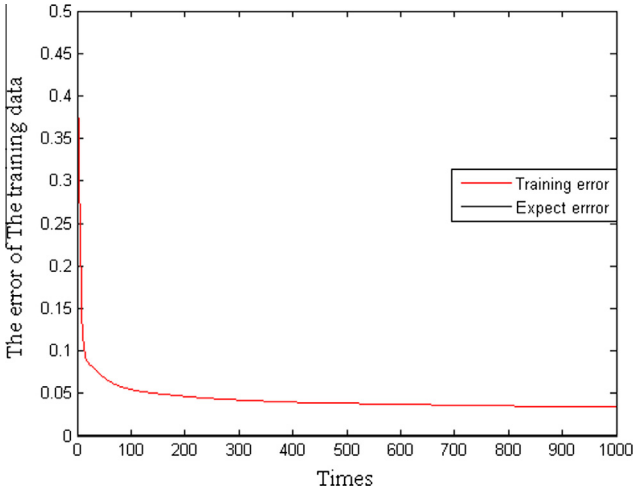

Fig. 11. Trends of the training error and the expected error.

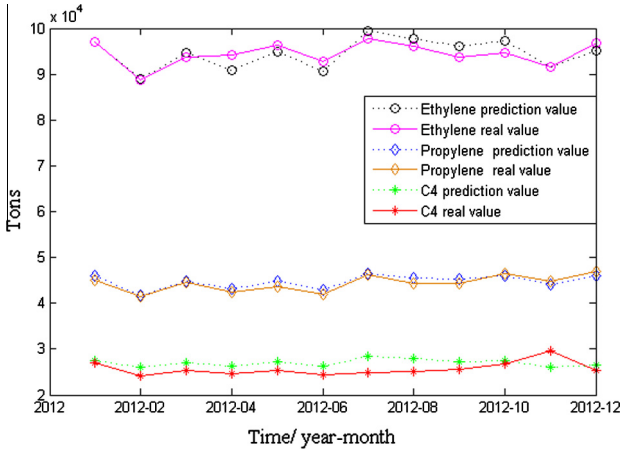

Fig. 12. Predictive and practical values of ethylene, propylene and C4 of the plant 1.

Third, the proposed optimization modeling method is applied to optimize and predict the energy status of ethylene plants effectively, and the parameters of model are adjusted by experiments. Therefore, we will improve our model that parameters can be adjusted automatically, such as designing a self-organizing DEA-ANN model, which is more suitable to the real world application.

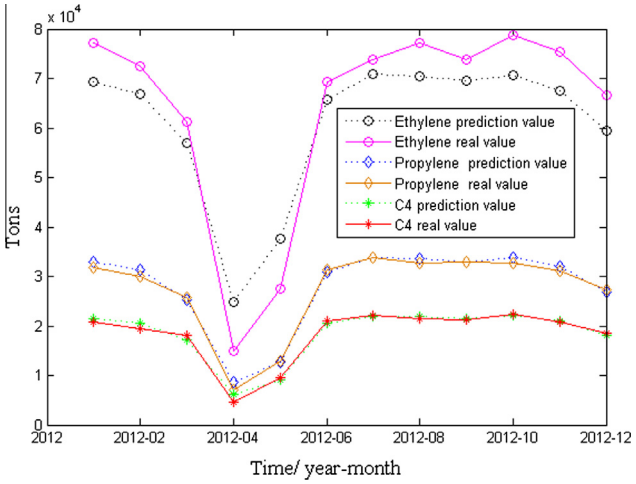

Fig. 13. Predictive and practical values of ethylene, propylene and C4 of the plant 2.

5. Conclusion

This paper proposes a DEA-BP approach. The proposed algorithm could help to build a multi-input-multi-output energy optimization and prediction model and analyze the energy consumption conditions based on the production data. It has better objectivity and resolving power for decision-making. Moreover, in terms of accuracy and training time, the robustness and effectiveness of the DEA-BP model is better than the BP model through the standard data set from the UCI repository. Finally, the proposed method is applied in optimizing and predicting the energy usage of ethylene production plants. The results show that the proposed method could obtain the rational consumption allocation of crude oils, steams, fuel, water and electricity. And it could also get the optimal production situation and improve energy efficiency ratio of about 2% under different input and output configurations reasonably. Meanwhile, the method successfully provides information on energy consumption indexes that decision makers could utilize to improve the energy efficiency effectively. And the information could also serve as the guide for reasonable ethylene production process.

For further research, we will take the effect of economic development and environmental planning on the energy consumption of ethylene industries into account. Moreover, the DEA integrated the ELM (Extreme Learning Machine) would be used to analyze and forecast the scale efficiency and the input-output energy measurement of ethylene production process to compare with current studies. Furthermore, energy usage prediction of thermal power industries could also be analyzed by the self-organizing DEA-ANN model.

Acknowledgment

This research was partly funded by National Natural Science Foundation of China (61533003, 61374166) and the Natural Science Foundation of Beijing, China (4162045).

References

[1] Ma GF, Xu YH, Guo X. Ethylene business review of China petrochemical in 2014. Ethyl Indust 2015;27(1):1–5.
[2] Zhang LJ, Hu J. Ethylene business review of china petroleum in 2014. Ethyl Indust 2015;27(1):6–10.
[3] Tao R, Patel M, Blok K. Olfins from conventional and heavy feedstocks. Energy 2006;31:425–51.
[4] Baileya JA, Gordona R, Burtonb D, Yiridoe EK. Energy conservation on Nova Scotia farms: baseline energy data. Energy 2008;33:1144–54.

[5] Han YM, Geng ZQ, Liu QY. Energy efficiency evaluation based on DEA integrated AHP in ethylene production. Chin J Chem Eng 2014;22(12):1279–84.
[6] Han YM, Geng ZQ. Energy efficiency hierarchy evaluation based on data envelopment analysis and its application in a petrochemical process. Chem Eng Technol 2014;37(12):2085–95.
[7] Dias WO, Wanner EF, Cardoso RTN. A multiobjective optimization approach for combating *Aedes aegypti* using chemical and biological alternated step-size control. Math Biosci 2015;269:37–47.
[8] Grom M, Stavber G, Drnovšek P, Likozar B. Modelling chemical kinetics of a complex reaction network of active pharmaceutical ingredient (API) synthesis with process optimization for benzazepine heterocyclic compound. Chem Eng J 2016;283:703–16.
[9] Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. Eur J Oper Res 1978;6(2):429–44.
[10] Cook WD, Seiford LM. Data envelopment analysis (DEA) thirty years on. Eur J Oper Res 2009;192:1–17.
[11] Sueyoshi T, Goto M, Ueno T. Performance analysis of US coal-fired power plants by measuring three DEA efficiencies. Energy Policy 2010;38(4):1675–88.
[12] Sueyoshi T, Goto M, Sugiyama M. DEA Window Analysis for Environmental Assessment in a Dynamic Time Shift: Performance Assessment of US Coal-fired Power Plants. Energy Econ 2013;40:845–57.
[13] 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.
[14] Sueyoshi T, Goto M. DEA radial measurement for environmental assessment: a comparative study between Japanese chemical and pharmaceutical firms. Appl Energy 2014;115:502–13.
[15] Bi GB, Song W, Zhou P, Liang L. Does environmental regulation affect energy efficiency in China's thermal power generation? Empirical evidence from a slacks-based DEA model. Energy Policy 2014;66:537–46.
[16] Zhu QX, Chen X, He YL, Lin XY, Gu XB. Energy efficiency analysis for ethylene plant based on PCA-DEA. Chin J Chem Eng (China) 2015;66(1):278–83.
[17] Han Y, 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.
[18] Ruggiero J. Impact assessment of input omission on DEA. Int J Inform Technol Decision Making 2005;4(3):359–68.
[19] Cooper WW, Seiford LM, Tone K. Introduction to data envelopment analysis and its uses: with DEA-solver software and references. Springer; 2006.
[20] Cooper WW, Ruiz JL, Sirvent I. Choosing weights from alternative optimal solutions of dual multiplier models in DEA. Eur J Oper Res 2007;180(1):443–58.
[21] Emrouznejad A, Amin GR. DEA models for ratio data: convexity consideration. Appl Math Model 2009;33(1):486–98.
[22] Wei QL. Data envelopment analysis. Chin Sci Bull 2000;45(17):1793–7.
[23] Rumelhart DE, McCelland JL. Parallel distributed processing. Volsland 2: MIT Press; 1986.
[24] Avramovic JM, Velickovic AV, Stamenkovic OS, Rajkovic KM, Veljkovic VB. Optimization of sunflower oil ethanolysis catalyzed by calcium oxide: RSM versus ANN-GA. Energy Convers Manage 2015;105:1149–56.
[25] 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.
[26] Meo S, Zohoori A, Vahedi A. Optimal design of permanent magnet flux switching generator for wind applications via artificial neural network and multi-objective particle swarm optimization hybrid approach. Energy Convers Manage 2016;110:230–9.
[27] McClelland, Rumelhart. Exploration in parallel distributed processing. MIT Press; 1988.
[28] Wang XG, Tang Z, Tamura H, Ishii M. A modified error function for the back-propagation algorithm. Neurocomputing 2004;57:477–84.
[29] Erdogmus D, Fontenla RO, Principe JC, Alonso-Betanzos A, Castillo E. Linear-least-squares initialization of multilayer perceptrons through back-propagation of the desired response. IEEE Trans Neural Networks 2005;16(2):325–37.
[30] Dai Q, Ma ZC, Xie QY. A two-phased and ensemble scheme integrated back-propagation algorithm. Appl Soft Comput 2014;24:1124–35.
[31] Sun JY. Learning algorithm and hidden node selection scheme for local coupled feedforward neural network classifier. Neurocomputing 2012;79:158–63.
[32] Anam S, Md Islam S, Kashem MA, Islam MN, Islam MR, Islam MS. Face recognition using genetic algorithm and back propagation neural network. In: Proceedings of the international multiconference of engineers and computer scientists 2009, vol. I. Hong Kong: IMECS; 2009.
[33] Wu WW. Beyond travel & tourism competitiveness ranking using DEA, GST, ANN and Borda count. Expert Syst Appl 2011;38:12974–82.
[34] Olanrewaju OA, Jimoh AA, Kholopane PA. Integrated IDA-ANN-DEA for assessment and optimization of energy consumption in industrial sectors. Energy 2012;46:629–35.
[35] Olanrewaju OA, Jimoh AA. Review of energy models to the development of an efficient industrial energy model. Renew Sustain Energy Rev 2014;30:661–71.
[36] Charnes A, Cooper WW, Wei QL, Gong LG, Sun DB, Huang ZM. Computability and economic interpretation for "cone-ratio" DEA model. The University of Texas at Austin, Center for Cybernetic Studies Report; 1987.
[37] MirHassani SA, Alirezaee MR. An efficient approach for computing non-Archimedean ε in DEA based on integrated models. Appl Math Comput 2005;166(2):449–56.
[38] Alirezaee MR, Khalili M. Recognizing the efficiency, weak efficiency and inefficiency of DMUs with an epsilon independent linear program. Appl Math Comput 2006;183(2):1323–7.
[39] Zhang BC, Wang WL, Lin WF, Du G, Wu YH. Review on DEA models involving the non-Archimedean infinitesimal. J Syst Eng 2010;25(3):407–14.
[40] Lippmann RP. An introduction to computing with neural nets. IEEE ASSP Mag 1987;4:4–22.
[41] Hecht-Nielsen R. Theory of the back-propagation neural network. Int Joint Conf Neural Networks 1989;1:593–605.
[42] Gao DQ. On structures of supervised linear basis function feedforward three-layered neural networks. Chin J Comput 1998;21(1):80–5.
[43] González-Romera E, Jaramillo-Morán MA, Carmona-Fernández D. Monthly electric energy demand forecasting with neural networks and Fourier series. Energy Convers Manage 2008;49:3135–42.
[44] Xu LJ, Davim JP, Cardoso R. Prediction on tribological behaviour of composite PEEK-CF30 using artificial neural networks. J Mater Process Technol 2007;189(1–3):374–8.
[45] 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;51:8526–34.
[46] China Standards: the general computing guide of special energy consumption (GB/T2589-2008); 2008.
[47] China Standards: the limitation of energy consumption for ethylene product (DB37/751-2007); 2008.
[48] Zhang H, Tong TS, Zhang F. Consensus measurement in micro-inertial sensors. J Transducer Technol 2001;10:40–1.
[49] Calculation method for energy consumption in petrochemical engineering design (SH/T3110-2001); 2002.