Data Envelopment Analysis Model and Its Application on Technical Efficiency Evaluation of Coal Enterprise

F. Pan[1], N. L. Li[1], J. Gao[2]
[1]School of mines, China University of Mining and Technology, Xuzhou, P. R. China
[2]Eco-Materials and Renewable energy Research Center, Nanjing University, Nanjing, P. R. China
(cumt_lnl@126.com)

Abstract - **A math model based on Data Envelopment Analysis was proposed to make comprehensive technical efficiency evaluation for coal enterprise with multi-input and multi-output systems generally, and this makes it complex and difficult to make precise technical efficiency evaluation, for example, it is hard to set evaluating weights index and difficult to get sound evaluation when the evaluation function is unknown. However, these difficulties can be avoided with DEA model. The C^2GS^2 model and EC^2GS^2 model derived from C^2GS^2 model were proposed and their efficiency judgment criterions were given in this paper. Then based on existing data, the model was applied to make technical efficiency evaluation for Chinese coal enterprises to estimate their technical performance.**

Keywords - **EC^2GS^2 model, C^2GS^2 model, Coal Enterprise, Data Envelopment Analysis**

I. INTRODUCTION

Data Envelopment Analysis (DEA) is an Operational Research Method proposed by Charnes and Cooper in 1978 basing on the concept of relative effectiveness [1]. It is applied to measure productive efficiency of Decision Making Unit (DMU) empirically and provides benchmarks to promote DMUs efficiency. In a DEA model, the efficiency of DMU is defined as the ratio of the weighted sum of its outputs to the weighted sum of its inputs. The performance of DMUs depends only on the identified efficient frontier [2]. DEA finds the most favorable set of weights for each DMU (i.e., the set of weights that optimizes the DMU's efficiency rating under the constraint that the efficiencies of all DMUs are not more than 1).

Substantial research has been conducted on developing new DEA methodology to calculate efficiency subsequently. There are primarily two models for DEA models: C^2R model [1] and C^2GS^2 model [2], and C^2GS^2 model is applied to assess the technical efficiency of DMUs and provides benchmarks for inefficient DMUs to improve. Improved and extended model were proposed as Assurance Region model [3], Cone-ratio model [4][5] and Stochastic Chance-constrained Model [6][7][8][9].

As a nonparametric approach, when applied to estimation of the efficiency of complex systems such as Coal enterprises, DEA is characterized with four main advantages: (1) the ability to assume relationships between inputs and outputs and set their weights automatically, without having to set weights for every index, (2) the ability to handle multiple inputs and outputs automatically without requiring any assumption of certain functions, (3) the ability to evaluate index dimension without consistent (4) the ability to calculate relative efficiency with linear programming method which outweigh absolute efficiency.

Since came into being, DEA has been introduced by a large number of literatures and has been applied so solve various sets of problems. And it has found ways into such fields as bank business [10], Air transportation networks [11], urban community construction [12][13] and hospital comparison [14].

However, little research has been carried out about application of DEA in estimating coal enterprises, especially large-scale ones with complex system of multiple inputs and outputs. Application of DEA on coal enterprises shows significant efficiency on both recognizing self operation condition and representing guidelines on taking action to improve efficiency and profit.

In this paper, the C^2GS^2 model and EC^2GS^2 model of DEA were applied to evaluate coal enterprise technical efficiency with multiple inputs and outputs structured systems like many other big enterprises.

This paper is organized as follows. In section 2, the structure and principle of C^2GS^2 and EC^2GS^2 of DEA were introduced and the defects of C^2GS^2 were pointed out. In section 3, a case study with eight coal enterprises was conducted. Advantages and disadvantages were discussed in section 4 and in section 5 a conclusion was given.

II. METHODOLOGY

A. *DEA analyzed C^2GS^2 model*

Suppose that there are n+1 DMUs and each of which has m input indexes and s output indexes. The $X_j = (x_{1j}, x_{2j}, ..., x_{mj})^T > 0$ is the vector of actual inputs and $Y_j = (y_{1j}, y_{2j}, ..., y_{sj})^T > 0$ is the vector of actual outputs for DMU_j. And x_{ij} is the actual value of input and y_{ij} is the actual value of output for DMU_j ($x_{ij}>0$, $y_{ij}>0$). Take DMU_0 as an example, the linear programming model is described as formulation (1):

978-1-4244-3672-9/09/$25.00 ©2009 IEEE

$$Max\frac{U^T Y_0 + u_0}{V^T X_0}$$

$$s.t.\begin{cases} \dfrac{U^T Y_j + u_0}{V^T X_j} \le 1 \\ U \ge 0, V \ge 0 \\ u_0 \text{ without symble constraint} \\ j = 0,...,n \end{cases} \quad (1)$$

Where U means the vector of inputs weights and V means the vector of outputs weights. Equivalent linear programming can be obtained by Charnes-Cooper transformation and is given as formulation (2):

$$Max(U^T Y_0 + u_0)$$

$$s.t.\begin{cases} \omega^T X_j - \mu^T Y_j - u_0 \ge 0 \\ \omega^T X_0 = 1 \\ \omega \ge 0, \mu \ge 0 \\ u_0 \text{ without symble constraint} \\ j = 0,...,n \end{cases} \quad (2)$$

The linear programming can be obtained by introducing non-Archimedean infinitesimal quantity to dual programming (3):

$$Min[\theta - \varepsilon(\hat{e}^T S^- + \hat{e}^T S^+)$$

$$(D_\varepsilon) s.t.\begin{cases} \sum_{j=0}^{n} X_j \lambda_j + S^- = \theta X_0 \\ \sum_{j=0}^{n} Y_j \lambda_j - S^+ = Y_0 \\ \sum_{j=0}^{n} \lambda_j = 1 \\ S^-, S^+, \lambda_j \ge 0 \\ j = 0,...,n \end{cases} \quad (3)$$

Where Θ means the efficiency score for evaluated DMU_0, S^- and S^+ means the input and output slack variables respectively, λ_j means the weights attached to inputs and outputs. With this dual programming it is possible to estimate DMUs' DEA technical efficiency. Here are efficiency judgment criterions of C^2GS^2 model:

Suppose λ^*, S^{*-}, S^{*+}, Θ^* are the optimal solutions of this linear programming problem, then:

1) If $\Theta^*=1$, $S^{*-}=0$ and $S^{*+}=0$, then DMU_0 is DEA efficient (C^2GS^2). That means in the system formed by n DMUs, the output Y_0 has been optimized under the input X_0;

2) If $\Theta^*=1$, DMU_0 is weak DEA efficient (C^2GS^2). That means in the system formed by these n DMUs, for the input X_0, decreasing S^{*-} and maintaining the output Y_0 are needed in order to be optimum;

3) If $\Theta^*<1$, DMU_0 is DEA inefficient (C^2GS^2).

In the case of an inefficient DMU, the DEA models generate nonzero λ, and then the other DMUs are regarded as the role model for DMU_0. Thus, multiple inputs and outputs can be incorporated appropriately with DEA models when calculating technology efficiency and are suitable for analyzing the efficiency of DMUs, especially those that are that complex and have qualitative factors.

B. Extended C^2GS^2 modle-EC^2GS^2

However, the variation of relative efficiency when the output or input of evaluating DMU varies is hardly reflected. For example, the DMU_2 (see table 1 and 2) relative efficiency remains unchanged whereas it should because the input has decreased from 50 to 45. This means the improvement in relative efficiency resulted from technology development is not reflected and also means that this model can be more precise. In order to make more precise efficiency results and reflect the information of DMU more efficiently, an extend DEA model, C^2GS^2 model, was introduced.

Hao Hai, Gu Pei-liang, Lu Qi (2004) [15] put forward the concept of strong-efficiency and propose a new DEA model, EC^2GS^2, derived from C^2GS^2 model by decreasing the constraint conditions of C^2GS^2 model. The EC^2GS^2 model can be obtained by removing the constraint $\frac{U^T Y_0 + u_0}{V^T X_0} \le 1$ from the C^2GS^2 model:

TABLE I

INPUT AND OUTPUT OF DMU

	DMU_1	DMU_2
Input	30	50
Output	5	10

TABLE II

INPUT AND OUTPUT OF DMU

	DMU_1	DMU_2
Input	30	45
Output	5	10

TABLE III

DATA INDEX OF COAL ENTERPRISE OF CHINA

DMU	Input Index			Output Index		
	Total Assets (10^6 Yuan)	Staff Size	R&D Investment (10^6 Yuan)	Net Profit (10^6 Yuan)	Coal Yield (10^6 t)	Social Contribution (%)
DMU_1	188814.00	148000	650.18	13366.62	149.68	18.85
DMU_2	45879.85	88675	489.19	3085.18	71.86	23.20
DMU_3	33782.39	153531	710.52	202.28	56.68	25.57
DMU_4	43911.13	146340	1140.00	156.20	60.81	19.43
DMU_5	25671.48	53863	43.66	416.16	30.06	23.65
DMU_6	20876.11	134407	298.25	594.99	32.06	37.45
DMU_7	35590.89	258852	110.49	694.56	52.71	23.30
DMU_8	24363.29	77100	538.17	463.78	11.08	16.56

$$Max \frac{U^T Y_0 + u_0}{V^T X_0}$$

$$s.t. \begin{cases} \dfrac{U^T Y_j + u_0}{V^T X_j} \le 1 \\ U \ge 0, V \ge 0 \\ u_0 \text{ without symble constraint} \\ j = 1, \ldots, n \end{cases} \tag{4}$$

Equivalent linear programming can be obtained by Charnes-Cooper transformation and introducing non-Archimedean infinitesimal quantity and slack variables the dual programming can be obtained as (5):

$$Min[\theta - \varepsilon(\hat{e}^T S^- + \hat{e}^T S^+)$$

$$(D_\varepsilon) s.t. \begin{cases} \sum_{j=1}^{n} X_j \lambda_j + S^- = \theta X_0 \\ \sum_{j=1}^{n} Y_j \lambda_j - S^+ = Y_0 \\ \sum_{j=1}^{n} \lambda_j = 1 \\ S^-, S^+, \lambda_j \ge 0 \\ j = 1, \ldots, n \end{cases} \tag{5}$$

Suppose λ^*, S^{*-}, S^{*+}, Θ^* are the optimal solutions of this linear programming problem, then:

1) If $\Theta^*=1$, $S^{*-}=0$ and $S^{*+}=0$, DMU_0 is DEA efficient (EC^2GS^2);

2) If $\Theta^*=1$, DMU_0 is weak DEA efficient (EC^2GS^2);

3) If $\Theta^*>1$, DMU_0 is strong DEA efficient (EC^2GS^2).

III. APPLICATION ON COAL ENTERPRISES

Different coal enterprises can be taken as of one type of DMU for comparison since the input and efficiency evaluating index for all coal enterprises are almost the same. In this paper, eight representative coal enterprises were taken as example to show how the application of DEA in efficiency analyzing based on 2006 analysis report on top 100 Chinese coal enterprises [16].

The technical efficiency of coal enterprise is influenced by different factors such as total assets, machining technology, facilities, staff size, management level, and investment in research and development (R&D) [17]. In this paper, according to the demand of DEA, the goal of evaluation and data Availability, three indexes (total assets, staff size, R&D investment) are regarded as input index, and three other indexes (Net profit, coal yield, social contribution) are regarded as output index. Then according to table-1, the proposed method was applied to identify the consistently best performing DMUs, eight effectiveness result of DMUs' calculated by model C^2GS^2 and EC^2GS^2 were given in table-4 and table-5:

From table 4, DMU_1, DMU_2, DMU_3, DMU_5, DMU_6, DMU_7 are DEA (C^2GS^2) technical efficiency. Based on C^2GS^2 model, from the whole, technical efficiency is at good state in each enterprise, even for the inefficient ones, the Θ^* is not too small. For one reason, the technology in mining is advanced; another reason is that each enterprise applies appropriate mining technical measures.

TABLE IV

EVALUATION RESULT WITH MODEL C^2GS^2

C^2GS^2	DMU_1	DMU_2	DMU_3	DMU_4	DMU_5	DMU_6	DMU_7	DMU_8
θ_V^*	1	1	1	0.86	1	1	1	0.99
s_1^{*-}	0	0	0	0	0	0	0	0
s_2^{*-}	0	0	0	0	0	0	0	0
s_3^{*-}	0	0	0	39168.16	0	0	0	42029.50
s_1^{*+}	0	0	0	125653.54	0	0	0	352.75
s_2^{*+}	0	0	0	0	0	0	0	1955.20
s_3^{*+}	0	0	0	6.17	0	0	0	11.04

TABLE `V

EVALUATION RESULT WITH MODEL C^2GS^2

EC^2GS^2	DMU_1	DMU_2	DMU_3	DMU_4	DMU_5	DMU_6	DMU_7	DMU_8
θ_V^*	0.75	1.47	1.08	0.85	4.01	2.38	1.56	0.99
s_1^{*-}	9618223	0	0	0	7243244	1119990	0	0
s_2^{*-}	22678.96	0	59051.84	0	0	166020.2	329291.8	0
s_3^{*-}	0	16948.64	34801.65	39168.16	0	0	0	42029.50
s_1^{*+}	-1028144	15887.46	193312.40	125653.50	24413.43	0	213618.7	352.75
s_2^{*+}	-7782.00	0	0	0	1554.36	2739.43	0	1955.20
s_3^{*+}	4.35	0.41	3.07	6.17	4.52	-12.08	0	11.04

But when EC^2GS^2 is applied, the result shows enormous technical efficiency differences among different DMUs. The DMU_5 is most efficient, as presented by model EC^2GS^2, the technical efficiency remains as 1 while the input is 4.01 times larger. The DMU_1 seems to be of the least efficient since its technical efficient remains to be 1 while the input has decreased by 25 %. Obviously, In the light of different efficiency, DMUs will be ordered at a sensible sequence, that is: DMU_5, DMU_6, DMU_7, DMU_2, DMU_3, DMU_4, DMU_8, DMU_1.

As indicated in Table 4 and 5, further comparison for the DMU of weak efficient can be made.

IV. DISCUSSION

As an estimating tool, DEA is something like "black box" which has made it convenient for estimating work. However, some shortcomings are still with the "black box" and DEA is limited by these short comings in the actual use. For example, the front line face estimated according to the partial liner estimation often result in unreasonable weights and too many efficient DUMs. And the assumption of identity of Units makes it impossible to estimate units with different external properties.

To get rid of the existing shortcomings, new theories were introduced to the DEA to explore new and more efficient ways to estimate performance efficiency. More objective evaluation mechanism and reasonable weights is in need. DEA and other method were cooperated to establish integrate estimate model for specific task, such as proposing new DEA model by introducing game theory model [17][18] of which "Max-Min" is typical model [19]. In this way, the theory can be developed further and this is important for theory study, what is more, the market prospect is promising and can be more likely applied practically.

V. CONCLUSION

In this paper, DEA was applied to measure the efficiency for coal enterprise with a series of complex characteristics, for instances, it has a multiple inputs and outputs structured system, is hard to set evaluating index and difficult to get sound evaluation when the evaluation function is unknown. In such a complexion, DEA model will yield relatively efficient enterprise by comparing each similar type DMU. In addition, by analyzing coal enterprise's redundant input and deficient output, DEA affords the direction to optimize its technical inefficiency.

The structure and principle of C^2GS^2 and EC^2GS^2 of DEA was introduced firstly in this paper; the defects of C^2GS^2 were pointed out. Then in the case study, basing on eight Chinese coal enterprises as DMUs, the two models were applied to estimate DMUs technical performance. By comparing C^2GS^2 model with EC^2GS^2 model, the relation between them, advantages and defects of both models were studied, and the advantage of EC^2GS^2 are mainly shown in the following facets: (1). The relative efficiency in EC^2GS^2 model can excess 1, and the issue of not enough or excess can be reflected by this mode but this is impossible for C^2GS^2 model. (2) The relative efficiency varies when the output (input) of evaluation DMU in EC^2GS^2 model is changed whereas those of the C^2GS^2 model cannot. (3) The DMUs whose relative efficiencies is 1 can be compared with EC^2GS^2 model whereas cannot with C^2GS^2. In sum, compared with model C^2GS^2, model EC^2GS^2 make a much more precise evaluating result on technical efficiency evaluation of coal enterprise.

REFERENCES

[1] A. Charnes, W. W. Cooper and E. Rhodes, "Measuring the Efficiency of Decision Making Units," *European Journal of Operational Research*, vol. 2, no. 6, pp. 429-444, Nov. 1978.

[2] R. D. Banker, A. Charnes and W. W. Cooper, "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis," *Management Science*, vol. 30, no. 9, pp. 1078-1092, Sep. 1984.

[3] R. G. Thompson, F. D. Singleton, R. M. Thrall, B. A. Smith and M. Wilson, "Comparative Site Evaluations for Locating a High-Energy Physics Lab in Texas," *Interfaces*, vol. 16, no. 6, pp. 35-49, Nov. – Dec. 1986.

[4] A. Charnes, W. W. Cooper, Q. L. Wei and Z. M. Huang, "Cone Ratio Data Envelopment Analysis and Multi-objective Programming," *International Journal of Systems Science*, vol. 20, no. 7, pp. 1099-1118, Jul. 1989.

[5] A. Charnes, W. W. Cooper, Z. M. Huang and D. B. Sun, "Polyhedral Cone-Ratio DEA Models with an Illustrative Application to Large Commercial Banks," *Journal of Econometrics*, vol. 46, no. 1-2, pp. 73-91, Oct. – Nov. 1990.

[6] K. C. Land, C. A. K. Lovell and S. Thore, "Productive Efficiency under Capitalism and State Socialism: The Chance-Constrained Programming Approach," presented at the *47th Congress of the International Institute of Public Finance*, St Petersburg, 1992.

[7] K. C. Land, C. A. K. Lovell and S. Thore, "Productive Efficiency under Capitalism and State Socialism: An Empirical Inquiry Using Chance-Constrained Data Envelopment Analysis," *Technological Forecasting and Social Change*, vol. 46, no. 2, pp. 139-152, Jun. 1994.

[8] O. B. Olesen and N. C. Petersen, "Chance Constrained Efficiency Evaluation," *Management Science*, vol. 41, no. 3, pp. 442-457, Mar. 1995.

[9] W. W. Cooper, Z. M. Huang, V. Lelas, S. X. Li and O. B. Olesen, "Chance Constrained Programming Formulations for Stochastic Characterizations of Efficiency and Dominance in DEA," *Journal of Productivity Analysis*, vol. 9, no. 1, pp. 53-79, Jan. 1998.

[10] J. A. Clark, "Economic Cost, Scale Efficiency, and Competitive Viability in Banking," *Journal of Money Banking and Credit*, vol. 28, no. 3, pp. 342-364, Aug. 1996.

[11] N. Adler and B. Golany, "Evaluation of Deregulated Airline Networks Using Data Envelopment Analysis Combined with Principal Component Analysis with an Application to Western Europe," *European Journal of Operational Research,* vol. 132, no. 2, pp. 260-273, Jul. 2001.

[12] G. Q. Xiong and J. P. Cheng, "Application of DEA in Performance Evaluation of Urban Community Construction" (in Chinese), *Operations Research and Management Science*, vol. 12, no. 4, pp. 120-123, Aug. 2003.

[13] K. Z. Yang and X. Xie, "The Analysis about DEA-Efficiency of China Cities Input-Output" (in Chinese), *Geography and Territorial Research*, vol. 18, no. 3, pp. 45-47, Aug. 2002.

[14] S. J. Weng, T. Wu, J. Blackhurst and G. Mackulak, "An Extended DEA Model for Hospital Performance Evaluation and Improvement," *Health Services and Outcomes Research Methodology*, vol. 9, no. 1, pp. 39-53, Mar. 2009.

[15] H. Hao, P. L. Gu and Q. Lu, "Extended Data Envelopment Analysis Model Based on C^2GS^2 Model - EC^2GS^2" (in Chinese), *Journal of Systems Engineering*, vol. 19, no. 1, pp. 1-7, Feb. 2004.

[16] F. M. Tao, *2006 Analysis Report on Top 100 Chinese Coal Enterprises*. Beijing: Economic Science Press, 2007, pp. 12-20, 76-80.

[17] G. J. Yang and L. J. Wang, "Empirical Analysis on Relative Efficiency of Coal Enterprise Based on DEA" (in Chinese), *China Coal*, vol. 33, no. 10, pp. 20-23, Oct. 2007.

[18] R. D. Banker, "A Game Theoretic Approach to Measuring Efficiency," *European Journal of Operational Research*, vol. 5, no. 4, pp. 262-266, Oct. 1980.

[19] J. J. Rousseau and J. H. Semple, "Two-Person Ratio Efficiency Games," *Management Science*, vol. 41, no. 3, pp. 435-441, Mar. 1995.

[20] D. S. Wu, Q. Shi and M. Wang, "Max-Min DEA Model and Corresponding Nash Bargaining Solution" (in Chinese), *Journal of Management Sciences in China*, vol. 8, no. 5, pp. 90-94, Oct 2005.