



Comprehensive evaluation of cleaner production in thermal power plants based on an improved least squares support vector machine model

Minquan Ye^{1*}, Jingyi Sun², Shenhai Huang³

¹State Grid Fujian Economic Research Institute, Fuzhou 350000, China

²Beijing Electric Power Engineering Company Limited, Beijing 100071, China

³State Grid Jiaxing Economic Technology Research Institute, Jiaxing 314000, China

ABSTRACT

In order to alleviate the environmental pressure caused by production process of thermal power plants, the application of cleaner production is imperative. To estimate the implementation effects of cleaner production in thermal plants and optimize the strategy duly, it is of great significance to take a comprehensive evaluation for sustainable development. In this paper, a hybrid model that integrated the analytic hierarchy process (AHP) with least squares support vector machine (LSSVM) algorithm optimized by grid search (GS) algorithm is proposed. Based on the establishment of the evaluation index system, AHP is employed to pre-process the data and GS is introduced to optimize the parameters in LSSVM, which can avoid the randomness and inaccuracy of parameters' setting. The results demonstrate that the combined model is able to be employed in the comprehensive evaluation of the cleaner production in the thermal power plants.

Keywords: Analytic hierarchy process, Comprehensive evaluation, Grid search, Least squares support vector machine, Thermal power plants cleaner production

1. Introduction

With the burgeoning development of industry, cleaner production, which is a new theory about pollution control and environmental governance, has been proposed. It provides a new idea for effective environmental protection so as to deal with more and more serious energy consumption and environmental pollution [1]. Based on the implementation of cleaner production, the achievement of energy saving and emission reduction during the production process can be greatly promoted.

In terms of cleaner production theory, Sanz et al. [2] analyzed the development status of micro and small sized enterprises in developing countries and applied the theory of cleaner production to Systematic Approach to Social Inclusion (SASI) method. In the light of social and economic constraints, the results showed that the micro and small sized enterprises needed to execute

the cleaner production in all fields and expedite the establishment of supervisory system and systemic policy mechanisms. Vieira and Amaral [3] thought the clean production theory was a form of sustainable development in the enterprise, and analyzed the reason why it was not widely adopted. The results indicated that the implementation of clean production was related to local culture, policy, education, social pressure and so on. Therefore, it is necessary to unite the business, academics, governments and community to spread this theory. Wang et al. [4] put forward that the cleaner production was the imperative method for thermal power enterprises to reduce environmental pollution and improve energy efficiency. Through taking two power plants in China as an example, this paper studied the relationship among energy consumption ratio, resource consumption and cleaner production management by combining qualitative evaluation with quantitative analysis. In accordance with the results, im-



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2019 Korean Society of Environmental Engineers

Received September 28, 2018 Accepted November 17, 2018

* Corresponding author
Email: hdymq2014@163.com
Tel: +86-152-3027-9980

proved strategies were proposed.

At present, there are few studies on the evaluation of clean production at home and abroad. Sun and Liang [5] employed particle swarm optimization (PSO) algorithm to obtain the optimal parameters in LSSVM. Comparing with the traditional LSSVM model, the average relative error (RE) of the proposed technique was less than 0.285%, which showed the developed technique was more effective in the comprehensive evaluation of power plants cleaner production. Wang and Cheng [6] derived the evaluation values through setting up a fuzzy matrix and calculating the weights of indexes. The results of this study can be used as an objective basis for the government or intermediary institutions to assess the implementation level of cleaner production and to promote the emission trading system. Wen [7] hold the view that clean production was an effective way to reduce energy and environmental pollution in the power industry. However how to judge the performance of clean production was a primary problem. In this paper, an evaluation index system was established and an intuitionistic fuzzy cluster decision method was proposed to assess the performance of clean production in thermal power plants.

In the field of electricity and energy, the thermal power capacity is increasing rapidly. However, a series of problems have appeared, such as greenhouse effect, acid rain, ozone depletion, etc. Accordingly, the application of cleaner production theory to thermal power enterprises is necessary. It can promote the realization of low energy consumption and low pollution of thermal power generation as well as avoid the lag of traditional terminal control. The evaluation of cleaner production in thermal power plants can help managers clarify the actual implementation and optimize the relevant measures. In this way, the competitiveness of thermal power plants can be comprehensively enhanced, not only saving energy and reducing emissions, but also reducing the risk of environmental liabilities, which can improve the economic benefits of enterprises [8].

This paper selects GS to optimize the parameters in LSSVM, and establishes this hybrid evaluation model for the thermal power plants cleaner production. The rest of paper is organized as follows: Section 2 builds a comprehensive evaluation index system for thermal power plants cleaner production; Section 3 provides a brief description of AHP, LSSVM and GS as well as the specific steps of the developed approach; Section 4 carries out a case study to validate the hybrid model; Section 5 concludes this paper.

2. Establishment of Comprehensive Evaluation Index System of Thermal Power Plants Cleaner Production

The comprehensive evaluation index system is the core of the whole assessment. The layers of the built index system are interrelated with each other, meanwhile, the lower level index should play a good role in analysis of the expansion for the upper level index. The evaluation index system, which is revised jointly by the National Development and Reform Commission, Ministry of Industry and Information Technology of the People's Republic of China and the environmental protection departments, is taken as a basis. Production situation of the thermal power plants cleaner production is the first layer of the index system. The second layer of index system is composed of 7 parts including energy consumption, resource consumption, comprehensive utilization, pollutant discharge, the compliance of cleaner production technologies, cleaner production management and the compliance of environmental protection. In this paper, on the foundation of the characteristics of the power industry, the second level indexes are divided into 30 third level indicators. The specific content is shown in Table 1.

Table 1. The Thermal Power Plants Cleaner Production Evaluation Index System

Top level	Middle level	Bottom level
The cleaner production evaluation index system of thermal power plants (U)	Energy consumption (A)	Coal consumption for power generation (A1)
		Power supply coal consumption (A2)
		Consumption of crude oil (A3)
	Resource consumption (B)	Water consumption for power generation unit (B1)
		Consumption of industrial water (B2)
		Industrial water reuse rate (B3)
		Rate of steam/water waste (B4)
	Comprehensive utilization (C)	Utilization ratio of fly ash (C1)
		Utilization ratio of desulfurized gypsum (C2)
		Total displacement of fly ash (D1)
		Soot emissions power generation units (D2)
		Pollutant discharge (D)
	NO _x emissions from power generation units (D4)	
	Factory noise (D5)	

Table 1. Continued

Top level	Middle level	Bottom level
The cleaner production evaluation index system of thermal power plants (U)	The compliance of cleaner production technologies (E)	Shut down situation of violation units (E1)
		Transformation performance of turbine flow passage (E2)
		Use situation of fuel ignition technology (E3)
		Reconstruction situation of pumps and fans (E4)
		Construction situation of operation monitoring device (E5)
		Situation of SO ₂ governance (E6)
	Cleaner production management (F)	Implementation of low-NO _x combustion technology (E7)
		Status of waste water treatment and reuse in the whole plant (E8)
		Situations of carrying out fuel, heat, energy and water balance (F1)
	The compliance of environmental protection (G)	Status of carrying out coal source control (F2)
		Circumstances of carrying out a comprehensive audit of cleaner production (F3)
		Situations of establishing an environmental and certified management system case (G1)
		Implementation of environmental protection (G2)
		Implementation of environmental impact assessment system (G3)
		Implementation of pollution source control (G4)
		Situation of pollutant emission control (G5)

3. Comprehensive Evaluation Model of Cleaner Production Based on AHP and LSSVM Optimized by GS

3.1. AHP

The AHP is a structured technique for complex decision problems. Firstly, the decision target is determined, then the objective is decomposed into multiple criteria and further into diverse indicators. Each criterion and indicator can be analyzed independently, which provides a comprehensive and rational framework for the decision problem. These elements are related to the overall goal through quantifying and the evaluation results can be finally obtained [9]. Specifically speaking, AHP converts these evaluations which are derived from subjective scores of experts into numerical values that are able to be processed and compared over the entire range of the problem. Afterwards, the quantitative processed indexes are taken to analyze and sort. The essence of AHP is the ranking of various indexes and criteria.

3.2. LSSVM

As an improved algorithm based on support vector machine (SVM), LSSVM is widely used in intelligent computing. It transforms the inequality constraints of traditional SVM into equality constraints, as well as considers sum squares error and loss function as the loss experience of the training set, which converts quadratic programming problems into linear equations ones. Accordingly, the computing speed and accuracy have been improved [10]. Here, the training set is treated as $\{(x_k, y_k) | k = 1, 2, \dots, n\}$, in which $x_k \in R^n$

is the input data, and $y_k \in R^n$ is the output data. $\varphi(\bullet)$ is the nonlinear mapping function which transfers the samples into much higher dimensional feature space $\phi(x_k)$. Thus, the optimal decision function is established in the high-dimensional feature space [11]:

$$y(x) = \omega^T \cdot \varphi(x) + b \quad (1)$$

Given the training data set $\{x_i, y_i\} (i = 1, 2, \dots, N)$, the optimization problem of LSSVM can be defined as Eq. (2):

$$\min_{\omega, b, e} (\omega, e) = \frac{1}{2} \omega^T \omega + \frac{1}{2} C \sum_{i=1}^n \xi_i^2 \quad (2)$$

Where ω equals n -dimensional weight vector; ξ_i represents the training error; $C > 0$ is the regularization parameter that makes a balance between the training error and model complexity.

The core of this approach is the symmetric function that satisfies the Mercer condition, which is generally called as a kernel function. Gaussian (RBF) is used in this paper,

$$K(x, x_i) = e^{-\frac{(x-x_i)^2}{2\sigma^2}} \quad (3)$$

where σ^2 is the kernel width.

In RBF, the value of C is associated with tolerable error. The larger C allows smaller errors. In addition, the kernel width σ is related to the input spatial extent or width of learning samples. The larger the sample input space is, the greater the values are.

3.3. Parameter Optimization Based on GS

3.3.1. Cross validation (CV)

CV is a statistical method widely employed to validate the classifier performance. This technique divides the original data into two parts, namely training set which is adopted to train the classifier as well as validation set that is used to test the trained model. Thus, the evaluation index of the classifier performance can be derived.

K-fold CV is presented as a representative to divide original data into K groups (generally based on mean values). Each subset of the data is taken as a validation set in turn and the remaining $K-1$ subsets are used as training sets, therefore K models can be obtained. Here, the average of classification precision derived from K validation sets is regarded as the classifier performance, wherein $K \geq 2$ and K starts from 3 actually. Only when the number of original data is small, $K = 2$. K-CV can effectively address the problems of over-learning or less-learning and obtain convincing results.

3.3.2. GS

The basic principle of GS is to divide the grid in a range by the parameters that need to be optimized and traverse all the points to achieve parameters' setting. For the determined parameters, K-CV approach is introduced to derive the classification precision of the corresponding training set. Hence, the parameters with highest accuracy of classification are selected as the optimal ones.

3.4. Approaches of Comprehensive Evaluation Model

The comprehensive evaluation method of cleaner production in thermal power plants incorporating AHP, GS and LSSVM is constructed as follows:

- i) Collect the data of the established index system from five thermal power plants and apply dimensionless processing to them.
- ii) Calculate the weights of the indexes based on AHP. The judgment matrixes are obtained through expert scoring method and consistency check is carried out here to test the rationality.
- iii) Measure the comprehensive evaluation values of cleaner production of each thermal power plants on the foundation of processed results above.

iv) 15 samples are selected as a training set, and the remaining data of 10 samples are used as a test set. The radial basis function (RBF) is exploited as the kernel function in this paper. Simultaneously, the determination of these two parameters, regularization parameter C and the kernel width σ is generally based on GS and CV.

Divide the ranges of $\log_2 C$ and $\log_2 \sigma$ into a few grids, as well as all samples into k groups by cross validation. C and σ are fixed on the grid, afterwards $k-1$ groups, namely training samples are taken into the proposed approach to achieve the optimal evaluation model. Mean square errors (MSE) of n samples are calculated as follows:

$$MSE = \frac{1}{n} \sum_{i=1}^n \left(\hat{y}_i - y_i \right)^2 \quad (4)$$

where \hat{y}_i and y_i are the actual and forecasted evaluation results, respectively.

i) The achievement of minimum value for MSE means the optimal parameters C and σ in LSSVM model. Thus, a comprehensive evaluation method of cleaner production is established.

4. Case Study

Based on the established comprehensive evaluation index system in Table 1, the domestic-related data are collected from five thermal power plants from 2012 to 2016. The 15 groups of data from plant power plant V, W and X are used as training samples, while the 10 remaining groups of data from power plant Y and Z are selected as testing samples.

4.1. Calculation of Index Weights

The weights of comprehensive evaluation indexes can be obtained by AHP integrated expert scoring method with consistency check. The weights of indexes in each level are presented in Table S1.

According to the weights of the basic indicators and the second level that are respectively illustrated in Table S1, the weights of the basic indexes relative to the goal can be calculated. The results are listed in Table S2.

4.2. Calculation of Sample Output Value

This paper applies expert scoring method to estimation of the qualitative indicators about the thermal power plants cleaner production. The experts are asked according to relevant national policies over the standard criterion-referenced objective scoring. The scoring range is in $[0,1]$, where $[0,0.2]$, $[0.2,0.4]$, $[0.4,0.6]$, $[0.6,0.8]$ and $[0.8,1]$ represents the worst, worse, general, better and excellent, partly. The higher the score is, the better the performance of index is. The average value of expert scoring is selected as the final result.

To avoid the disadvantage of large value scope of the quantitative index, as well as the incomparable unit values between the indexes, this paper introduces dimensionless processing to address this problem. The values of all indexes range between 0 and 1. Due to the existence of positive index and inverse one, the dimensionless processing formula is different, which are listed as follows:

The dimensionless processing formula for positive index

$$y_i = \frac{x_i}{\max_{1 \leq i \leq n} x_i} \quad (5)$$

The dimensionless processing formula for inverse index

$$y_i = \frac{\max x_i + \min x_i - x_i}{\max x_i} \quad (6)$$

In line with the results of both dimensionless processing and expert scoring for each index, the comprehensive evaluation values of each thermal power plants cleaner production for testing can be derived based on Eq. (7). The results are regarded as the expected output values of the samples.

$$E = \sum s \cdot w \quad (7)$$

where E equals the expected output value, s represents the score of each index and w is the weight. The final 25 expected output values are presented in Table2.

Table 2. The Final 25 Expected Output Values

	2011	2012	2013	2014	2015
V	0.7964	0.8176	0.8501	0.9430	0.9455
W	0.7855	0.7997	0.8372	0.8732	0.9095
X	0.7590	0.7707	0.8118	0.8436	0.8927
Y	0.7933	0.8199	0.8252	0.8921	0.9276
Z	0.7165	0.7441	0.7842	0.8096	0.8479

4.3. Training of Comprehensive Evaluation Model

4.3.1. Parameter optimization

In accordance with the proposed evaluation algorithm, 25 sample data from V,W,X,Y,Z thermal power plants should be pre-processed by normalization method. The range of $\log_2 C$ and $\log_2 \sigma$ is $[-10, 10]$, and mesh width is set to 0.4. Five-fold cross validation is implemented on the training samples. The results are illustrated in the contour map and 3D view plot of GS, which are, respectively shown in Fig. 1 and Fig. 2.

Obviously, in the light of curve trend, the best values of regularization parameter C and kernel width parameter σ are 3.0314 and 0.0272, respectively. Correspondingly, MSE of cross-validation equals 0.0027.

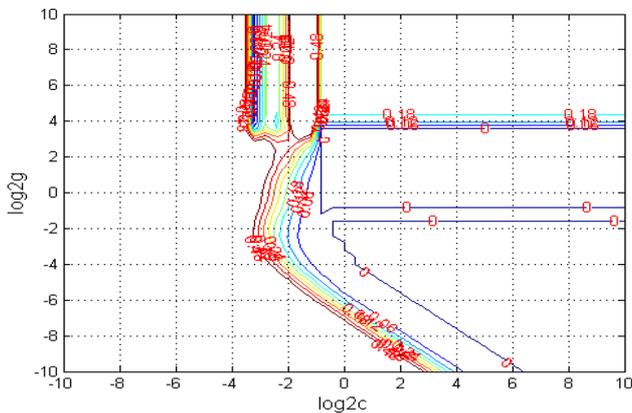


Fig. 1. Contour map of GS.

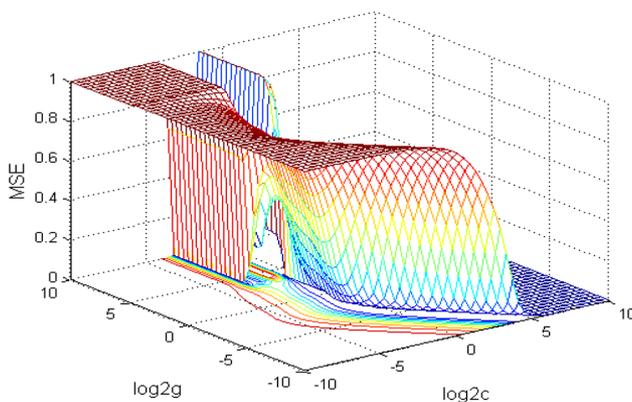


Fig. 2. 3D view plot of GS.

4.3.2. Training results of comprehensive evaluation model

In this paper, RBF is used as the kernel function. Here, the optimal parameter C^* , σ^* and training samples are taken into LSSVM model so that the trained evaluation model is obtained. To examine the performance of this approach, the RE and the mean absolute percentage error (MAPE) are proposed to measure the evaluation accuracy. The formulas are defined as follows:

$$RE(i) = \frac{\hat{y}_i - y_i}{y_i} \times 100\% \tag{8}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \tag{9}$$

where y_i and \hat{y}_i represent the expected output and evaluation result, respectively.

Table 3 illustrates the training results of the established evaluation model.

As shown in Table 3, the MAPE of the former 15 sets equals 0.245%. Notably, the maximum RE is 0.316%, and the minimum is less than 0.1%, which indicates that the training results are satisfactory. Thus, the proposed model presents high training accuracy. Based on the above results, the remaining 10 samples can be applied to test.

Table 3. Training Results of GS-LSSVM

Serial number	Expected output value	Training output value	RE (%)
1	0.7964	0.7987	0.287%
2	0.8176	0.8198	0.269%
3	0.8501	0.8504	0.031%
4	0.9430	0.9407	-0.247%
5	0.9455	0.9438	-0.185%
6	0.7855	0.7832	-0.298%
7	0.7997	0.8021	0.298%
8	0.8372	0.8349	-0.271%
9	0.8732	0.8721	-0.124%
10	0.9095	0.9072	-0.254%
11	0.7165	0.7188	0.316%
12	0.7441	0.7464	0.314%
13	0.7842	0.7819	-0.300%
14	0.8096	0.8114	0.225%
15	0.8479	0.8457	-0.260%
MAPE (%)			0.245%

4.4. Test and Analysis of Results

The evaluation results of the rest 10 sets are shown in Table 4. It can be seen that the actual values are close to the output with a low fluctuation range, which indicates the developed GS-LSSVM model owns high accuracy and strong generalization

Table 4. Testing Results of GS-LSSVM, PSO-LSSVM and LSSVM

Test sample number	Actual value	GS-LSSVM		PSO-LSSVM		LSSVM	
		Output value	RE	Output value	RE	Output value	RE
1	0.7933	0.7776	-1.968%	0.8099	1.664%	0.7964	0.395%
2	0.8199	0.8059	-1.710%	0.8366	1.670%	0.8157	-0.513%
3	0.8252	0.8341	1.079%	0.837	1.180%	0.832	0.824%
4	0.8921	0.8911	-0.107%	0.9123	2.020%	0.8805	-1.300%
5	0.9276	0.9250	-0.288%	0.9175	-1.014%	0.8984	-3.152%
6	0.7590	0.7650	0.790%	0.7835	2.446%	0.7967	4.961%
7	0.7707	0.7865	2.048%	0.7935	2.283%	0.8072	4.740%
8	0.8118	0.8298	2.217%	0.8572	4.538%	0.8371	3.114%
9	0.8436	0.8584	1.758%	0.8394	-0.418%	0.8618	2.160%
10	0.8927	0.9010	0.930%	0.8792	-1.354%	0.8784	-1.607%
MAPE		1.290%		1.818%		2.277%	

ability. Thus it is suitable for the comprehensive evaluation of cleaner production in thermal power plants.

In order to verify the validity of the above technique and the evaluation accuracy, LSSVM optimized by PSO, namely PSO-LSSVM and single LSSVM are introduced in this paper as the comparable models. The parameters in single LSSVM are set as $C = 20, \sigma = 5$. The test results are shown in Table 4.

Compared with PSO-LSSVM and LSSVM, it can be seen from Table 4 and Fig. 3 that the output value tested by GS-LSSVM is most close to actual one of cleaner production in thermal power plants. Among the 10 test samples, there are 7 output values superior to corresponding results by PSO-LSSVM and LSSVM. Thus, the best performance belongs to GS-LSSVM, which indicates the GS optimization part can effectively avoid the randomness of the parameters' setting in LSSVM model.

As shown in Table 4 and Fig. 4, the values of MAPE derived from GS-LSSVM, PSO-LSSVM and LSSVM are 1.290%, 1.818% and 2.277%, respectively. It obviously proves that GS-LSSVM presents the highest evaluation precision. In addition, according to the comparable results of RE shown in Fig. 4, the range of RE by GS-LSSVM is [-1.968%, 2.217%], while the values of RE calculated

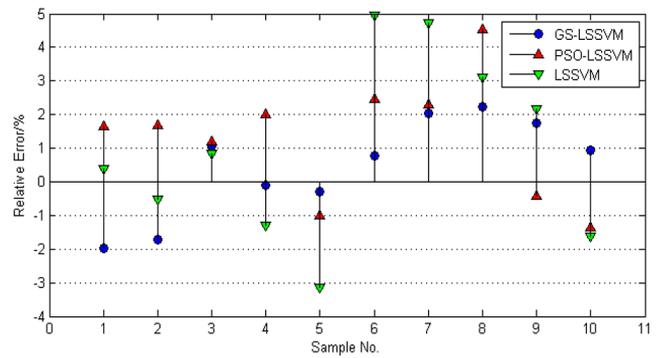


Fig. 4. Comparison of RE.

by PSO-LSSVM and LSSVM are both larger, namely [-1.354%, 4.538%] and [-3.152%, 4.961%], partly. Therefore it can be concluded that the error fluctuation of GS-LSSVM is lower than that of PSO-LSSVM and LSSVM.

Above all, the generalization and evaluation precision of GS-LSSVM is superior to PSO-LSSVM and single LSSVM model. Hence, the proposed approach is able to solve the problem of cleaner production evaluation in thermal power plants effectively.

5. Conclusions

Considering the rapid development of thermal power plants, the ecological environment has been influenced by a series of problems, such as greenhouse effect, acid rain, ozone depletion, etc. Therefore, it's of great necessity to apply the cleaner production theory at home and abroad to thermal power plants and evaluate the implementation effects. This paper proposes a hybrid method based on AHP and LSSVM optimized by GS for comprehensive evaluation of cleaner production. The case study indicates that the proposed model is efficient enough to compete with existing techniques with high accuracy and strong generalization ability. Thus, this established algorithm is able to provide a basis for cleaner pro-

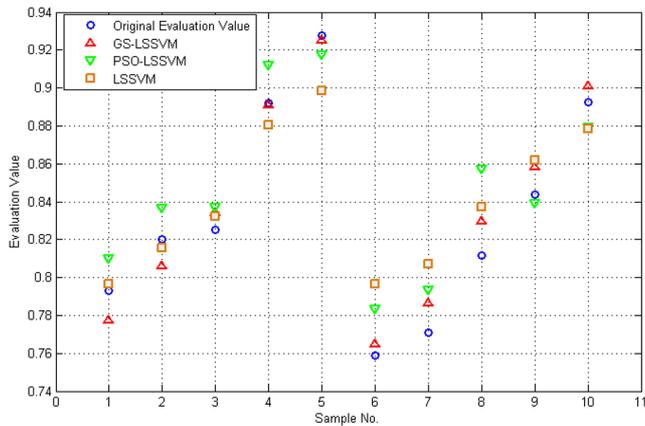


Fig. 3. Result comparison.

duction comprehensive evaluation of thermal power plants.

In order to further put the theory of cleaner production into practice, the following measures can be taken based on the evaluation results and index system: (a) Strengthen the management of raw coal blending technology and constantly optimize it so as to realize the reduction of coal consumption in production. (b) Eliminate old equipment and technical standards that don't meet the requirements of cleaner production. The power generation facilities should be updated in time and the latest technology of saving oil as well as coal needs to be introduced duly. (c) Improve the audit mechanism of cleaner production. The corresponding audit institution should be constructed with independent power. (d) Make periodic cleaner production inspection system so that the defects during the implementation of schemes in each stage can be found out and amended them accordingly. (e) Combine the implementation effect of cleaner production with performance incentives in department assessment so that the realization of target can be ensured.

References

1. Lin YH, Tseng ML, Chiang JH, Chen YM. Prioritization of competitive priority in cleaner production implementation. In: Industrial engineering and engineering management, 2007 IEEE International Conference on IEEE; 2-4 December 2007; Singapore. p. 104-108.
2. Sanz M, Siebel MA, Ahlers R, Gupta J. New approaches to cleaner production: Applying the SASI method to micro-tanneries in Colombia. *J. Clean. Prod.* 2016;112:963-971.
3. Vieira LC, Amaral FG. Barriers and strategies applying cleaner production: A systematic review. *J. Clean. Prod.* 2016;113:5-16.
4. Wang H, Xu J, Wang D, Zhang T, Liu Y. Cleaner production based on sustainable development in Chinese power plants. *Environ. Eng. Sci.* 2015;32:461-469.
5. Sun W, Liang Y. Comprehensive evaluation of cleaner production in thermal power plants using particle swarm optimization based least squares support vector machines. *J. Inform. Comput. Sci.* 2015;12:1993-2000.
6. Wang LX, Cheng K. The research on comprehensive fuzzy evaluation model of enterprise cleaner production. In: Information management, innovation management and industrial engineering, 2008. ICIII'08. International Conference on IEEE; 19-21 December 2008. p. 296-300.
7. Wen L. Thermal power plants cleaner production performance evaluation based on intuitionistic fuzzy set group decision-making. *Adv. Mater. Res.* 2010;113-116:1311-1315.
8. Hilson G. Pollution prevention and cleaner production in the mining industry: An analysis of current issues. *J. Clean. Prod.* 2000;8:119-126.
9. Zyoud SH, Kaufmann LG, Shaheen H, Samhan S, Fuchs-Hanusch D. A framework for water loss management in developing countries under fuzzy environment: Integration of Fuzzy AHP with Fuzzy TOPSIS. *Expert Syst. Appl.* 2016;61:86-105.
10. Suykens JAK, Vandewalle J. Least squares support vector machine classifiers. *Neural Process. Lett.* 1999;9:293-300.
11. Mesbah M, Soroush E, Azari V, Lee M, Bahadori A, Habibnia S. Vapor liquid equilibrium prediction of carbon dioxide and hydrocarbon systems using LSSVM algorithm. *J. Supercrit. Fluid.* 2015;97:256-267.