Applied Thermal Engineering 115 (2017) 280–291

Contents lists available at ScienceDirect

Applied Thermal Engineering

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

Research Paper

Energy and carbon emissions analysis and prediction of complex petrochemical systems based on an improved extreme learning machine integrated interpretative structural model

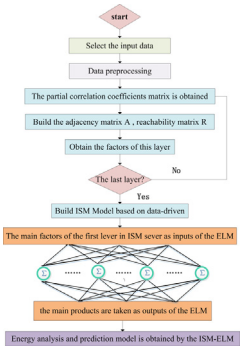
Yongming Han, Qunxiong Zhu, Zhiqiang Geng *, Yuan Xu *

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

HIGHLIGHTS

- The ELM integrated ISM (ISM-ELM) method is proposed.
- The proposed method is more efficient and accurate than the ELM through the UCI data set.
- Energy and carbon emissions analysis and prediction of petrochemical industries based ISM-ELM is obtained.
- The proposed method is valid in improving energy efficiency and reducing carbon emissions of ethylene plants.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 16 September 2016
Revised 15 December 2016
Accepted 19 December 2016
Available online 21 December 2016

Keywords:
Energy saving
Carbon emissions reduction
Extreme learning machine
Interpretative structural model
Complex petrochemical systems

ABSTRACT

Energy saving and carbon emissions reduction of the petrochemical industry are affected by many factors. Thus, it is difficult to analyze and optimize the energy of complex petrochemical systems accurately. This paper proposes an energy and carbon emissions analysis and prediction approach based on an improved extreme learning machine (ELM) integrated interpretative structural model (ISM) (ISM-ELM). ISM based the partial correlation coefficient is utilized to analyze key parameters that affect the energy and carbon emissions of the complex petrochemical system, and can denoise and reduce dimensions of data to decrease the training time and errors of the ELM prediction model. Meanwhile, in terms of the model accuracy and the training time, the robustness and effectiveness of the ISM-ELM model are better than the ELM through standard data sets from the University of California Irvine (UCI) repository. Moreover, a multi-inputs and single-output (MISO) model of energy and carbon emissions of complex ethylene systems is established based on the ISM-ELM. Finally, detailed analyses and simulations using

Abbreviations: ELM, extreme learning machine; ISM, interpretative structural model; UCI, University of California Irvine; MISO, multi-inputs and single-output; DEA, data envelopment analysis; AHP, analytic hierarchy process; SLFNs, single-hidden-layer feedforward neural networks; ARGE, average relative generalization error; RMSE, root mean square error; $NetOut_i^{inver}$, the inverse normalized network output for all samples; $ExpectOut_i^{inver}$, the expected output; $NetOut_i^{norm}$, the normalized network output for all samples; NC, Naphtha; DC, light diesel oil; HC, hydrogenation tail oil; RC, raffinate; C_{345}, carbon3, carbon4, carbon5; LC, light hydrogenation tail oil; OC, other force; BW, boiler feed water; IW, industrial water; RW, recycled water; E, electricity; LS, low pressure steam; MS, medium pressure steam; HS, high pressure steam; SS, ultra-high pressure steam; HF, heavy oil; LF, light oil; FF, fuel gas; EL, ethylene; PL, propylene; n, the number of samples; k, the number of features; m, the number of outputs; O, the hidden nodes; δ, the weight matrix between the input and output layer nodes; B, the threshold of neurons in hidden layer; β_d, the weight matrix between the hidden and output layer; H^+, the Moore-Penrose generalized matrix; x_i, the i-th parameters of a data set; y_j, the j-th parameters of a data set; c, the inverse correlation matrix; A, the adjacency matrix; R, the reachability matrix; S_i, the reachable set; B_j, the first set; L_i, the i-th layer.

* 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), xyfancy@163.com (Y. Xu).

http://dx.doi.org/10.1016/j.applthermaleng.2016.12.081
1359-4311/

the real ethylene plant data demonstrate the effectiveness of the ISM-ELM and can guide the improvement direction of energy saving and carbon emissions reduction in complex petrochemical systems.

1. Introduction

Nowadays, all countries in the world focus on energy saving and emission reduction. Particularly in the complex petrochemical process, energy saving and carbon emissions reduction is an important means of reaching the economic and environmental goals. In addition, the ethylene industrial level plays a key role in estimating the industrial development level of a country. Based on the statistic, the ethylene production in China was less than that in United States and ranked second in the world [1], and the standard oil (average fuel plus power consumption) in China was about 616.7 kg per ton for producing a ton of ethylene in 2012 [2]. But the standard oil of the advanced countries was 440.2 kg for producing a ton of ethylene in 2012 [3]. Therefore, the energy efficiency of China was far lower. Meanwhile, over half of the operation cost in ethylene production plants is occupied by the energy consumption cost [4], and the carbon emissions of China accounted for 20.85% of carbon emissions in the total world [5]. During the 13th five-year plan period, energy conservation and emission reduction of petrochemical industries and eco-friendly development are the path China must take to achieve prosperity and development. Consequently, studying the energy saving and carbon emissions analysis of petrochemical industries is beneficial for the sustainable development and the environment and economy of China.

Generally, the mean method and the optimal index method are used to analyze the energy status [6]. But these methods cannot introduce energy optimization knowledge to guide the improvement direction of energy saving. Energy efficiency estimation of ethylene systems based data fusion methods has obtained better results, but these methods did not take the impact indicators of energy saving and carbon emissions into consideration [7,8]. Chen et al. [9–16] analyzed the characteristics and increased the energy efficiency of the iron and steel production process and system by optimizing the recovery and utilization of residual energy and heat, distributing the sintering proportioning, recovering residual heat and decreasing the exergy loss, etc. However, the economic cost of reforming production plants was not taken into account. Additionally, using the data envelopment analysis (DEA) based on analytic hierarchy process (AHP) and DEA-cross model, Geng et al. [17,18] analyze the energy status of ethylene processes and chiller systems. However, the efficiency discrimination of the DEA model is poor when many efficiency values of decision making units are 1 [19,20]. In allusion to the disadvantages existing in energy saving and estimation of the petrochemical industry, this paper presents an improved extreme learning machine (ELM) integrated interpretative structural model (ISM) (ISM-ELM) to analyze and predict energy and carbon emissions for the petrochemical industry.

Artificial neural networks have significant accuracy to solve complex mathematical and industrial problems which consists of multiple nonlinear variables, and multiple advantages. In 2006, Huang [21] first proposed a novel learning algorithm named extreme learning machine (ELM) for single-hidden-layer feedforward neural networks (SLFNs). The only parameter of the ELM needed to be defined is the number of the hidden node. But the input weights and biases of hidden nodes including are randomly assigned and need not be tuned. And the output weights can be analytically confirmed by the simple generalized inverse operation. Compared with other traditional learning algorithms such as SLFNs, the ELM provides extremely faster learning speed and better generalization performance. Therefore, the ELM is widely applied in analysis and prediction of objects like mortality of bladder cancer, real-time fault diagnosis, dynamic voltage stability status, wind power density, etc. [22–25]. Wong et al. [26] studied advanced machine learning methods to confirm the optimal biodiesel ratio that can achieve the goal of fewer emissions. Salcedo-Sanz et al. [27] successfully solved this problem of feature selection in short-term wind speed forecasting using a hybrid coral reefs optimal ELM method. Na et al. [28] proposed a single well production prediction model to improve the precision of oilfield production prediction using an improved ELM method. Sun et al. [29] applied the ELM model integrating the linear programming for analyzing the port oil transportation system with five and eight operation states. Ning et al. [30] studied two effective ELM methods with prediction interval to predict the temperature of slabs based on process parameters in the steel refining and casting process. However, the ELM has a problem with the validity of prediction [31]. An ISM is exploited to obtain the main factors among input variables to improve the reliability and the prediction accuracy.

In the practical prediction, energy saving of the petrochemical industry is affected by many indexes. And each index has a certain amount of information. Meanwhile, the relationship among indexes influences each other. If the ELM algorithm is directly used for energy prediction, the phase of overlapping information will reduce the running rate, and the effect will not be ideal. Therefore, the ISM-ELM method is presented to analyze and predict the energy saving and carbon emissions of the petrochemical industry.

The ISM developed by Warfield has emerged as a promising technique for handling high-dimensional data of the complex system, with increasing application in decision support units. Kuo and Faisal et al. [32–35] divided these barriers into a hierarchical structure of the quality management, the product service system and the logistics provider using the ISM method. Geng et al. [36] used DEA integrated ISM to estimate the energy status of ethylene industries in China. The experiment results mentioned above showed the practicality and effectiveness of the ISM. Generally, in order to establish the ISM, the initial step is to create an adjacency matrix mostly based the expert experience. But its shortcoming is subjective and inconsistent. While Yu et al. [37,38] discovered that the partial correlation function was more realistic to reflect the correlation relationship among variables in the time series data analysis. Therefore, we used the adjacency matrix of correlation coefficients and partial correlation coefficients to build the reachability matrix and obtain the ISM model.

First, main factors can be obtained by analyzing the standard data set from the University of California Irvine (UCI) repository. Meanwhile, in terms of the model accuracy and training time, the robustness and effectiveness of the ISM-ELM model is better than ELM. Moreover, the ethylene product system in the petrochemical industry is a complex industrial system, so the ISM can be utilized to construct a hierarchy model based on the energy data. And it can easily use accessible monthly operational data to further obtain the key parameters and basic reasons and reduce the redundant features to shorten the prediction identification time and improve the accuracy. Meanwhile, considering the

process knowledge and mechanism model, the complexity of modeling process of the ethylene production system can be avoided based on the ISM-ELM. Furthermore, the ISM-ELM based on key parameters can be used to analyze and predict energy status and carbon emissions in petrochemical industries reasonably, and provide the operation guidance for energy saving and carbon emissions reduction.

The structure of the paper is organized as follows. Section 1 introduces the research status of energy analysis and prediction in the petrochemical industry with the ELM and the ISM. Section 2 provides the details of the ELM methods. The energy and carbon emissions analysis and prediction framework based on ISM-ELM in the ethylene production industry is described in Section 3. Section 4 shows the robustness and effectiveness of the ISM-ELM model compared the ELM through standard data sets from the UCI repository. Section 5 presents an application about energy and carbon emissions analysis and prediction of the ethylene production system based on the ISM-ELM. Finally, Discussion and Conclusion are obtained in Section 6 and Section 7, respectively.

2. Extreme learning machine

Given $X = \begin{bmatrix} x_{11} & \cdots & x_{1k} \\ \cdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{nk} \end{bmatrix}_{n*k}$ and $Y = \begin{pmatrix} Y_1 \\ \cdots \\ Y_1 \end{pmatrix}_{n*m} = \begin{bmatrix} Y_{11} & \cdots & Y_{1m} \\ \cdots & \cdots & \cdots \\ Y_{n1} & \cdots & Y_{nm} \end{bmatrix}_{n*m}$ are inputs and outputs of n samples, respectively. And for each sample, k and m are the number of features and outputs, respectively. Where $X_i = (x_{i1}, x_{i2}, \ldots\ldots, x_{ip}) \in R^k$ and $Y_i = (Y_{i1}, Y_{i2}, \ldots\ldots, Y_{im}) \in R^m$ for n arbitrary distinct samples (X_i, Y_i). Setting the hidden nodes to O, and the standard SLFNs with activation function $G(x)$ are mathematically modeled as Eqs. (1) and (2).

$$\mathrm{H} = G(X * \delta + B) = \begin{bmatrix} G(X_1 * \delta_1 + B_1) & \cdots & G(X_1 * \delta_O + B_1) \\ \cdots & \cdots & \cdots \\ G(X_n * \delta_1 + B_n) & \cdots & G(X_n * \delta_O + B_n) \end{bmatrix}_{n*O} \tag{1}$$

$$Y = H * \eta = \begin{pmatrix} Y_1 \\ \cdots \\ Y_n \end{pmatrix}_{n*m} = \begin{bmatrix} y_{11} & \cdots & y_{1m} \\ \cdots & \cdots & \cdots \\ y_{n1} & \cdots & y_{nm} \end{bmatrix}_{n*m} \tag{2}$$

wherein $\delta = (\delta_1, \delta_2, \ldots\ldots, \delta_O)_{k*O}$ is the weight matrix between the input and output layer nodes, Where $\delta_i = (\delta_{i1}, \delta_{i2}, \ldots\ldots, \delta_{ik})^T, (i = 1, 2, \ldots\ldots, O)$. Meanwhile, $B = (\beta_d, \beta_d, \ldots\ldots, \beta_d)^T$ is the threshold of neurons in hidden layer,$\beta_d = (\beta_{d1}, \beta_{d2}, \ldots\ldots, \beta_{dO}), (d = 1, 2, \ldots\ldots, n)$.
$\eta = (\eta_1, \eta_2, \ldots\ldots, \eta_O)^T$ is the weight matrix between the hidden and output layer, and $\eta_i = (\eta_{i1}, \eta_{i2}, \ldots\ldots, \eta_{im}), (i = 1, 2, \ldots\ldots, O)$. Note that the operator '•' is the multiplication of two matrixes.

In general, the excitation function $G(x)$ is designed as Sigmoid, Sine or RBF [39,40]. In this paper, we use the sigmoid function in Eq. (3).

$$G(x) = \frac{1}{1 + e^{-x}} \tag{3}$$

In order to approximate the error of these n samples to zero error as $\sum_{i=1}^{n}(\widetilde{Y}_i - Y_i) = 0$, where $\widetilde{Y} = \begin{pmatrix} \widetilde{Y}_1 \\ \cdots \\ \widetilde{Y}_n \end{pmatrix}_{n*m} = \begin{bmatrix} \widetilde{Y}_{11} & \cdots & \widetilde{Y}_{1m} \\ \cdots & \cdots & \cdots \\ \widetilde{Y}_{n1} & \cdots & \widetilde{Y}_{nm} \end{bmatrix}_{n*m}$ are corresponding target outputs of n samples. Therefore, the weight matrix between the hidden layer and output layer is shown in Eq. (4).

$$\eta = H^+ \bullet \widetilde{Y} \tag{4}$$

where H^+ is the Moore-Penrose generalized matrix of H. The structure chart of the ELM is described in Fig. 1.

3. An energy analysis and prediction framework based on ELM integrating ISM

When the network nodes of the hidden layer too more, the network training errors become fewer, and the prediction accuracy will be improved, but then the iteration times of the network increase, the network will become more complicated. While the network nodes of the hidden layer are too little, the network will not be trained to training, or the prediction accuracy is poor. Therefore, in order to determine a reasonable number of hidden layer nodes, we need to analyze the complexity and errors of the network structure by the ISM.

3.1. Data preprocessing

Multidimensional data has complex nonlinear timing relationship, including the noise and abnormal data, etc. And dimension for variables of the timing data is generally different, so we used the general transformation method named the proportion to make values of variables incomparable [8]. However, the described subjects of the timing data need to be taken into account. Generally, different variables make different interactions for the same subject [41]. Some of variables are positive and other variables are negative, which their formulas are shown in Eqs. (5) and (6), respectively.

$$x'_{ij} = \frac{x_{ij} - x_j^{\min}}{x_j^{\max} - x_j^{\min}} \tag{5}$$

$$x'_{ij} = 1 - \frac{x_{ij} - x_j^{\min}}{x_j^{\max} - x_j^{\min}} = \frac{x_j^{\max} - x_{ij}}{x_j^{\max} - x_j^{\min}} \tag{6}$$

in which $x_j^{\max} = \max\{x_{1j}, x_{2j}, \ldots, x_{tj}\}$, $x_j^{\min} = \min\{x_{1j}, x_{2j}, \ldots, x_{tj}\}$, $i = 1, 2, \ldots, t$; $j = 1, 2, \ldots, m$. The inverse normalization of variables in the network is taken in Eqs. (7) and (8).

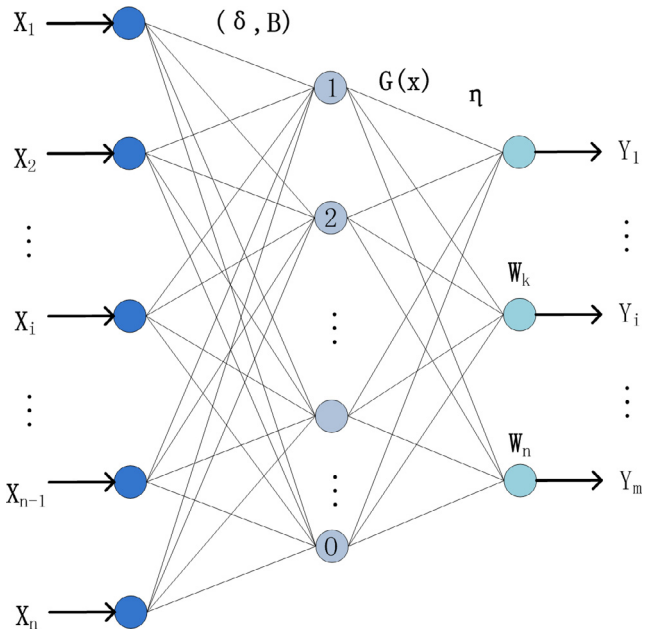


Fig. 1. The structure chart of the ELM.

$$x_{ij} = x'_{ij} \times \left(x_j^{\max} - x_j^{\min}\right) + x_j^{\min} \tag{7}$$

$$x_{ij} = x_j^{\max} - x'_{ij} \times \left(x_j^{\max} - x_j^{\min}\right) \tag{8}$$

3.2. The ISM method

The partial coefficient analysis can remove the influences of rest irrelevant parameters and obtain the real relationships among different elements. Meanwhile, the procedure based on data-driven analysis of obtaining the adjacency matrix will be objective and consistent, thus it can be systematically to build the ISM.

In practical, the correlation coefficient only takes the relationship between two elements into consideration, so it is difficult to determine directly the inner relationship between elements. However, because of deduct or fix the effect of other variables beside the relationship between two elements, the partial correlation relationship can reflect the real linkage of different elements [42]. If the absolute value of the partial correlation coefficient is greater, the relationship between both variables will be stronger. Therefore, due to other things being equal influence, the partial correlation coefficient reflects the real correlation between the dependent and independent elements, which value is between −1 and 1 without the unit.

Let x_i and y_i ($i = 1, 2, \ldots, m$) is the i-th and j-th parameters of a data set, then the correlation coefficient between them can be calculated by Eq. (9).

$$r_{xy} = \frac{\sum_{i=1}^{M}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{M}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{M}(y_i - \bar{y})^2}} \tag{9}$$

where $\bar{x} = \frac{x_1+x_2+\cdots+x_m}{m}$ and $\bar{y} = \frac{y_1+y_2+\cdots+y_m}{m}$, respectively.

And then we can obtain the correlation coefficient matrix based on the Eq. (10).

$$r = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n-1} & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n-1} & r_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{m1} & r_{m2} & \cdots & r_{m-1n-1} & r_{mn} \end{bmatrix}_{m*n} \tag{10}$$

We get the inverse correlation matrix c which can be obtained by the correlation coefficient matrix in Eq. (11).

$$c = inv(r) = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n-1} & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n-1} & c_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{m1} & c_{m2} & \cdots & c_{m-1n-1} & c_{mn} \end{bmatrix}_{m*n} \tag{11}$$

$inv(r)$ denotes the inverse correlation matrix of r, Meanwhile, the partial correlation coefficient between two variables is calculated as:

$$R_{ij} = -\frac{c_{ij}}{\sqrt{c_{ii} * c_{jj}}} \tag{12}$$

where $i = 1, 2, \ldots, m; j = 1, 2, \ldots, m$. Finally, the partial correlation coefficient matrix can be obtained by the inverse correlation matrix.

$$P = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ & \cdots & \cdots & \\ P_{m1} & P_{m2} & \cdots & P_{mn} \end{bmatrix}_{m*n} \tag{13}$$

Due to the different definition, the partial correlation coefficient is related or not in different industries. Generally, according to the scope of partial correlation coefficients [43], we can obtain the relationship as shown in Table 1.

If R_{ij} is a positive number and bigger than the threshold value, then $a_{ij} = 1$ and $a_{ji} = 0$. Otherwise, $a_{ij} = 0$ and $a_{ji} = 1$. The adjacency matrix A is as follows:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n-1} & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{n-1n-1} & a_{nn} \end{bmatrix}_{n*n} \tag{14}$$

Suppose the n-order identity matrix as

$$U = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}_{n*n} \tag{15}$$

If the equation $A + U = (A + U)^2 = \cdots = (A + U)^{n-1} = (A + U)^n$ exits, then $R = (A + U)^{n-1}$, which is the reachability matrix R by transferring the adjacency matrix A.

$$R = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1n-1} & R_{1n} \\ R_{21} & R_{22} & \cdots & R_{2n-1} & R_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ R_{n1} & R_{n2} & \cdots & R_{n-1n-1} & R_{nn} \end{bmatrix}_{n*n} \tag{16}$$

In the ith row, if $R_{ij} = 1 (j = 1, 2, \ldots, n)$, then the element R_{ij} is put into the reachable set S_i. Meanwhile, in the jth column, if $R_{ij} = 1$ ($i = 1, 2, \ldots, n$), then the element R_{ij} is put into the first set B_j.

If $S_j \cap B_j = S_j$, the influence elements can be stratified, and then the highest impact element of the first layer L_1 is identified. The row and column corresponding in L_1 are removed from the reachability matrix R. Based on the same decision rule, we can obtain the rest factors of the other layers L_2 $L_3 \ldots L_k$ until identify all of the factors at each level. The hierarchy structure of the ISM can be established based on each level of L.

3.3. The energy analysis and prediction process based on ISM-ELM

Step 1: Select the input factors, and carry on data normalization by Eqs. (5) and (6).
Step 2: the correlation coefficients matrix r is obtained by Eqs. (9) and (10).
Step 3: the partial correlation coefficients matrix P is established by using the inverse matrix of r based on Eqs. (11) and (12).
Step 4: the adjacency matrix A is transformed to the reachability matrix R based on Eqs. (13)–(15) and the correlation coefficient thresholds as shown in Table 1.
Step 5: The advanced set B_j and the reachable set S_i of the reachability matrix R, and the first level of elements can be obtained based the ISM.
Step 6: The corresponding column and row of the elements are removed, which included in the first level of the reachability matrix R, and then we rebuild the reachability matrix again, and run back to step 2. If there is no element in reachability matrix, then the loop will not stop.

Table 1
The scope of partial correlation coefficient corresponding to relationships.

The scope of partial correlation coefficient	Relationship between both variables
$0 \leqslant \lvert R_{ij}\rvert < 0.1$	No relationship
$0.1 \leqslant \lvert R_{ij}\rvert < 0.3$	Low correlation
$0.3 \leqslant \lvert R_{ij}\rvert < 0.5$	Medium correlation
$0.5 \leqslant \lvert R_{ij}\rvert < 0.8$	Strong correlation
$0.8 \leqslant \lvert R_{ij}\rvert \leqslant 1$	Extremely strong

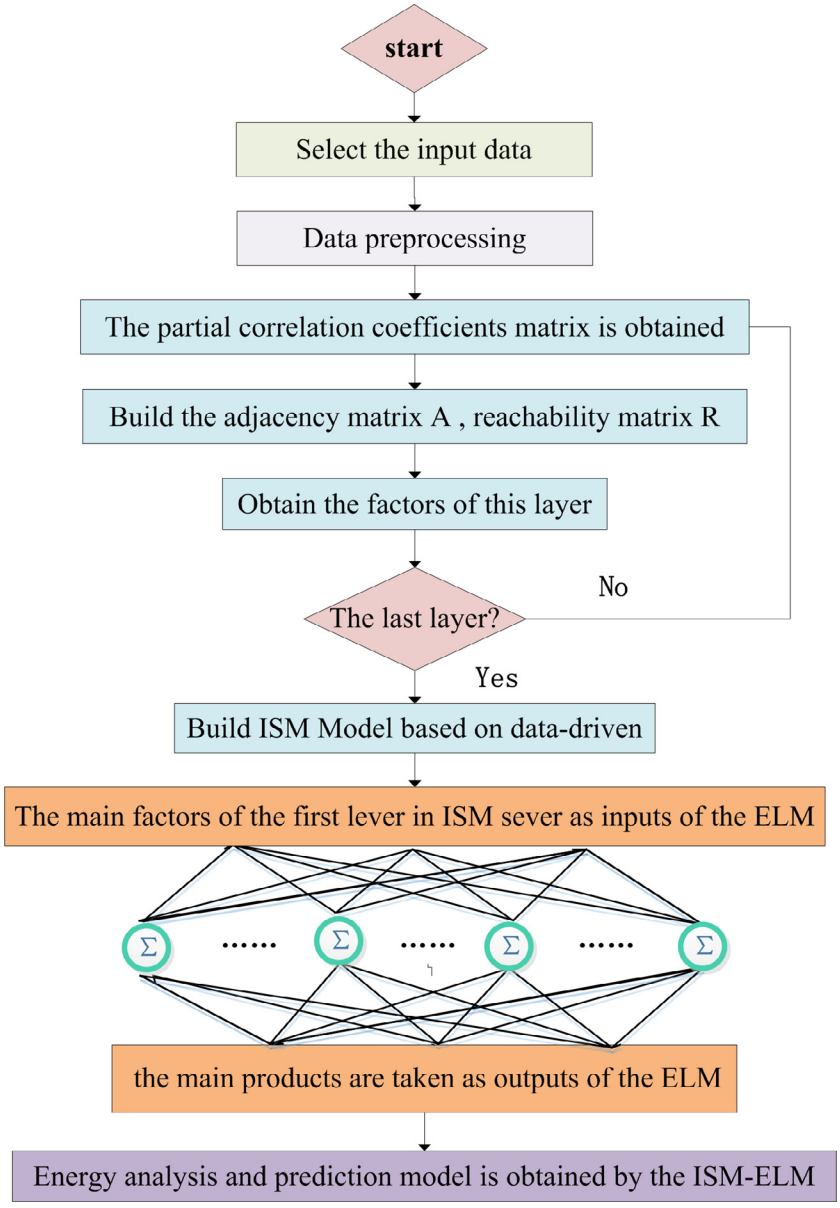


Fig. 2. The energy analysis and prediction framework based on ELM integrating ISM.

Step 7: The ISM based on the elements of each level is built.
Step 8: The main factors are obtained by the ISM and the main productions as the inputs and outputs of the ELM, respectively.
Step 9: Select the related parameters of the ELM, including the hidden layer nodes and the excitation function to establish the energy prediction model using Eqs. (1)–(4).
Step 10: Energy analysis and prediction based on the ELM network.

Diagrammatically, the energy analysis and prediction framework based on ELM integrating ISM with data-driven is described in Fig. 2.

4. Algorithm comparison based on standard data sets

In order to validate the effectiveness and robustness of the ISM-ELM method, the paper uses some standard data sets from UCI repository [44] to compare the ISM-ELM with the ELM algorithm. The specification of these standard data sets is listed in Table 2.

We build the ELM and the ISM-ELM network models by setting a learning factor as 0.01, 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. (17) and (18), respectively.

$$\begin{cases} \begin{cases} RGE_i = Abs\left(\frac{NetOut_i^{inver} - ExpectOut_i^{inver}}{ExpectOut_i^{inver}}\right), \ if \ ExpectOut_i^{inver} \neq 0 \\ RGE_i = Abs(NetOut_i^{inver} - ExpectOut_i^{inver}), \ else \end{cases} \\ ARGE = \frac{\sum_{i=1}^{n} RE_i}{n} \end{cases} \tag{17}$$

$$\begin{cases} MSE_i = \left(NetOut_i^{norm} - ExpectOut_i^{norm}\right)^2 \\ RMSE = \sqrt{\frac{\sum_{i=1}^{n} MSE_i}{n}} \end{cases} \tag{18}$$

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

Table 2
Specification of standard data sets.

Data sets	#Samples		#Attributes	
	Training	Testing	Inputs	Outputs
Wine	1000	599	11	1
CMC	1000	473	9	1
Zoo	68	33	16	1

Table 3
Comparison of the ARGE and RMSE of ELM and ISM-ELM.

Data Sets	Nodes	ELM			ISM-ELM			
		ARGE	RMSE	Time	ARGE	RMSE	Time	MFs
Wine	2	0.1252	0.3328	0.0049	0.1227	0.3283	0.0044	20
	3	0.1208	0.3265	0.0058	0.1216	0.3261	0.0046	
	5	0.1207	0.3282	0.0066	0.1213	0.3258	0.0047	
CMC	2	0.4047	0.8737	0.0050	0.4001	0.7414	0.0047	10
	3	0.3998	0.8689	0.0045	0.3970	0.7436	0.0035	
	5	0.4008	0.8686	0.0055	0.3963	0.7452	0.0054	
Zoo	2	0.7229	0.8288	0.0036	0.6684	0.7471	0.0035	7
	3	0.6858	0.7292	0.0040	0.5888	0.7175	0.0033	
	5	0.7798	0.7092	0.0046	0.6052	0.6938	0.0035	

Based on the relationship scope of partial correlation coefficient as shown in Table 1, we obtain that if the threshold value of relationships is set to 0.4 of these UCI data sets, the variables are relation. Therefore, we obtain 8 main factors of Wine data set, 7 main factors of CMC data set and 10 main factors of Zoo data set based on the ISM. Meanwhile, we increase the nodes of ELM network and perform the experiment again. The ARGE and the RMSE are listed in Table 3.

It can be seen from Table 3 that the training time, the ARGE and the RMSE of the ISM-ELM model are the best when the hidden nodes are set as 2, 3 and 5 for CMC and Zoo data sets. The ARGE and the RMSE of ELM are close to that of ISM-ELM when nodes of hidden layer decrease, however, the computing time is longer along with increasing hidden nodes. Meanwhile, we set the hidden nodes of ISM-ELM model as 2, 3 and 5 for Wine data sets, the training time and the RMSE of the ISM-ELM model are better than those of the ISM, but the ARGE of the ISM-ELM is bigger and is close to that of ELM when nodes of hidden layer increase. According to Table 3, the effectiveness of the ISM-ELM model is validated.

5. Case study: Energy and carbon emissions analysis and prediction of Chinese ethylene production industries

In Chinese ethylene product industries, there are about seven common process technologies [41]. In this article, the actual ethylene production data of ethylene production scales with from 200,000 to 800,000 and over 800,000 under different technologies are taken as a case study to illustrate the effectiveness and robustness of the ISM-ELM method.

Ethylene production plants are a key member of fossil fuel consumed energy sources that can be divided into two parts: the cracking section and the separation section. The energy consumption of the separation section mainly includes a rapid cooling, a compression and a separation part. And 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 after-heat in the flue gas. A schematic flow diagram of an ethylene plant is shown in Fig. 3 as follows.

5.1. Energy data analysis of the ethylene production process

Based on the statistics, the energy consumption fees occupy over 50 percent of overall cost in the ethylene production process. The main factor which affects the whole energy consumption of ethylene plants is the quality of crude oils, the second is the adopted cracking technology, the third is the separation process, and the last are the periphery project and the public project [45]. Over 70 percent of the whole cost in ethylene production is occupied by crude oils (raffinate, carbon3, carbon4, carbon5, light diesel oil, Naphtha, hydrogenation tail oil and other materials). The ethylene utilization boundary included the main energy types listed as follows [46]: crude including the Naphtha (NC), the light diesel oil (DC), the hydrogenation tail oil (HC), the Raffinate (RC), the carbon3, carbon4, carbon5 (C_{345}), the light hydrogenation tail oil (LC) and other force (OC); the water including the boiler feed water (BW), the industrial water (IW) and the recycled water (RW); the electricity (E); steams including the low pressure steam (LS), the medium pressure steam (MS), the high pressure steam (HS)and the ultra-high pressure steam (SS); fuels including the heavy oil (HF), the light oil (LF) and the fuel gas (FF); the compressing air and N_2. Due to the lowest consumption of the compressing air and N_2 in energy types, we did not taken them into consideration when analyzing and predicting energy status of the ethylene product process. Thus the crude oils and energy types serve as the inputs and the productions of ethylene (EL), propylene (PL) and C_4 are taken as the outputs [17,18,41].

Ethylene production data have complex nonlinear timing relationship including abnormal and noise data. And dimension for variables of the timing data is usually different, so we used the general transformation method named the proportion to dispose values of variables incomparable [8]. Meanwhile, based on the analysis of the DEA model above, the efficiency value of the DEA is relative. And the accuracy of the estimation result is affected by multiple input and output variables of ethylene plants and their precision. Therefore, normalized, uniform dimension disposal of units and consistency test are used to preprocess the data. The following Eq. (19) will lead to erroneous judgment for the short-length data, but it is well applied to dispose the long-length data of ethylene plants, and [8]. And the long-length data can be pro-

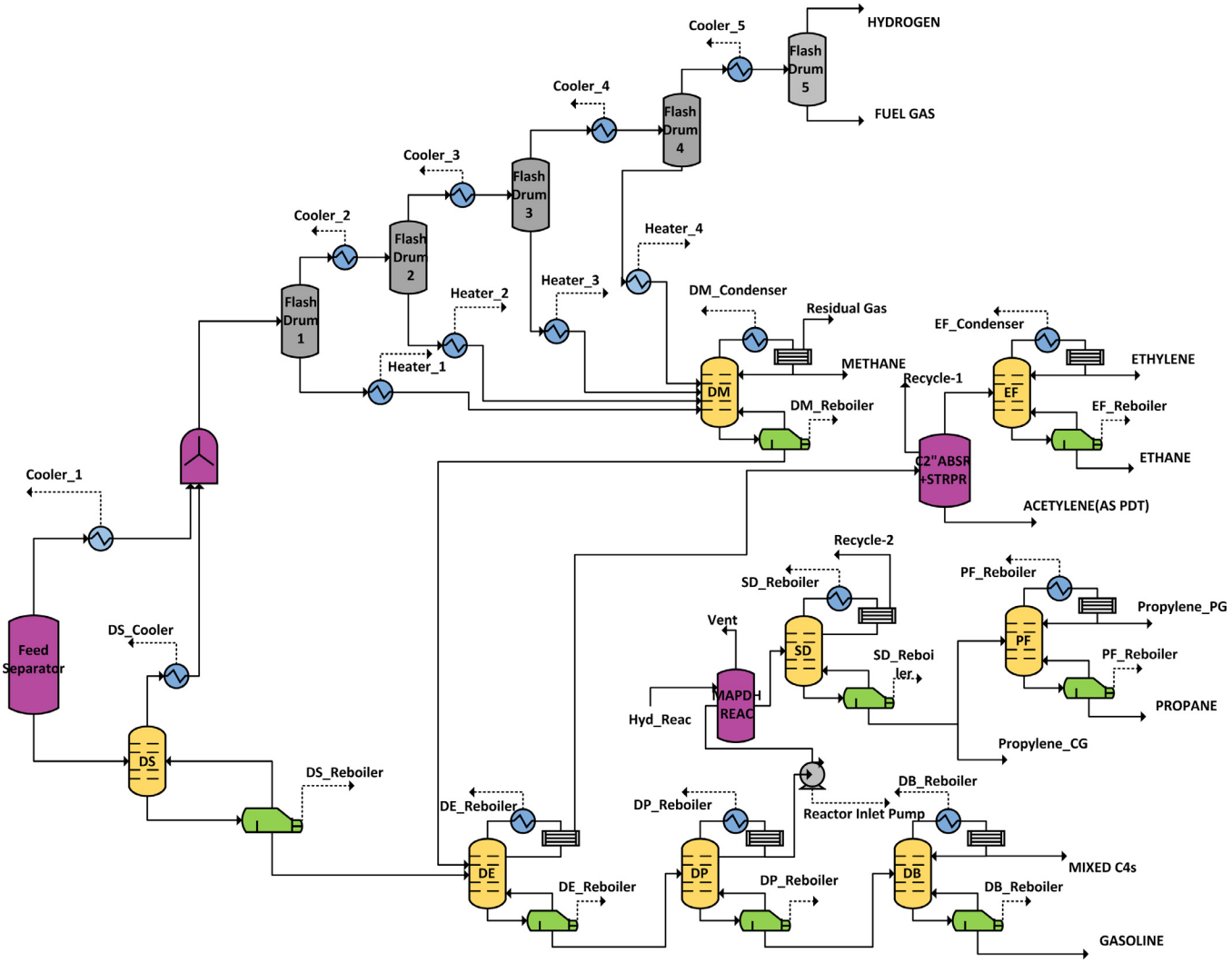


Fig. 3. Schematic flow diagram of an ethylene plant.

cessed based on Grubbs criterion [46]: if $\varphi \geqslant \varphi(n,\alpha)$, then x_i is eliminated, n and α denote the number of data and the significant level, respectively.

$$\varphi = \frac{|V|}{S} = \frac{|x_i - \overline{x}|}{S} \tag{19}$$

where $\overline{x} = \frac{1}{Q}\sum_{i=1}^{Q} x_i$, $S = \sqrt{\frac{1}{Q}\sum_{i=1}^{Q}(x_i - \overline{x})^2}$. And the value of $\varphi(n,\alpha)$ refers to literature Geng et al. [41].

Generally, due to characteristics of energy data of ethylene production systems, the measure unit of energy type parameters including fuel, electricity, water and steam are transformed to the uniform GJ based on energy consumption calculation method of petroleum chemical design (SH/T3110-2001) [47].

5.2. Discussion of the ethylene production prediction

The ethylene production process is a multiple-input and multiple-output process. The noise and abnormal data produced by the ethylene production process could be removed by normalized, consistency test and uniform dimension disposal of units. And then the multiple-input and multiple-output data can be sorted more objectively and accurately. Meanwhile, the precision of energy and carbon emissions analysis of ethylene production plants is significantly improved by using the ISM-ELM method. These preprocess works and the proposed method can demonstrate the advantage of improving energy efficiency and increasing the output. The energy and carbon emissions analysis and prediction flowchart of the ethylene production process based on the ISM-ELM method are shown in Fig. 4.

In order to analyze and predict the energy status and carbon emissions of Chinese ethylene production, the real monthly data of ethylene production scales with from 200,000 to 800,000 under different technologies are application in the empirical analysis. There are about twenty ethylene plants in China, and the monthly energy data of ethylene plants in the latest 10 years came largely from those twenty plants. The required crudes, the production of propylene, ethylene and C_4 and energy types (water, steams, fuels, electricity) of one ethylene plant from 2003 to 2013 are described in Figs. 5and 6, respectively.

First, the monthly data from different ethylene plants are applied to analyze the energy consumption of ethylene plants under the ethylene production scales with from 200,000 to 800,000, and the consistency test is used to remove outliers according to correlation Eq. (13). Second, the normalization is applied to obtain the uniform unit, of energy data of the ethylene plant by different techniques. Based on Eqs. (5)–(8), we can obtained the partial correlation matrix as shown in Table 4.

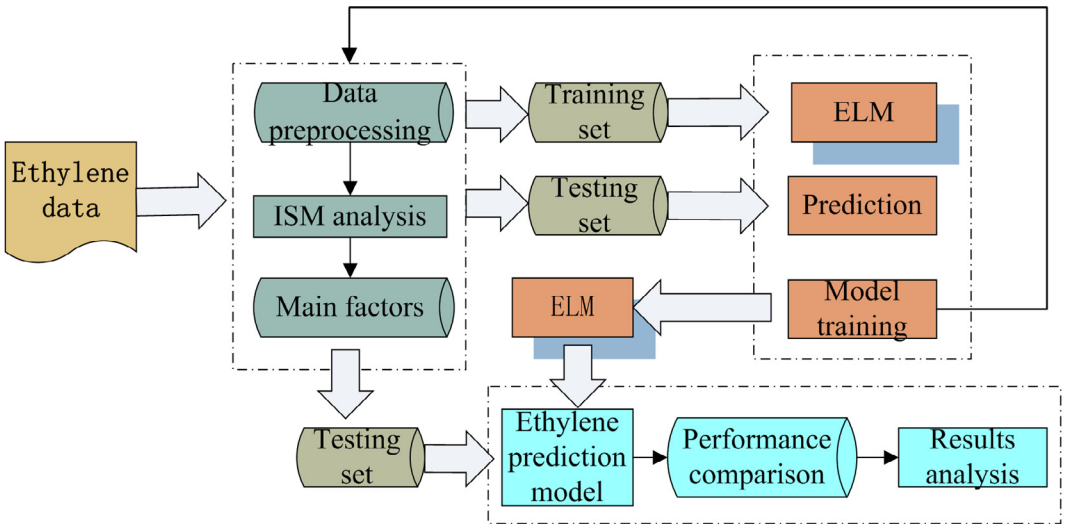


Fig. 4. Flowchart of energy and carbon emissions analysis and prediction.

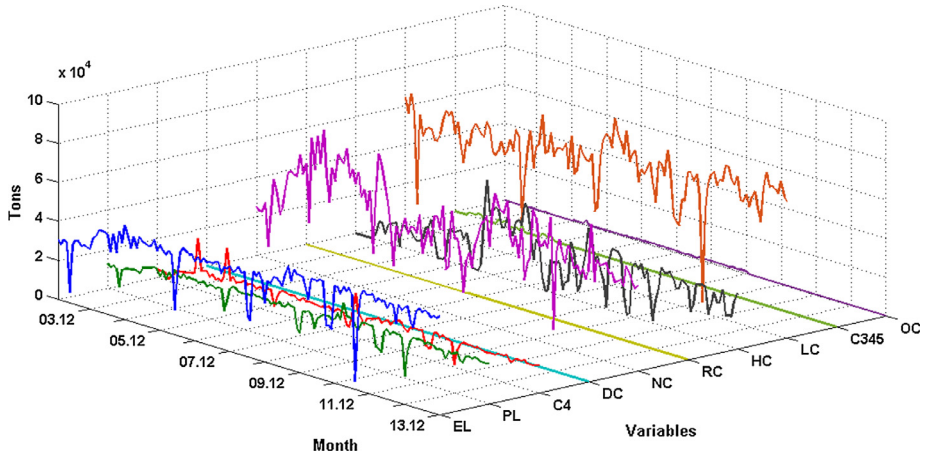


Fig. 5. The required crudes and the main productions of one ethylene plant from 2003 to 2013.

In the ethylene production process, if the correlation coefficient obtained by the public project and energy-related materials of ethylene production such as crude oils, steam, fuel, water, electricity is greater than 0.3, they can be called as be related. To analyze and predict the effect of the ethylene production capacity, if the correlation coefficient is no less than 0.4 [36], then the value between two elements in the adjacency matrix is 1. If $R_{ij} \geqslant 0.4$, then $a_{ij} = 1$ and $a_{ji} = 0$, otherwise, if $R_{ij} \leqslant -0.4$, then $a_{ij} = 0$ and $a_{ji} = 1$, which is called negative correlation.

The reachability matrix R can be made in Table 5 through the thresholds of related variables based on Eqs. (9)–(12).

It can be seen from Table 5 that the light oil has a direct impact on boiler water while the high-pressure steam has a direct impact on the ultra-high pressure steam. Similarly, the low pressure steam has a direct impact on electricity while the medium-pressure steam has a direct impact on fuel oil, ultra-high pressure steam and high-pressure steam.

The reachable set S_i, the first set B_j and their intersection $S_i \cap B_j$ in the first level can be got by the reachability matrix R. Meanwhile, the variables that affect energy efficiency of the ethylene plant in the first level can be obtained, and then we removed the entries that are included in the first level of the reachability matrix R. according to getting the variables in each level, we established the ISM model as shown in Fig. 7.

It can be seen from Fig. 7 that the main significant factors that play a key role in the energy efficiency of ethylene plants are Raffinate, Light diesel oil, Light hydrogenation tail oil, C_{345}, other force, heavy oil, fuel gas, ultra-high pressure steam, low pressure steam, industrial water, recycled water, boiler food water, electricity, etc. The second layer indexes, such as Naphtha, Hydrogenation tail oil, light oil, medium pressure steam and high pressure steam, are affected by the first level indexes of water, the main crude oils and electricity.

The main indexes in the first lever of the ISM and all indexes are taken as the input of the ELM, respectively. And the production sum of propylene, ethylene and C_4 sever as the output of the ELM. Then in this experiment the number of hidden nodes is set as 5, and according to the standard testing, the number of training samples is 824, the number of testing samples is 413, the training time is set as 1000, and the maximum allowed error is set as 0.001. We can obtain the ethylene production prediction model in Fig. 8.

It can be seen from Fig. 8 that the ARGE, the RMSE and the training time of ethylene industries based on ISM-ELM are 0.4843, 0.5026 and 0.0069, respectively. And the ARGE, the RMSE and the training time of ethylene industries based on ELM are 0.5170, 0.5725 and 0.0077, respectively. These indicate that the relatively accurate model prediction ability of the ISM-ELM is better than

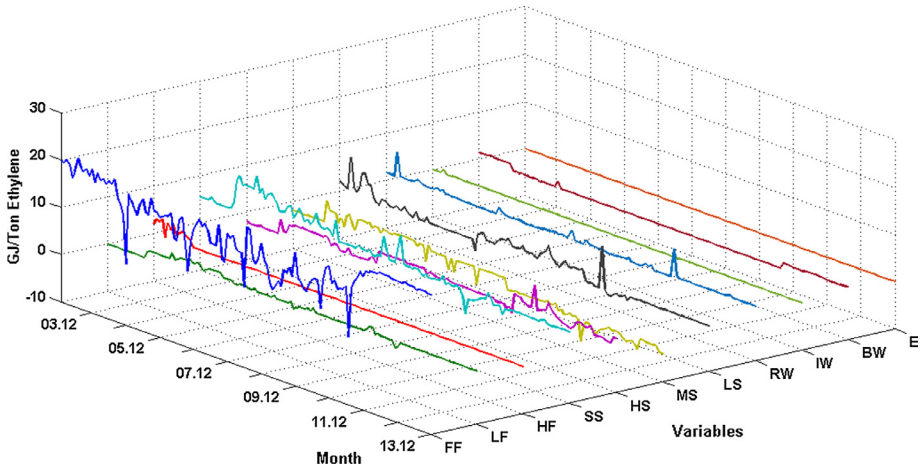


Fig. 6. Energy types (water, steams, fuels, electricity) of one ethylene plant from 2003 to 2013.

Table 4
The partial correlation matrix of the energy consumption data.

	DC	NC	RC	HC	LC	C_{345}	OC	FF	LF	HF	SS	HS	MS	LS	RW	IW	BW	E
DC	−1.00	−0.36	0.34	−0.11	−0.30	−0.15	−0.15	−0.05	−0.10	−0.12	0.20	−0.23	0.02	−0.01	−0.03	−0.03	0.02	−0.17
NC	−0.36	−1.00	0.43	−0.21	−0.22	−0.25	−0.02	−0.10	−0.23	−0.29	0.06	−0.35	−0.11	−0.10	0.07	−0.02	−0.03	−0.29
RC	0.34	0.43	−1.00	0.22	−0.05	0.15	0.01	0.14	0.16	0.10	−0.08	0.20	0.19	0.15	−0.04	0.01	−0.13	0.24
HC	−0.11	−0.21	0.22	−1.00	0.11	0.47	0.13	−0.05	0.26	−0.09	−0.01	−0.10	−0.08	−0.16	−0.04	−0.02	−0.11	−0.13
LC	−0.30	−0.22	−0.05	0.11	−1.00	−0.30	−0.22	−0.12	−0.16	−0.24	0.20	−0.02	−0.12	0.21	−0.02	0.04	−0.21	−0.02
C_{345}	−0.15	−0.25	0.15	0.47	−0.30	−1.00	−0.27	−0.08	−0.25	−0.18	−0.07	−0.19	−0.28	0.25	0.17	0.09	−0.28	−0.15
OC	−0.15	−0.02	0.01	0.13	−0.22	−0.27	−1.00	−0.01	0.03	0.09	−0.13	−0.07	−0.09	0.03	−0.05	0.01	−0.23	−0.11
FF	−0.05	−0.10	0.14	−0.05	−0.12	−0.08	−0.01	−1.00	−0.41	−0.14	−0.22	−0.34	−0.30	−0.08	0.00	0.01	0.01	−0.13
LF	−0.10	−0.23	0.16	0.26	−0.16	−0.25	0.03	−0.41	−1.00	−0.28	−0.03	−0.30	0.05	−0.09	−0.09	−0.01	0.15	−0.23
HF	−0.12	−0.29	0.10	−0.09	−0.24	−0.18	0.09	−0.14	−0.28	−1.00	0.22	−0.28	−0.22	0.18	−0.14	0.01	0.13	0.06
SS	0.20	0.06	−0.08	−0.01	0.20	−0.07	−0.13	−0.22	−0.03	0.22	−1.00	−0.62	−0.25	0.15	0.24	0.01	−0.21	−0.28
HS	−0.23	−0.35	0.20	−0.10	−0.02	−0.19	−0.07	−0.34	−0.30	−0.28	−0.62	−1.00	−0.18	0.09	0.10	−0.01	−0.09	−0.22
MS	0.02	−0.11	0.19	−0.08	−0.12	−0.28	−0.09	−0.30	0.05	−0.22	−0.25	−0.18	−1.00	0.40	0.07	0.06	−0.05	0.07
LS	−0.01	−0.10	0.15	−0.16	0.21	0.25	0.03	−0.08	−0.09	0.18	0.15	0.09	0.40	−1.00	−0.13	0.00	−0.02	−0.10
RW	−0.03	0.07	−0.04	−0.04	−0.02	0.17	−0.05	0.00	−0.09	−0.14	0.24	0.10	0.07	−0.13	−1.00	−0.05	0.36	0.36
IW	−0.03	−0.02	0.01	−0.02	0.04	0.09	0.01	0.01	−0.01	0.01	0.01	−0.01	0.06	0.00	−0.05	−1.00	0.03	0.01
BW	0.02	−0.03	−0.13	−0.11	−0.21	−0.28	−0.23	0.01	0.15	0.13	−0.21	−0.09	−0.05	−0.02	0.36	0.03	−1.00	−0.29
E	−0.17	−0.29	0.24	−0.13	−0.02	−0.15	−0.11	−0.13	−0.23	0.06	−0.28	−0.22	0.07	−0.10	0.36	0.01	−0.29	−1.00

that of the ISM, which proves the effectiveness and accuracy of the proposed method.

Then we predict the sum of outputs of plant 1 and plant 2 under the production scale with from 200,000 to 800,000 Tons in 2012. The results are shown in Fig. 9.

Seen from Fig. 9, the predictive value each month of Plant 1 and Plant 2 fluctuates little, and the predictive production values of Plant 1 are approximate with the practical values, which indicate the steady output level and good production conditions of the plant 1. However, the practical production of Plant 2 decrease from June to September in 2012 and is less than the predictive value, which shows that Plant 2 do not achieve the full production. Based on the predictive results of the former four months, the input data of October has been adjusted, and the plant production will return back to the normal level since this month. As a whole, producing the same scale of ethylene, propylene and C_4 in 2012, Plant 1 uses 2,418,999 Tons crude oil, and Plant 2 used 3,301,771 Tons crude oil. According to carbon emission factor for different types of fuels [48], comparing with Plant 2, the CO_2 emission of Plant 1 reduce 517,039.77 Tons. Obviously, the production conditions of the plant 2 are a little worse and less steady than those of the plant 1.

In the same way, the sum of outputs of plant 1 and plant 2 under the production scale with from 200,000 to 800,000 Tons in 2013 can be obtained in Fig. 10.

In Fig. 10, the predictive values of all productions of the plant 1 are the same with the practical values, while the outputs of ethylene, propylene and C4 in 2013 went down to the lowest till September and October. And plant 1 used 177084.82 Tons crude oil, 21.7 GJ fuel, 2.81 GJ steam, 2.5 GJ water and 1.34 GJ electricity per ton of ethylene in September, while in June this plant used 196,298 Tons crude oil, 21.7 GJ fuel, 1.92 GJ steam, 2.14 GJ water and 1.21 GJ electricity per ton of ethylene, which shows that the production status attain the best, but more produces 11253.16 Tons CO_2 emission at the same time. Similarly, the energy efficiency condition of Plant2 is basically stable. As a whole, producing the same scale of ethylene, propylene and C_4 in 2013, Plant 1 used 2,346,970 Tons crude oil, and Plant 2 used 2,246,948 Tons crude oil. Combing with Plant 1, the CO_2 emission of Plant 2 reduced 58,582.77 Tons.

If we compare the prediction and real values of the two plants in Figs. 9 and 10, we can obtain that the crude oil computation decreases and the energy types (water, steam, fuel, electricity) production computation increases, especially the CO_2 emission of Plant 2 reduced 31.9%, which shows that plant 1 has changed the technology or increased the CO_2 recovery to reduce carbon emissions. However, the production sum of the plant 1 is bigger than the plant 2, which shows that the production scale and the input–output ratio of the plant 2 run normally. If the plant 1 stud-

Table 5
The reachability matrix of the energy consumption data.

	DC	NC	RC	HC	LC	C_{345}	OC	FF	LF	HF	SS	HS	MS	LS	RW	IW	BW	E
DC	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NC	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RC	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HC	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
LC	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
C_{345}	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
OC	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
FF	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
LF	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
HF	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
SS	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
HS	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
MS	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
LS	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
RW	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
IW	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
BW	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
E	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

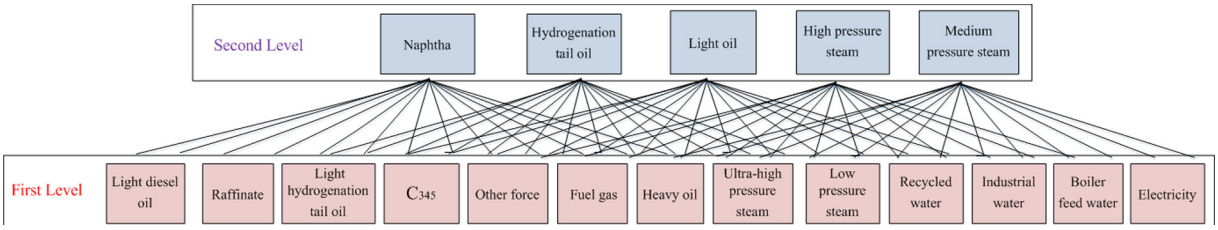


Fig. 7. The hierarchical relation of energy data of the scale with from 200,000 to 800,000 based on ISM model.

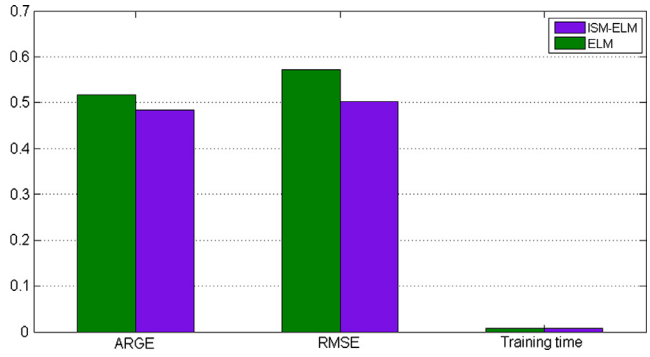


Fig. 8. The ARGE, RMSE and time of ethylene industries based on ELM and ISM-ELM.

ies the technology and experience of the plant 2, it can reduce carbon emissions in future.

By analyzing energy and carbon emissions of this production scale, we can obtain the main indexes and basic factors that affect the energy and carbon emissions of the ethylene production process by the ISM. Moreover, compared with the ELM, the improvement direction of energy and carbon emissions of the different ethylene plants are analyzed better by the analysis and prediction model based the ISM-ELM, and the integrating analysis and prediction method can be applied in other scales of the ethylene industry. Moreover, Fossil energy consumption (Naphtha, light diesel oil, Raffinate, etc.) rather than non-fossil energy consumption is a primary driver of CO_2 emissions. Different energy structures (i.e., the percentages of energy sources in total energy consumption) result

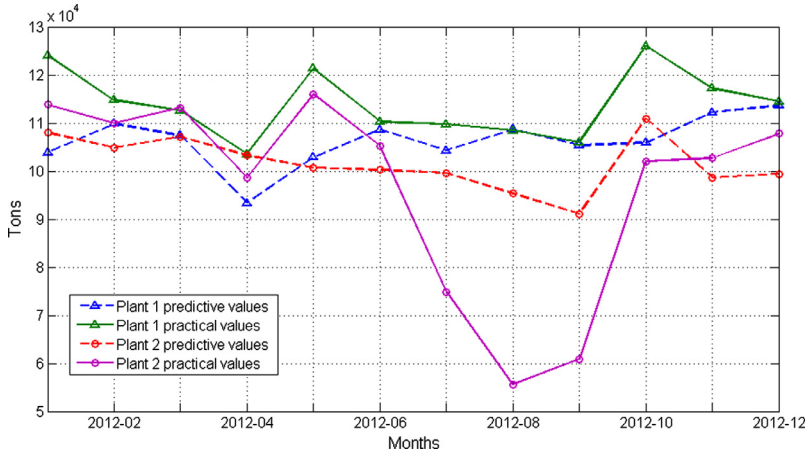


Fig. 9. The prediction production value of plant 1 and Plant 2 based on the ISM-ELM in 2012.

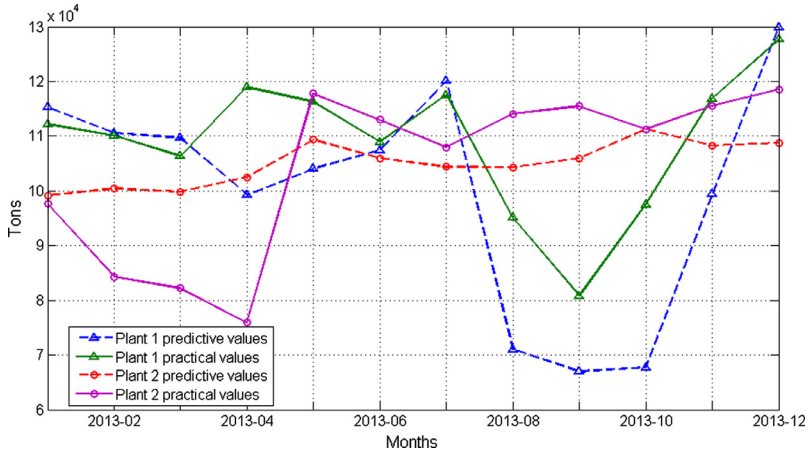


Fig. 10. The prediction production value of plant 1 and Plant 2 based on the ISM-ELM in 2013.

in different carbon emission reductions [49]. Therefore, optimization of the energy structure and energy saving has significant impacts on regional energy efficiencies and CO_2 emissions accordingly.

6. Discussion

First, the energy and carbon emissions analysis and prediction approach based on the improved ISM-ELM is proposed. Compared with the traditional ELM, the robustness and effectiveness of the ISM-ELM is validated based on the UCI standard dataset.

Second, this proposed method is used to analyze and predict the energy and carbon emissions of ethylene production system in the petrochemical industry. The key parameter and basic reasons are obtained to reduce the redundant features to shorten the prediction identification time and improve the accuracy. Meanwhile, the optimum value and the direction of energy saving and carbon emissions reduction are figured out by the ISM-ELM to provide the operation guidance of ethylene production systems.

Third, the proposed method is applied to analyze and predict the energy and carbon emissions status of ethylene production systems effectively. However, the inputs cannot be separated into energy and non-energy inputs of ethylene production systems. Therefore, we will improve our method so that the input and output can be separated reasonably, such as designing a self-adapting ISM-ELM model, which is more suitable for the real world application.

7. Conclusion

The paper presents an energy and carbon emissions analysis and prediction framework based on the ISM-ELM method. The ISM based on differences in the data themselves can overcome the subjectivity shortcoming. Meanwhile, the ISM is applied to further obtain the main indexes and basic factors and reduce the redundant features to shorten the prediction identification time and improve the accuracy. Moreover, in terms of the model accuracy and the training time, the robustness and effectiveness of the ISM-ELM model is better than the ELM based the standard data set from the UCI repository. Finally, the analysis and prediction framework based on the ISM-ELM is verified by using the monthly energy data among different ethylene production plants under the production scale with from 200,000 to 800,000 in the practical ethylene production process. The main indexes that affect the energy and carbon emissions of the ethylene product process could be analyzed objectively. Further, it is used to analyze and predict energy and carbon emissions in complex petrochemical plants reasonably, offer the operation guidance for energy saving and energy and carbon emissions reduction. Furthermore, the proposed method is also applied to analyze and predict energy and carbon emissions of other complex systems in the petrochemical industry.

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 ISM-ELM would be used to analyze and forecast the technical efficiency and the input–output energy measurement of ethylene production process to compare with current studies. Furthermore, energy usage and carbon emissions prediction of thermal power industries could also be analyzed by the self-organizing ISM-ELM.

Acknowledgments

This research was partly funded by National Natural Science Foundation of China (61603025, 61533003 and 61673046) and the Natural Science Foundation of Beijing, China (4162045).

References

[1] L. Xu, Development of ethylene industry in China, Speciality Petrochemicals 31 (2014) 81–84.
[2] L.J. Zhang, J. Hu, Review of sinopec's ethylene production In 2014, Ethylene Ind. 27 (2015) 6–10.
[3] Department of energy statistics, National bureau of statistics, People's Republic of China, China Energy Statistical Yearbook 2013, In China statistics Press, Beijing, 2013
[4] R. Tao, M. Patel, Olefins from conventional and heavy feedstocks: energy use in steam cracking and alternative processes, Energy 31 (2006) 425–451.
[5] World Development Indicators, World Bank, Washington, 2009.
[6] J.A. Baileya, R. Gordona, D. Burtonb, E.K. Yiridoe, Energy conservation on Nova Scotia farms: baseline energy data, Energy 33 (2008) 1144–1154.
[7] Z.Q. Geng, Y.M. Han, Y.Y. Zhang, X.Y. Shi, Data fusion-based extraction method of energy consumption index for the ethylene industry, Lect. Notes Comput. Sci. 6329 (2010) 84–92.
[8] Z.Q. Geng, X.Y. Shi, X.B. Gu, Q.X. Zhu, Hierarchical linear optimal fusion algorithm and its application in ethylene energy consumption indices acquisition, J. Chem. Ind. Eng. (China) 61 (2010) 2056–2060.
[9] L.G. Chen, B. Yang, X. Shen, Z.H. Xie, F.R. Sun, Thermodynamic optimization opportunities for the recovery and utilization of residual energy and heat in China steel industry: a case study, Appl. Therm. Eng. 86 (2015) 151–160.
[10] C.X. Liu, Z.H. Xie, F.R. Sun, L.G. Chen, System dynamics analysis on characteristics of iron-flow in sintering process, Appl. Therm. Eng. 82 (2015) 206–211.
[11] C.X. Liu, Z.H. Xie, F.R. Sun, L.G. Chen, Optimization for sintering proportioning based on energy value, Appl. Therm. Eng. 103 (2016) 1087–1094.
[12] X. Liu, L.G. Chen, X.Y. Qin, F.R. Sun, Exergy loss minimization for a blast furnace with comparative analyses for energy flows and exergy flows, Energy 93 (2015) 10–19.
[13] X. Shen, L.G. Chen, S.J. Xia, F.R. Sun, Numerical simulation of sinter cooling processes in vertical tank and annular cooler, Sci. China: Tech. Sci. 46 (2016) 36–45.

[14] L.G. Chen, H.J. Feng, Z.H. Xie, Generalized thermodynamic optimization for iron and steel production processes: a theoretical exploration and application cases, Entropy 18 (2016) 353–390.
[15] L.G. Chen, X. Shen, S.J. Xia, F.R. Sun, Thermodynamic analyses for recovering residual heat of high temperature basic oxygen gas (BOG) by the methane reforming with carbon dioxide reaction, Energy (2016), http://dx.doi.org/10.1016/j.energy.2016.10.105 (in press).
[16] X. Shen, L.G. Chen, S.J. Xia, F.R. Sun, Numerical simulation and analyses for sinter cooling process with convective and radiative heat transfer, Int. J. Energy Environ. 7 (2016) 303–316.
[17] Y.M. Han, Z.Q. Geng, Energy efficiency hierarchy evaluation based on data envelopment analysis and its application in a petrochemical process, Chem. Eng. Technol. 37 (2014) 2085–2095.
[18] Y.M. Han, Z.Q. Geng, X.B. Gu, Z. Wang, Performance analysis of china ethylene plants by measuring malmquist production efficiency based on an improved data envelopment analysis cross-model, Ind. Eng. Chem. Res. 54 (2015) 272–284.
[19] J. Ruggiero, Impact assessment of input omission on DEA, Int. J. Inform. Technol. Decis. Making 4 (2005) 359–368.
[20] W.W. Cooper, L.M. Seiford, K. Tone, Introduction to Data Envelopment Analysis and Its Uses: With DEA-Solver Software and References, Springer, 2007.
[21] G.B. Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: theory and applications, Neurocomputing 70 (2006) 489–501.
[22] G.J. Wang, K.M. Lam, Z.H. Deng, K.S. Choi, Prediction of mortality after radical cystectomy for bladder cancer by machine learning techniques, Comput. Biol. Med. 63 (2015) 124–132.
[23] P.K. Wong, Z.X. Yang, C.M. Vong, J.H. Zhong, Real-time fault diagnosis for gas turbine generator systems using extreme learning machine, Neurocomputing 128 (2014) 249–257.
[24] M.H. Velayati, N. Amjady, I. Khajevandi, Prediction of dynamic voltage stability status based on Hopf and limit induced bifurcations using extreme learning machine, Electr. Power Energy Syst. 69 (2015) 150–159.
[25] K. Mohammadi, S. Shamshirb, P.L. Yee, D. Petkovi, M. Zamani, S. Ch, Predicting the wind power density based upon extreme learning machine, Energy 86 (2015) 232–239.
[26] K.I. Wong, P.K. Wong, C.S. Cheung, C.M. Vong, Modeling and optimization of biodiesel engine performance using advanced machine learning methods, Energy 55 (2013) 519–528.
[27] S. Salcedo-Sanz, A. Pastor-Sánchez, L. Prieto, A. Blanco-Aguilera, R. García-Herrera, Feature selection in wind speed prediction systems based on a hybrid. Coral reefs optimization-Extreme learning machine approach, Energy Convers. Manage. 87 (2014) 10–18.
[28] W.B. Na, Z.W. Su, Y.F. Ji, Research of single well production prediction based on improved extreme learning machine, Appl. Mech. Mater. 333–335 (2013) 1296–1300.
[29] Z.L. Sun, K.M. Ng, J. Soszýnska-Budny, M.S. Habibullah, Application of the LP-ELM model on transportation system lifetime optimization, IEEE Trans. Intell. Transp. Syst. 12 (2011) 1484–1494.
[30] K.F. Ning, M. Liu, M.Y. Dong, C. Wu, Z.S. Wu, Two efficient twin ELM methods with prediction interval, IEEE Trans. Neural Networks Leaning Syst. 26 (2015) 2058–2071.
[31] Z.Y. Cao, J.C. Xia, M. Zhang, J.S. Jin, L. Deng, X.Y. Wang, J. Qu, Optimization of dear blank preforms based on a new R-GPLVM model utilizing GA-ELM, Knowl.-Based Syst. 83 (2015) 66–80.
[32] T.C. Kuo, H.Y. Ma, S.H. Huang, A.H. Hu, C.S. Huang, Barrier analysis for product service system using interpretive structural model, Int. J. Adv. Manuf. Technol. 49 (2010) 407–417.
[33] F. Talib, Z. Rahman, M.N. Qureshi, Analysis of interaction among the barriers to total quality management implementation using interpretive structural modeling approach, Benchmarking: Int. J. 18 (2011) 563–587.
[34] S. Chandramowli, M. Transue, F.A. Felder, Analysis of barriers to development in landfill communities using interpretive structural modeling, Habitat Int. 35 (2011) 246–253.
[35] K. Govindan, M. Palaniappan, Q. Zhu, D. Kannan, Analysis of third party reverse logistics provider using interpretive structural modeling, Int. J. Prod. Econ. 140 (2012) 204–211.
[36] Y.M. Han, Z.Q. Geng, X.B. Gu, Q.X. Zhu, Energy efficiency analysis based on DEA integrated ISM: a case study for chinese ethylene industries, Eng. Appl. Artif. Intell. 45 (2015) 80–89.
[37] N.L. Yu, D.Y. Yi, X.Q. Tu, Analyze auto-correlations and partial-correlations function in time series, Mathe. Theory Appl. 27 (2007) 54–57.
[38] A. Vargha, L.R. Bergman, H.D. Delaney, Interpretation problems of the partial correlation with nonnormally distributed variables, Qual. Quant. 47 (2013) 3391–3402.
[39] W.Y. Deng, Q.H. Zheng, L. Chen, X.B. Xu, Research on extreme learning of neural networks, Chinese J. Comput. 33 (2010) 279–287.
[40] G.B. Huang, L. Chen, Convex incremental extreme learning machine, Neurocomputing 70 (2007) 3056–3062.
[41] Z.Q. Geng, Y.M. Han, X.B. Gu, Q.X. Zhu, Energy efficiency estimation based on data fusion strategy: case study of ethylene product industry, Ind. Eng. Chem. Res. 51 (2012) 8526–8534.
[42] L.J. Zhang, Application of energy conservation and emission reduction techniques in ethylene plants, Sino-global Energy 14 (2009) 90–94.
[43] Q. Zhang, Z.H. Wang, G.Z. Zhang, A algorithm for discovering positive & negative association rules based on correlation coefficient, J. Shanxi Univ. Technol. 21 (2005) 35–38.
[44] Benchmarks. Available: <http://archive.ics.uci.edu/ml/datasets>.
[45] D.R. Wu, K. He, H.F. Zhu, Analysis of energy consumption and energy saving technology in ethylene complex, Chem. Eng. (China) 35 (2007) 66–71.
[46] H. Zhang, T.S. Tong, F. Zhang, H.L. Liu, Consensus measurement in micro-inertial mensors, J. Transducer Technol. 10 (2001) 40–41.
[47] Calculation method for energy consumption in petrochemical engineering design SH/T 3110, 2011.
[48] 2006 IPCC guidelines for national greenhouse gas inventories: volume Japan: the Institute for Global Environmental Strategies, http://www. ipcc.ch/ipccreports/Methodology-reports.Htm, IPCC, 2008.
[49] Y.W. Bian, P. He, H. Xu, Estimation of potential energy saving and carbon dioxide emission reduction in China based on an extended non-radial DEA approach, Energy Policy 63 (2013) 962–971.