Data envelopment analysis: an efficient duo linear programming approach

Saber Saati

Department of Mathematics,
Tehran-North Branch, Islamic Azad University,
P.O. Box 19585-936,
Tehran, Iran
E-mail: s_saatim@iau-tnb.ac.ir

Adel Hatami-Marbini

Louvain School of Management,
Center of Operations Research and Econometrics (CORE),
Universite Catholique de Louvain,
34 Voie du Roman Pays,
Louvain-la-Neuve B-1348, Belgium
Fax: +32 10 47 4301
E-mail: adel.hatamimarbini@uclouvain.be
E-mail: adel_hatami@yahoo.com

Madjid Tavana*

Lindback Distinguished Chair of Information Systems,
La Salle University,
Philadelphia, PA 19141, USA
Fax: +1 267 295 2854
E-mail: tavana@lasalle.edu
*Corresponding author

Abstract: Data envelopment analysis (DEA) is a powerful mathematical method that utilises linear programming (LP) to determine the relative efficiencies of a set of functionally similar decision-making units (DMUs). Evaluating the efficiency of DMUs continues to be a difficult problem to solve, especially when the multiplicity of inputs and outputs associated with these units is considered. Problems related to computational complexities arise when there are a relatively large number of redundant variables and constraints in the problem. In this paper, we propose a three-step algorithm to reduce the computational complexities and costs in the multiplier DEA problems. In the first step, we identify some of the inefficient DMUs through input–output comparisons. In the second step, we specify the efficient DMUs by solving a LP model. In the third step, we use the results derived from the second step and another LP model to obtain the efficiency of the inefficient DMUs. We also present a numerical example to demonstrate the applicability of the proposed framework and exhibit the efficacy of the procedures and algorithms.

Keywords: DEA; data envelopment analysis; duo linear programming; DMU; decision-making unit; multiplier; efficiency determination.

Reference to this paper should be made as follows: Saati, S., Hatami-Marbini, A. and Tavana, M. (2011) 'Data envelopment analysis: an efficient duo linear programming approach', *Int. J. Productivity and Quality Management*, Vol. 7, No. 1, pp.90–103.

Biographical notes: Saber Saati is an Assistant Professor of Mathematics and Chairman of the Department of Mathematics at Islamic Azad University (IAU) – Tehran North Branch in Iran. He received his PhD in Operations Research at Science and Research Branch of IAU in 2002. His research interests include data envelopment analysis, fuzzy programming and fuzzy MADM. He has published in journals such as *Ricerca Operativa, Fuzzy Optimization and Decision Making, Journal of Interdisciplinary Mathematics, Far East Journal of Applied Mathematics, Applied Mathematics and Computation, Iranian Journal of Fuzzy Systems, Advance in Fuzzy Sets and Systems, Int. J. Mathematical Analysis, Applied Mathematical Science, Journal of Industrial Engineering International*, *Australian Journal of Basic and Applied Sciences* and *Journal of Applied Sciences*, Mathematical Sciences and Applied Soft Computing.

Adel Hatami-Marbini is a Doctoral Student and Research Assistant at the Louvain School of Management and the Center for Operations Research and Econometrics (CORE) of the Université Catholique de Louvain (UCL) in Belgium. He holds an MSc and a BSc in Industrial Engineering from Islamic Azad University in Iran. His research interests are in operations research, data envelopment analysis, multi criteria decision making and fuzzy sets theory. He has published in journals such as *Computers and Industrial Engineering*, *Applied Soft Computing, Journal of Applied Sciences*, *Applied Mathematical Sciences*, *Int. J. Mathematics in Operational Research*, *Int. J. Operations Research and Information Systems, Int. J. Information and Decision Sciences, Industrial Engineering International*, etc.

Madjid Tavana is a Professor of Management Information Systems and Decision Sciences and the Lindback Distinguished Chair of Information Systems at La Salle University where he served as a Chairman of Department of Management and Director of the Center for Technology and Management. He has been a distinguished Research Fellow at NASA's Kennedy Space Center, NASA's Johnson Space Center, Naval Research Laboratory – Stennis Space Center and Air Force Research Laboratory. He was awarded the prestigious Space Act Award by NASA in 2005. He holds an MBA, a PMIS and a PhD in Management Information Systems. He received his Post-Doctoral Diploma in strategic information systems from the Wharton School in the University of Pennsylvania. He is the Editor-in-Chief for the *Int. J. Strategic Decision Sciences*, the *Int. J. Enterprise Information Systems*, and the *International Journal of Applied Decision Sciences*. He has published in journals such as *Decision Sciences*, *Interfaces*, *Information Systems*, *Annals of Operations Research*, *Information and Management*, *Journal of the Operational Research Society*, *Computers and Operations Research*, *Advances in Engineering Software*, etc.

1 Introduction

Efficiency measurement in organisations has enjoyed a great deal of interest among researchers studying performance analysis. Data envelopment analysis (DEA) has been successfully employed to study the comparative performance of units that consume similar inputs to produce similar outputs. Generally, these units are referred to as decision-making units (DMUs). Charnes et al. (1978) originally proposed the first DEA model known as the CCR model. DEA is non-parametric and utilises linear programming (LP) to measure the relative efficiency of the DMUs without *a priori* specification of input and output weights (or multipliers). A score of one is assigned to the frontier (efficient) units. The frontier units in DEA are those with maximum output levels for given input levels or with minimum input levels for given output levels.

DEA is used to compare the efficiency of DMUs in terms of several inputs and outputs. Within the DEA context, problems of discrimination between efficient and inefficient DMUs often arise, particularly when there are a relatively large number of variables. Consequently, there is a direct correlation between the number of DMUs and the number of LP models needed to solve a DEA problem. The greater the number of input and output variables in a DEA, the higher is the dimensionality of the LP solution space and the less discerning is the analysis (Jenkins and Anderson, 2003). This phenomenon is particularly noticeable in the multiplier form of DEA where the number of constraints could be as much as the number of DMUs. In this study, we propose a three-step algorithm to reduce the computational complexities in the multiplier DEA problems.

This paper is organised as follows. Section 2 presents the literature review. Section 3 provides some preliminaries and definitions and in Section 4, we present an overview of the CCR models. Section 5 describes the proposed algorithm. In Section 6, we present a numerical example to demonstrate the applicability of the proposed algorithm and in Section 7, we conclude with our conclusions and future research directions.

2 Literature review

The number of efficient DMUs in DEA relies on the number of input and output variables (Serrano Cinca and Mar Molinero, 2004). As a result, a greater number of variables in DEA results in less discerning analysis (Jenkins and Anderson, 2003). Therefore, it is useful to reduce the dimensionality or the number of variables in a DEA structure. Two methodologies are suggested in the literature as potential approaches to reduce the dimensionality or the number of variables without requiring additional preferential information in DEA. The two approaches are principal component analysis combined with DEA (PCA–DEA) and variable reduction based on a partial covariance analysis. Adler and Golany (2002) used PCA–DEA to calculate the input- or output-oriented principal components to, respectively, replace the original output or input data in the DEA model. Similar PCA–DEA approaches have been applied by Adler and Golany (2001), Adler and Berechman (2001) and Serrano Cinca and Mar Molinero (2004). Shanmugam and Johnson (2007) provide a complete review of the PCA–DEA methods in the literature.

The PCA–DEA method assumes that separation of variables representing similar themes, and the removal of principal components with little or no explanatory power, improves the categorisation of efficient and inefficient DMUs. Adler and Yazhemsky (2010) used a simulation technique to generalise the comparison between the PCA–DEA method and the variable reduction based on a partial covariance analysis. Monte Carlo simulation was used to generate a large number of DMUs, based on various production functions, inefficiency distributions and correlation between variables and sample sizes. Furthermore, to ensure the generality of their conclusions, they analysed various forms of misspecification of the DEA. They concluded that PCA–DEA provides a more powerful tool than variable reduction with consistently more accurate results. PCA–DEA was applied to all basic DEA models and guidelines for its application were presented. Nevertheless, Adler and Yazhemsky (2010) were only able to show that the PCA–DEA method is useful when analysing relatively small datasets, removing the need for additional preference information.

Several studies have proposed limiting the number of variables relative to the number of DMUs. In general, the number of input and output variables in the DEA model should be no more than one-third of the number of DMUs (Boussofiane et al., 1991; Friedman and Sinuany-Stern, 1998). Nunamaker (1985) has shown that adding variables to a DEA model will expand the set of efficient DMUs. Golany and Roll (1989) also showed that a large number of variables in the model tend to shift the DMUs under consideration towards the efficient frontier resulting in a relatively large number of efficient DMUs.

Other DEA studies have suggested various methods for identifying the most relevant variables to be included in the DEA model. One approach involved judgemental screening of the list of variables by expert opinion (Golany and Roll, 1989). Another commonly used approach for reducing the number of variables in a DEA model is regression and correlation analysis (Lewin et al., 1982). This approach suggests that variables which are highly correlated are generally redundant and should not be included in the model (Chilingerian, 1995; Salinas-Jimenez and Smith, 1996). Another application of variable selection based on correlating the efficiency scores was proposed by Sigala et al. (2004). Golany and Roll (1989) have shown that regression tests on the inputs and outputs should not be regarded as a reliable rule for removing variables from a DEA model. Similarly, Jenkins and Anderson (2003) have suggested that an analysis of simple correlation is insufficient in identifying unimportant variables and removing even highly correlated variables could have a significant impact on the efficiency scores.

In contrast to correlation analysis, other approaches look directly at the effect on efficiency scores as input and output DEA variables are changed. Statistical tests developed by Banker (1993, 1996) have been used to evaluate the marginal impact on the efficiencies of when adding or removing a given variable. Kittelson (1993) presented an iterative method for building a DEA model involving statistical procedures while focusing on evaluating the statistical significance of the changes in the efficiencies. A similar methodology for variable selection involving statistical procedures was proposed by Pastor et al. (2002). They used statistical tests to determine the significance of the efficiency contribution of a particular variable in the reduced and extended forms of the DEA model.

Other researchers have studied this problem from different perspectives. Alirezaee and Afsharian (2007) argued that DMUs with extraordinary output lead to a monopoly in the reference set and introduced an algorithm for screening the DMUs into efficient layers. Wagner and Shimshaka (2007) focused on the choice of input and output variables and developed a stepwise selection algorithm to measure the effect or influence of variables directly on the efficiencies by considering their average change as variables are added or dropped from the analysis. Their method was intended to produce DEA models that include only those variables with the largest impact on the DEA results. Amin and Toloo (2007) proposed an integrated efficient DEA model to identify the most CCR-efficient DMU without using a trial and error method and without solving one LP model for each DMU. They showed that the improved integrated DEA model is always feasible and capable to rank the most efficient DMUs where the return to scale was constant. Amin (2009) showed that the integrated DEA model proposed by Amin and Toloo (2007) may obtain more than one efficient DMU and proposed an improved version of their model. Toloo and Nalchigarb (2009) later proposed a computationally efficient model with a wider range of applications to find the most BCC-efficient DMU by solving only one LP model where the return to scale was variable.

There are several DEA applications that involve large data. Barr and Durchholz (1997) reported results from banking using 8,000 DMUs, Wilson and Wheeloc (2000) worked with bank data and solved DEA problems with 15,000 DMUs and Dulá (2008) reported on a comprehensive computational study involving DEA problems with up to 100,000 DMUs. Chen and Cho (2009) proposed an accelerating procedure that identified a few 'similar' critical DMUs to compute DMU efficiency scores in a given set. Simulation results demonstrated that the proposed procedure was suitable for solving large-scale BCC problems when the percentage of efficient DMUs was high. Soleimani-Damaneh (2009) sketched an LP model with fewer constraints through some linear transformations and artificial data. They showed that the efficiency measure obtained, under the artificial input–output levels, using a reduced model is equal to the one obtained under the original data.

While it is advantageous to limit the number of input and output variables in DEA and reduce the computational complexities, there is no consensus on how best to do this (Wagner and Shimshaka, 2007). In this paper, we propose a three-step algorithm to reduce the computational complexities of the DEA problems. Initially, we identify the inefficient DMUs through input–output comparisons. Next, we determine the efficient DMUs by solving a LP model. Finally, we obtain the efficiency of the inefficient DMUs using a LP model and the results derived from the previous step.

3 Preliminary definitions

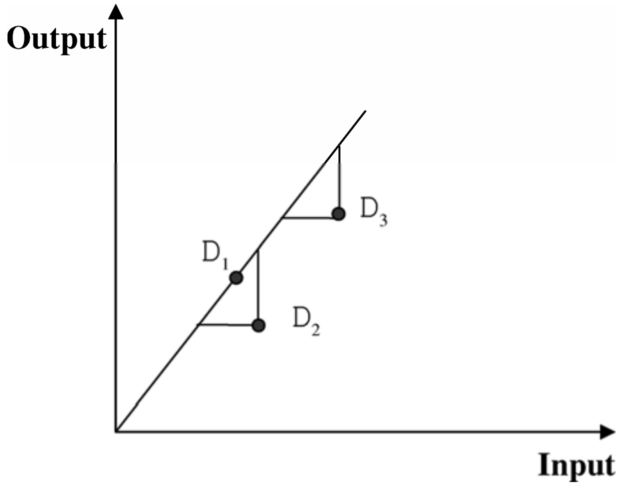
Assume a set of n DMUs, with each DMU$_j$ $(j = 1, 2, \dots, n)$ using m inputs x_{ij} $(i = 1, 2, \dots, m)$ to produce s outputs y_{rj} $(r = 1, 2, \dots, s)$.

Definition: *DMU$_p$ is a non-dominated unit if we cannot find a DMU$_k$ with the following properties:*

1. $x_{ik} < x_{ip}$ $(i = 1, 2, \dots, m)$
2. $y_{rk} > y_{rp}$ $(r = 1, 2, \dots, s)$

For instance, the CCR frontier of three DMUs, D_1, D_2 and D_3, is shown in Figure 1. DMU D_2 is a dominated by DMU D_1 which is able to produce more output with less input. In this example, DMUs D_1 and D_3 are non-dominated units and DMU D_2 is a dominated DMU. The efficient DMUs are non-dominated but non-dominated DMUs are not necessarily efficient. For example, as shown in Figure 1, DMU D_3 is a non-dominated unit but it is not efficient.

Figure 1 An example of non-dominated DMUs



4 The CCR models

Consider a set of n DMUs to be evaluated where each DMU$_j$ ($j = 1, 2, \ldots, n$) uses m inputs $x_{ij} (i = 1, 2, \ldots, m)$ to generate s outputs $y_{rj} (r = 1, 2, ..., s)$. Charnes et al. (1978) proposed the CCR model to evaluate the efficiency of a given DMU$_p$. The primal and dual LP statements for the (input oriented) CCR model are:

$$
\begin{aligned}
&\textit{Primal CCR model (input-oriented)} \\
&\min \quad \theta \\
&\text{s.t.} \quad \theta x_{ip} - \sum_{j=1}^{n} \lambda_j x_{ij} \geq 0, \quad \forall i \\
&\qquad \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{rp}, \quad \forall r \\
&\qquad \lambda_j \geq 0, \quad \forall j \\
&\textit{Dual CCR model (input-oriented)} \\
&\max \quad \sum_{r=1}^{s} u_r y_{rp} \\
&\text{s.t.} \quad \sum_{i=1}^{m} v_i x_{ip} = 1 \\
&\qquad \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \quad \forall j \\
&\qquad u_r, v_i \geq 0, \quad \forall r, i
\end{aligned}
\tag{1}
$$

where u_r and v_i are the weights assigned to the *r*th output and the *i*th input, respectively. The primal and dual models are referred to as the envelopment and the multiplier, respectively.

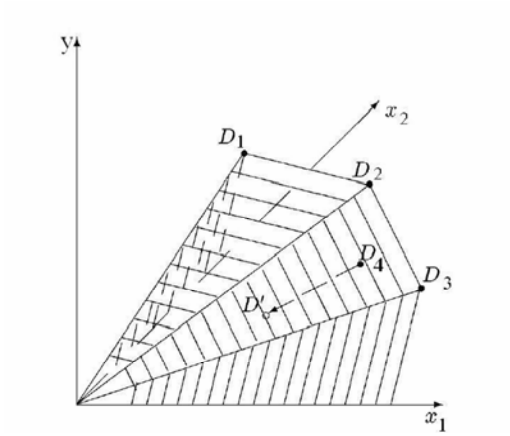
In the envelopment form of DEA, the DMUs are compared with a frontier constructed by a combination of $\left(\sum_{j=1}^{n}\lambda_j x_{ij}, \sum_{j=1}^{n}\lambda_j y_{rj}\right)$ $(\forall i, r)$. According to the envelopment model, if the *j*th DMU is efficient, λ_j is positive and the DMU lies on the frontier. On the other hand, if the *j*th DMU is inefficient, λ_j is zero. As a result, by determining the inefficient DMUs, we can reduce the number of decision variables of the primal CCR by the number of inefficient DMUs.

In the multiplier model (the dual of the envelopment model), the reduction of the number of the decision variable in the envelopment model will result in an equivalent reduction of the constraints in the multiplier model. The eliminated constraints correspond to the inefficient DMUs and do not have any effects on the construction of the efficient frontier. After eliminating the extra constrains, the LP model is solved to obtain the equations of the efficient frontier. Then the efficiency of the inefficient DMUs is evaluated in comparison with the efficient frontier.

5 Proposed algorithm

Let us first consider the envelopment form of the CCR model. In the evaluation of the *j*th DMU, the optimal value of λ_j is positive if the *j*th DMU lies on the segment of the efficient frontier which is constructed by the efficient DMU_j. For example, Figure 2 shows the CCR frontier of four DMUs, D_1, D_2, D_3 and D_4 with two inputs and one output. D_1, D_2 and D_3 are efficient while D_4 is inefficient. In the evaluation of D_4, this unit lies on the segment of the efficient frontier which is constructed by D_2 and D_3. Therefore, λ_2 and λ_3 are positive in the optimal solution of the envelopment form of the CCR model and λ_1 and λ_4 are zero.

Figure 2 An example of a CCR frontier



Accordingly, we consider the following model:

$$
\begin{aligned}
\min \quad & \sum_{j=1}^{n}\theta_j \\
\text{s.t.} \quad & \theta_1 x_{i1} - \sum_{j=1}^{n}\lambda_j^{(1)} x_{ij} \ge 0, \quad \forall i \\
& \sum_{j=1}^{n}\lambda_j^{(1)} y_{rj} \ge y_{r1}, \quad \forall r \\
& \vdots \\
& \theta_k x_{ik} - \sum_{j=1}^{n}\lambda_j^{(k)} x_{ij} \ge 0, \quad \forall i \\
& \sum_{j=1}^{n}\lambda_j^{(k)} y_{rj} \ge y_{rk}, \quad \forall r \\
& \vdots \\
& \theta_n x_{in} - \sum_{j=1}^{n}\lambda_j^{(n)} x_{ij} \ge 0, \quad \forall i \\
& \sum_{j=1}^{n}\lambda_j^{(n)} y_{rj} \ge y_{rn}, \quad \forall r \\
& \lambda_j^{(k)} \ge 0, \quad j,k = 1,\dots,n
\end{aligned}
\tag{2}
$$

Based on the decomposition algorithm, the solution of model (2) is identical to the solution of the envelopment form of the CCR model. Hence we can solve only one LP model instead of solving *n* LP models needed using the CCR model. The proposed model (2), which is a large LP model, consists of $n(n + 1)$ variables and $2n(m + s)$ constraints.

As described earlier, several variables in model (2) are zero and several constraints corresponding to the inefficient DMUs are redundant. While comparing the DMUs with each other, the DMU_p under consideration that has more inputs than the specific DMU_k $(x_{ip} \ge x_{ik})$ and less outputs than DMU_k $(y_{rp} < y_{rk})$ is inefficient. However, the inverse case is not true. As a result, loads of the inefficient DMUs can be determined without solving the LP problem. The proposed algorithm here can be described by the following nine steps:

Step 1 Assume a set of DMU index $(J = \{1, 2, \dots, n\})$.

Step 2 Determine the inefficient DMUs by comparing the inputs and outputs of the DMUs with each other.

Step 3 Place the index of the inefficient DMUs in J_N.

Step 4 Consider $J_{EN} = J - J_N$.

Step 5 Utilise the proposed model (2) for the DMUs which consist of the J_{EN} index. Hence the model is expressed as:

$$\begin{aligned}
\min \quad & \sum_{j \in J_{EN}} \theta_j \\
\text{s.t.} \quad & \theta_k x_{ik} - \sum_{j \in J_{EN}} \lambda_j^{(k)} x_{ij} \geq 0, \quad \forall i, \ \forall k \in J_{EN} \\
& \sum_{j \in J_{EN}} \lambda_j^{(k)} y_{rj} \geq y_{rk}, \quad \forall r, \ \forall k \in J_{EN} \\
& \lambda_j^{(k)} \geq 0, \quad j, k \in J_{EN}
\end{aligned} \tag{3}$$

Step 6 Determine a set of efficient DMUs, denoted by J_E, after solving the model (3).

Step 7 Consider $J_I = J - J_E$.

Step 8 Utilise the proposed model (2) for the DMUs which consist of the J_I index by using the efficient DMUs which is obtained in the sixth step. Hence the model is expressed as:

$$\begin{aligned}
\min \quad & \sum_{j \in J_I} \theta_j \\
\text{s.t.} \quad & \theta_k x_{ik} - \sum_{j \in J_E} \lambda_j^{(k)} x_{ij} \geq 0, \quad \forall i, \ \forall k \in J_I \\
& \sum_{j \in J_E} \lambda_j^{(k)} y_{rj} \geq y_{rk}, \quad \forall r, \ \forall k \in J_I \\
& \lambda_j^{(k)} \geq 0, \quad j \in J_E, k \in J_I
\end{aligned} \tag{4}$$

Step 9 Obtain the efficiency of the inefficient DMUs using model (4).

6 Numerical example

In this section, we present an example to demonstrate the applicability of the proposed framework and exhibit the efficacy of the procedures and algorithms. Let us consider the problem presented in Table 1 with five DMUs, two inputs and two outputs.

The proposed method is utilised to solve the problem presented in Table 1. The computational procedure for this problem is summarised below:

Step 1 $J = \{1, 2, 3, 4, 5\}$ representing five DMUs.

Steps 2 and 3 By comparing the inputs and outputs of the DMUs to each other, DMUs 3, 4 and 5 are considered inefficient in comparison with DMU 2. Hence $J_N = \{3,4,5\}$ represents the inefficient DMUs.

Step 4 We consider $J_{EN} = \{1,2\}$.

Step 5 Model (3) is formulated for $J_{EN} = \{1,2\}$ as follows:

$$
\begin{aligned}
\min \quad & \theta_1 + \theta_2 \\
s.t. \quad & 10\theta_1 - 10\lambda_1^{(1)} - 12\lambda_2^{(1)} \geq 0 \\
& 17\theta_1 - 17\lambda_1^{(1)} - 25\lambda_2^{(1)} \geq 0 \\
& 73\lambda_1^{(1)} + 75\lambda_2^{(1)} \geq 73 \\
& 5\lambda_1^{(1)} + 4\lambda_2^{(1)} \geq 5 \\
& 12\theta_2 - 10\lambda_1^{(2)} - 12\lambda_2^{(2)} \geq 0 \\
& 25\theta_2 - 17\lambda_1^{(2)} - 25\lambda_2^{(2)} \geq 0 \\
& 73\lambda_1^{(2)} + 75\lambda_2^{(2)} \geq 75 \\
& 5\lambda_1^{(2)} + 4\lambda_2^{(2)} \geq 4 \\
& \lambda_1^{(1)}, \lambda_2^{(1)}, \lambda_1^{(2)}, \lambda_2^{(2)} \geq 0
\end{aligned}
$$

Step 6 After running the above model, the results presented in Table 2 are obtained:

As shown in Table 2, the DMU 2 is also identified as inefficient and the only efficient DMU in this evaluation problem is DMU 2.

Steps 7 and 8 The model (4) for $J_I = \{3,4,5\}$ is as follows:

$$
\begin{aligned}
\min \quad & \theta_3 + \theta_4 + \theta_5 \\
\text{s.t.} \quad & 13\theta_3 - 10\lambda_1^{(3)} \geq 0 \\
& 31\theta_3 - 17\lambda_1^{(3)} \geq 0 \\
& 73\lambda_1^{(3)} \geq 68 \\
& 5\lambda_1^{(3)} \geq 3 \\
& 20\theta_4 - 10\lambda_1^{(4)} \geq 0 \\
& 41\theta_4 - 17\lambda_1^{(4)} \geq 0 \\
& 73\lambda_1^{(4)} \geq 25 \\
& 5\lambda_1^{(4)} \geq 2 \\
& 20\theta_5 - 10\lambda_1^{(5)} \geq 0 \\
& 40\theta_5 - 17\lambda_1^{(5)} \geq 0 \\
& 73\lambda_1^{(5)} \geq 20 \\
& 5\lambda_1^{(5)} \geq 1 \\
& \lambda_1^{(3)}, \lambda_2^{(3)}, \lambda_1^{(4)}, \lambda_2^{(4)}, \lambda_1^{(5)}, \lambda_2^{(5)} \geq 0
\end{aligned}
$$

Step 9 After running the above model, the efficiency of the inefficient DMUs 3, 4 and 5 presented in Table 3.

Table 1 The numerical problem

DMU	*Input 1*	*Input 2*	*Output 1*	*Output 2*
1	10	17	73	5
2	12	25	75	4
3	13	31	68	3
4	20	41	25	2
5	20	40	20	1

Table 2 The numerical results from model (3)

DMU	*Efficiency*	λ
1	1	$\lambda_1^{(1)} = 1,\ \lambda_2^{(1)} = 0$
2	0.86	$\lambda_1^{(2)} = 1.027,\ \lambda_2^{(2)} = 0$

Table 3 The numerical results from model (4)

DMU	*Efficiency*	λ
3	0.72	$\lambda_1^{(3)} = 0.93$
4	0.20	$\lambda_1^{(4)} = 0.4$
5	0.14	$\lambda_1^{(5)} = 0.27$

In addition, we evaluate this example with the primal CCR model to demonstrate the validity of the proposed algorithm. For instance, the primal (input oriented) CCR model for DMU 2 is as follows:

$$\begin{aligned}
\min\quad & \theta \\
\text{s.t.}\quad & 12\theta - \left(10\lambda_1 + 12\lambda_2 + 13\lambda_3 + 20\lambda_4 + 20\lambda_5\right) \geq 0 \\
& 25\theta - \left(17\lambda_1 + 25\lambda_2 + 31\lambda_3 + 41\lambda_4 + 40\lambda_5\right) \geq 0 \\
& \left(73\lambda_1 + 75\lambda_2 + 68\lambda_3 + 25\lambda_4 + 20\lambda_5\right) \geq 75 \\
& \left(5\lambda_1 + 4\lambda_2 + 3\lambda_3 + 2\lambda_4 + 1\lambda_5\right) \geq 4 \\
& \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5 \geq 0
\end{aligned}$$

Using the above model, we obtain the efficiency values of DMUs 1, 2, 3, 4 and 5 as 1, 0.86, 0.72, 0.20 and 0.14, respectively. These efficiencies are identical with the efficiencies obtained from our model.

Therefore, we could calculate the efficiency of the DMUs by using a number of comparisons and solving two LP models based on the non-dominated DMUs and the inherent nature of the CCR model.

7 Conclusions

The field of DEA has grown exponentially since the pioneering papers of Farrell (1957) and Charnes et al. (1978). DEA measures the relative efficiency of a DMU by comparing it against a peer group. The most popular DEA models are computationally easy and can be solved with standard LP models. However, evaluating the efficiency of DMUs continues to be a difficult problem to solve when the multiplicity of inputs and outputs associated with the DMUs is considered (Parkan, 2006). Problems related to computational complexities arise when there are a relatively large number of redundant variables and constraints in the problem. Generally, these redundant variables and constraints correspond to inefficient DMUs.

There are many studies in the DEA literature that have focused on the computational requirements of DEA models and provided some ways to reduce them. We proposed a three-step algorithm to reduce the computational complexities of the DEA problems by identifying the inefficient DMUs through input–output comparisons. We further used two LP models to identify the efficient DMUs and to obtain the efficiency of the inefficient DMUs. We also presented a numerical example and demonstrated the applicability of the proposed framework.

Acknowledgements

The authors would like to thank the anonymous reviewers and the editor for their insightful comments and suggestions.

References

Adler, N. and Berechman, J. (2001) 'Measuring airport quality from the airlines' viewpoint: an application of data envelopment analysis', *Transport Policy*, Vol. 8, No. 3, pp.171–181.

Adler, N. and Golany, B. (2001) '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.

Adler, N. and Golany, B. (2002) 'Including principal component weights to improve discrimination in data envelopment analysis', *Journal of the Operations Research Society*, Vol. 53, No. 9, pp.985–991.

Adler, N. and Yazhemsky, E. (2010) 'Improving discrimination in data envelopment analysis: PCA–DEA or variable reduction', *European Journal of Operational Research*, Vol. 202, No. 1, pp.273–284.

Alirezaee, M.R. and Afsharian, M. (2007) 'Model improvement for computational difficulties of DEA technique in the presence of special DMUs', *Applied Mathematics and Computation*, Vol. 186, No. 2, pp.1600–1611.

Amin, G.R. (2009) 'Comments on finding the most efficient DMUs in DEA: an improved integrated model', *Computers and Industrial Engineering*, Vol. 56, No. 4, pp.1701–1702.

Amin, G.R. and Toloo, M. (2007) 'Finding the most efficient DMUs in DEA: an improved integrated model', *Computers and Industrial Engineering*, Vol. 52, No. 1, pp.71–77.

Banker, R.D. (1993) 'Maximum likelihood, consistency and data envelopment analysis: a statistical foundation', *Management Science*, Vol. 39, No. 10, pp.1265–1273.

Banker, R.D. (1996) 'Hypothesis tests using data envelopment analysis', *Journal of Productivity Analysis*, Vol. 7, Nos. 2–3, pp.139–159.

Barr, R.S. and Durchholz, M.L. (1997) 'Parallel and hierarchical decomposition approaches for solving large-scale data envelopment analysis models', *Annals of Operations Research*, Vol. 73, pp.339–372.

Boussofiane, A., Dyson, R.G. and Thanassoulis, E. (1991) 'Applied data envelopment analysis', *European Journal of Operational Research*, Vol. 52, No. 1, pp.1–15.

Charnes, A., Cooper, W.W. and Rhodes, E.L. (1978) 'Measuring the efficiency of decision making units', *European Journal of Operational Research*, Vol. 2, No. 6, pp.429–444.

Chen, W.C. and Cho, W.J. (2009) 'A procedure for large-scale DEA computations', *Computers and Operations Research*, Vol. 36, No. 6, pp.1813–1824.

Chilingerian, J.A. (1995) 'Evaluating physician efficiency in hospitals: a multivariate analysis of best practices', *European Journal of Operational Research*, Vol. 80, No. 3, pp.548–574.

Dulá, J.H. (2008) 'A computational study of DEA with massive data sets', *Computers and Operations Research*, Vol. 35, No. 4, pp.1191–1203.

Farrell, M.J. (1957) 'The measurement of productive efficiency', *Journal of the Royal Statistical Society A*, Vol. 120, No. 3, pp.253–281.

Friedman, L. and Sinuany-Stern, Z. (1998) 'Combining ranking scales and selecting variables in the DEA context: the case of industrial branches', *Computers and Operations Research*, Vol. 25, No. 9, pp.781–791.

Golany, B. and Roll, Y. (1989) 'An application procedure for DEA', *OMEGA*, Vol. 17, No. 3, pp.237–250.

Jenkins, L. and Anderson, M. (2003) 'A multivariate statistical approach to reducing the number of variables in data envelopment analysis', *European Journal of Operation Research*, Vol. 147, No. 1, pp.51–61.

Kittelson, S.A.C. (1993) *Stepwise DEA: Choosing Variables for Measuring Technical Efficiency in Norwegian Electricity Distribution*. Memorandum No. 06/93, Department of Economics, University of Oslo, Norway.

Lewin, A.Y., Morey, R.C. and Cook, T.J. (1982) 'Evaluating the administrative efficiency of courts', *OMEGA*, Vol. 10, No. 4, pp.401–411.

Nunamaker, T.R. (1985) 'Using data envelopment analysis to measure the efficiency of non-profit organizations: a critical evaluation', *Managerial and Decision Economics*, Vol. 6, No. 1, pp.50–58.

Parkan, C. (2006) 'An alternative derivation of the basic DEA models for gauging the productive performances of operating entities', *Int. J. Productivity and Quality Management*, Vol. 1, No. 3, pp.253–271.

Pastor, J.T., Ruiz, J.L. and Sirvent, I. (2002) 'A statistical test for nested radial DEA models', *Operations Research*, Vol. 50, No. 4, pp.728–735.

Salinas-Jimenez, J. and Smith, P. (1996) 'Data envelopment analysis applied to quality in primary health care', *Annals of Operations Research*, Vol. 67, No. 1, pp.141–161.

Serrano Cinca, C. and Mar Molinero, C. (2004) 'Selecting DEA specifications and ranking units via PCA', *Journal of the Operational Research Society*, Vol. 55, No. 5, pp.521–528.

Sigala, M., Airey, D., Jones, P. and Lockwood, A. (2004) 'ICT paradox lost? A stepwise DEA methodology to evaluate technology investments in tourism settings', *Journal of Travel Research*, Vol. 43, No. 2, pp.180–192.

Shanmugam, R. and Johnson, C. (2007) 'At a crossroad of data envelopment and principal component analyses', *Omega*, Vol. 35, No. 4, pp.351–364.

Soleimani-Damaneh, M. (2009) 'An effective computational attempt in DDEA', *Applied Mathematical Modelling*, Vol. 33, No. 10, pp.3943–3948.

Toloo, M. and Nalchigarb, S. (2009) 'A new integrated DEA model for finding most BCC-efficient DMU', *Applied Mathematical Modelling*, Vol. 33, No. 1, pp.597–604.

Wagner, J.M. and Shimshaka, D.G. (2007) 'Stepwise selection of variables in data envelopment analysis: procedures and managerial perspectives', *European Journal of Operational Research*, Vol. 180, No. 1, pp.57–67.

Wilson, P.W. and Wheeloc, D.C. (2000) 'Why do banks disappear? The determinants of US bank failures and acquisitions', *Review of Economics and Statistics*, Vol. 82, No. 1, pp.127–338.