Computing the efficiency interval of decision making units (DMUs) having interval inputs and outputs with the presence of negative data

M. Rostamy-Malkhalifeh a*, F. seyed Esmaeili b

(a) Department of Mathematics, Science and Research Branch, Islamic Azad University, Tehran, Iran
(b) Department of Mathematics, South Tehran Branch, Islamic Azad University, Tehran, Iran

Received Autumn 2015, Accepted Winter 2016

Abstract

The basic assumption in data envelopment analysis patterns (DEA) (such as the CCR and BCC models) is that the value of data related to the inputs and outputs is a precise and positive number, but most of the time in real conditions of business, determining precise numerical value is not possible in for some inputs or outputs. For this purpose, different models have been proposed in DEA for imprecise data over recent years and also several researches have been conducted on DEA that are able to evaluate efficiency with negative data. The negative interval DEA pattern which has been introduced and used in the present study, addresses uncertainty both in inputs and outputs and provides user with more stable and reliable results for decision making.

Now, in this paper a model is presented that is able to compute efficiency interval of units with interval input and output that while some indicators can also be negative and then we prove that the efficiency interval that this model gives us is more precise compared to efficiency interval of models previously proposed and finally, ten decision making units (DMUs) with the negative imprecise (interval) data are investigated by the proposed model and the results of the proposed model are compared with the results of the previous models.

Keywords: Data envelopment analysis, interval data, negative data, evaluating the efficiency

* Corresponding author: Email: mohsen_rostamy@yahoo.com
1. Introduction

DEA is one of decision making branches that was introduced by Farrell in 1957 and was extended and developed by the fundamental paper of Charnes, Cooper and Rhodes (1978) [1]. In fact, the DEA is based on a series of optimization using linear programming that is also called nonparametric method and its purpose is to evaluate performance of units under study in which multiple inputs and inputs have been considered as non-negative [2]. But, in some problems we encounter units having several inputs and outputs which are negative. Thus, we need methods that can obtain the efficiency performance of units with negative data. Accordingly, Sharp et al 2006 [4] presented a model entitled as MSBM that is able to evaluate efficiency with negative data.

In classic DEA models, definitive and precise data are used to measure the efficiency of units, but in fact practically, there are situations where there is no precise information of the units' inputs and inputs. Determining the precise numerical value is not possible. In such conditions models should be used that evaluate efficiency of DMUs with regard to the imprecise data. Thus, another subject was proposed in DEA that is related to imprecise (interval) data and models were presented for using these data. Among these models is the Interval data envelopment analysis (IDEA) that was presented by Wang et al (2005) [6] and this is a new and appropriate technique for computing the efficiency in the uncertainty conditions.

In the present paper a model with interval and negative data is presented that is based on MSBM model and is able to provide a stronger and more precise efficiency interval than the obtained efficiency interval of previous paper.

The structure of this paper is as follows: MSBM model is presented in the second section. Then, imprecise (interval) DEA is reviewed in the third section. The proposed model for calculating efficiency of units with negative interval data is presented in the fourth section and the discussion is ended in fifth section by presenting conclusion.

2. MSBM model

Sharp et al (2006) [4] have introduced a modified slack based measure model known as MSBM for negative data that has the ability to handle negative inputs and inputs. Sharp et al (2006) rewrote SBM model for calculating the efficiency measure using SBM model in the presence of negative variables as well as applying
Portela method [5] and placing the improving directions \((R_{io}, R_{ro})\) and called it MSBM:

\[
\begin{align*}
\text{Min } e_o &= \frac{1 - \sum_{i=1}^{m} \frac{w_i s_i^-}{R_{io}}}{1 + \sum_{r=1}^{s} \frac{v_r s_r^+}{R_{ro}}} \\
\text{S.t } &\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io}, \quad i = 1, ..., m \\
&\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro}, \quad r = 1, ..., s \\
&\sum_{j=1}^{n} \lambda_j = 1 \\
&\lambda_j \geq 0, \quad s_i^- \geq 0, \quad s_r^+ \geq 0, \\
&j = 1, ..., n, \quad r = 1, ..., s, \quad i = 1, ..., m
\end{align*}
\]

(1)

Where:
- \(s_i^-\): the value of \(i^{th}\) input slack
- \(s_r^+\): the value of \(r^{th}\) output slack
- \(w_i\) and \(v_r\): the weights predetermined by decision maker (DM).

In addition, vectors in the model are as below:

\(R_{ro} = \max_{j} \{y_{rj}\} - y_{ro}, \quad R_{io} = x_{io} - \min_{j} \{x_{ij}\}\)

When \(R_{io}\) and \(R_{ro}\) are equal to zero, it is assumed that \(\frac{w_i s_i^-}{R_{io}}\) and \(\frac{v_r s_r^+}{R_{ro}}\) terms are eliminated from nominator and denominator.

3. Imprecise (interval) Data Envelopment Analysis

In the classic DEA models it has been assumed that inputs and outputs are assessed by the precise data. But, there are some cases that the precise and reliable values cannot be specified for inputs and outputs due to lack of certainty. Wang et al (2005) [6] proposed IDEA pattern for solving this problem in which interval data efficiency can be achieved by making changes in the DEA models; that each input and output is placed in the determined upper and lower bound ranges in intervals.

As it is observed in Table 1, the values of each inputs and outputs are within an interval that the value of these considered inputs and inputs can be variable in this interval. If every \(n\) units existing use different \(m\) units for producing different \(s\) outputs, then \(DMU_j, \ j = 1, ..., n\) applies the values \(X_j = [x_{1j}, x_{2j}, ..., x_{mj}]^t, i = 1, ..., m\) from inputs for producing \(Y_j = [y_{1r}, y_{2r}, ..., y_{sr}]^t, r = 1, ..., s\); that imprecise inputs and outputs are as follows:

\(y_{rj} \in [y_{lrj}^l, y_{lrj}^u], \quad x_{ij} \in [x_{ij}^l, x_{ij}^u]\)

\(x_{ij}^l\) and \(y_{rj}^l\) are lower bounds, and \(x_{ij}^u\) and \(y_{rj}^u\) are upper bounds for inputs and outputs.
The following model must be solved to calculate efficiency interval of the unit under evaluation in uncertainty conditions and presence of imprecise data:

\[
\bar{\bar{e}}_o = \text{Min} \frac{1 - \sum_{i=1}^{m} w_i^{s_i^-}}{1 + \sum_{r=1}^{s} v_r^{e_r^+}}
\]

S.t

\[
\begin{align*}
\sum_{j=1}^{n} \lambda_j x_{ij}^l + s_j^- &= x_{i0}^l, & i = 1, ..., m \\
\sum_{j=1}^{n} \lambda_j y_{ij}^- - s_j^+ &= y_{r0}^- & r = 1, ..., s \\
\sum_{j=1}^{n} \lambda_j &= 1 \\
\lambda_j &\geq 0, & s_i^- \geq 0, & s_i^+ \geq 0, \\
&j = 1, ..., n, & r = 1, ..., s, & i = 1, ..., m
\end{align*}
\]  

(2)

Esmaeili and Rostamy [3] presented a method that calculated efficiency interval for each the DMUs in such conditions. They presented two models (3) and (4) that put the evaluated DMU in the worst conditions and PPS border in its own best conditions to compute lower bound of the efficiency (the evaluated DMU is involved in making border in the worst its own conditions) and put the evaluated DMU in the best conditions and PPS border in its own worst conditions to compute upper bound of the efficiency (the evaluated DMU is involved in making border in the best its own conditions).

For now, we present a model in the next section that is able to compute a more precise interval compared to the efficiency interval presented in the paper [3].

\[
p_o^L = \text{min} \frac{1 - \sum_{i=1}^{m} w_i^{s_i^-}}{1 + \sum_{r=1}^{s} v_r^{e_r^+}}
\]

S.t

\[
\begin{align*}
\sum_{j=1}^{n} \lambda_j x_{ij}^l + s_j^- &= x_{i0}^l, & i = 1, ..., m \\
\sum_{j=1}^{n} \lambda_j y_{ij}^u - s_j^+ &= y_{r0}^u & r = 1, ..., s \\
\sum_{j=1}^{n} \lambda_j &= 1 \\
\lambda_j &\geq 0, & s_i^- \geq 0, & s_i^+ \geq 0, \\
&j = 1, ..., n, & r = 1, ..., s, & i = 1, ..., m
\end{align*}
\]  

(3)

\[
p_o^u = \text{min} \frac{1 - \sum_{i=1}^{m} w_i^{s_i^-}}{1 + \sum_{r=1}^{s} v_r^{e_r^+}}
\]

S.t

\[
\begin{align*}
\sum_{j=1}^{n} \lambda_j x_{ij}^u + s_j^- &= x_{i0}^u, & i = 1, ..., m \\
\sum_{j=1}^{n} \lambda_j y_{ij}^l - s_j^+ &= y_{r0}^l & r = 1, ..., s
\end{align*}
\]
\[
\sum_{i=1}^{n} \lambda_i = 1 \\
\lambda_i \geq 0, \quad s_i^- \geq 0, \quad s_r^+ \geq 0, \\
j = 1, \ldots, n, \quad r = 1, \ldots, s, \quad i = 1, \ldots, m
\]

(4)

4. The proposed model to calculate efficiency of units with negative interval data

We present two models for computing interval efficiency of model (2) in this section and using them compute the lower and upper bound of interval efficiency. We put the evaluated DMU in the worst conditions and other DMUs in its own best conditions to compute lower bound of the efficiency interval (the evaluated DMU is involved in making border in the worst its own conditions) and put the evaluated DMU in the best conditions and other DMUs in its own worst conditions to compute upper bound of the efficiency interval (the evaluated DMU is involved in making border in the best its own conditions). Therefore, we introduce DMU with negative imprecise data to identify lower and upper bound of the interval efficiency respectively in models (5) and (6). Model (5) shows a lower bound of unit efficiency \( J_0 \) interval:

\[
e_0^L = \text{Min} \left( \frac{1 - \sum_{i=1}^{m} \frac{w_i s_i^-}{R_{io}}}{1 + \sum_{r=1}^{s} \frac{v_r s_r^+}{R_{ro}}} \right)
\]

S.t \[
\sum_{i=1}^{n} \lambda_i x_i^0 + s_i^- + \lambda_o x_i^0 = x_i^0, \quad i = 1, \ldots, m
\]

\[
\sum_{j=1}^{n} \lambda_j Y_{ij}^u - s_i^+ + \lambda_o y_{ij}^u = y_{ij}^u, \quad r = 1, \ldots, s
\]

\[
\sum_{i=1}^{n} \lambda_i = 1, \quad \lambda_i \geq 0, s_i^- \geq 0, s_r^+ \geq 0,
\]

(5)

\[
\sum_{i=1}^{n} \lambda_i = 1, \quad \lambda_i \geq 0, s_i^- \geq 0, s_r^+ \geq 0,
\]

(6)

\[
\sum_{i=1}^{n} \lambda_i = 1, \quad \lambda_i \geq 0, s_i^- \geq 0, s_r^+ \geq 0,
\]

j = 1, \ldots, n, \quad r = 1, \ldots, s, \quad i = 1, \ldots, m

Model (6) shows an upper bound of unit efficiency \( J_0 \) interval:

\[
R_{ro} = y_{ro}^{Max} - y_{ro}^{Min}
\]

\[
R_{io} = x_{io}^{Max} - x_{io}^{Min}
\]

Now, it is illustrated in the following theorem that \( \bar{e}_o \in [e_o^L, e_o^u] \).

**Theorem 1:**

If \( \bar{e}_o, e_o^L \) and \( e_o^u \) are the optimums of target functions of models (2), (5), (6) respectively, then: \( e_o^L \leq \bar{e}_o \leq e_o^u \)

Proof: assume that \( \bar{\lambda} \) and \( \bar{s} \) is the optimum of model (2).
\[ \sum_{j=1}^{n} \tilde{\lambda}_j x_{ij} + \tilde{\lambda}_a x_{10} \leq \tilde{\chi}_{io} - \tilde{s}^- + \tilde{\lambda}_a (x_{10}^u - \tilde{\chi}_{io}) \]

Since, \( \sum_{j=1}^{n} \tilde{\lambda}_j \tilde{\chi}_{ij} + s^-_i = \tilde{\chi}_{io} \), therefore we have:

\[ \sum_{j=1}^{n} \tilde{\lambda}_j x_{ij} + \tilde{\lambda}_a x_{10}^u \leq \tilde{\chi}_{io} - s^- + \tilde{\lambda}_a (x_{10}^u - \tilde{\chi}_{io}) \]

\[ = \sum_{j=1}^{n} \tilde{\lambda}_j x_{ij} + \tilde{\lambda}_a x_{10}^u + s^-_i \leq x_{10}^u \]

There is also \( s^+ = \tilde{s}^- + t^+ \), thus:

\[ \sum_{j=1}^{n} \tilde{\lambda}_j y_{rj} + \tilde{\lambda}_a y_{10} - \tilde{s}^- = y_{10} \]

Then, \( s^-, \tilde{s}^+, \tilde{\lambda} \) is feasible solution for model (5) whose objective function value applies in the following conditions:

Since, \( \tilde{s}^- \geq s^- \) so \( \sum_{i=1}^{m} w_i \tilde{s}_i \geq \sum_{i=1}^{m} w_i s_i \)

\[ 1 - \sum_{i=1}^{m} w_i \tilde{s}_i \leq 1 - \sum_{i=1}^{m} w_i s_i \]

\[ 1 - \sum_{r=1}^{s} v_r \tilde{s}_r \leq 1 - \sum_{r=1}^{s} v_r s_r \]

And \( s^+ = \tilde{s}^- + t^+ \), thus:

\[ \sum_{j=1}^{n} \tilde{\lambda}_j y_{rj} + \tilde{\lambda}_a y_{10} - \tilde{s}^- = y_{10} \]

Therefore, the optimum of target function of model (5) equals to \( e_1 \), and is smaller or equal to the value of target function for the feasible solution of \( \tilde{s} \) and \( \tilde{\lambda} \).

In other words, \( e_1 \leq \tilde{e} \).

Similarly, it is proven that \( e^u \geq \tilde{e} \).

Now, after solving models (5) and (6) interval efficiency of \([e_1^l, e_1^u]\) is obtained for \( j \) unit.

Now, according to the obtained efficiency intervals, units can be divided into the following three categories:
E^{++} = \{ j \in J \mid e^l_j = 1 \}
E^+ = \{ j \in J \mid e^l_j < 1, e^U_j = 1 \}
E^- = \{ j \in J \mid e^U_j < 1 \}

In the above sets, if \( e^l_j = 1 \), then the \( j \)th decision-making unit is efficient for all values of input/output intervals.
However, if \( e^l_j < 1 \) and \( e^U_j = 1 \), the \( j \)th decision-making unit is only efficient for the upper bounds of input/output intervals.
If \( e^U_j < 1 \), the \( j \)th decision-making unit is not efficient for any values in the input/output intervals.

In the following theorem we prove that the efficiency interval obtained is more precise than the paper [3].

**Theorem 2:**
The efficiency interval obtained from models (5) and (6), is more precise than the efficiency interval calculated from models (3) and (4), i.e.: \( e^U_0 \leq p^u_0 \) and \( p^l_0 \leq e^l_0 \).

**Proof:** Suppose \( \hat{\lambda} \) and \( \hat{s} \) are optimum solution of model (5):

\[
\sum_{j=1}^{n} \hat{\lambda}_j x^l_{ij} \leq \sum_{j=1}^{n} \hat{\lambda}_j x^U_{ij} + \hat{\lambda}_0 x^u_{io} = x^u_{io} - \hat{s}_i^\gamma
\]

Therefore, \( \hat{s}_i^\gamma \geq \) exists that:

\[
\sum_{j=1}^{n} \hat{\lambda}_j x^l_{ij} + \hat{s}_i^\gamma = x^u_{io} - \hat{s}_i^\gamma
\]

As a result, we have:

\[
\sum_{j=1}^{n} \hat{\lambda}_j x^l_{ij} + (\hat{s}_i^\gamma + \hat{s}_i^\gamma) = x^u_{io}
\]

Now, for output variables we have:

\[
\sum_{j=1}^{n} \hat{\lambda}_j y^u_{ij} \geq \sum_{j=1}^{n} \hat{\lambda}_j y^u_{ij} + \hat{\lambda}_0 y^l_{io} = y^l_{io} + \hat{s}_r^\gamma
\]

Therefore, \( \hat{s}_r^\gamma \geq \) exists that:

\[
\sum_{j=1}^{n} \hat{\lambda}_j y^u_{ij} - \hat{s}_r^\gamma = y^l_{io} + \hat{s}_r^\gamma
\]

So, \( \sum_{j=1}^{n} \hat{\lambda}_j y^u_{ij} - (\hat{s}_r^\gamma + \hat{s}_r^\gamma) = y^l_{io} \)

Then, \( \hat{\lambda}, (\hat{s}_r^\gamma + \hat{s}_r^\gamma), (\hat{s}_r^\gamma + \hat{s}_r^\gamma) \) is a feasible solution for model (3) whose objective function value is equal to: \( (\hat{s}_r^\gamma + \hat{s}_r^\gamma) = k \) then

\[
k = \frac{1 - \sum_{i=1}^{m} w_i (\hat{s}_r^\gamma + \hat{s}_r^\gamma)}{R_{io} + \sum_{r=1}^{s} v_r (\hat{s}_r^\gamma + \hat{s}_r^\gamma) \frac{R_{ro}}{R_{io}}}
\]

That \( k \leq e^l_0 \) and since model (5) is minimization, for the optimum value of its objective function we have:

\( p^l_0 \leq k \leq e^l_0 \) then \( p^l_0 \leq e^l_0 \)

And similarly it is proved that \( e^U_0 \leq p^u_0 \) and the proof is complete.

Now, we obtain efficiency measure of 10 DMUs by the model proposed in this paper through software and compare the obtained results with the obtained efficiency measure from models (3) and (4).

**5. A numerical example**

Assume that there are ten DMUs with one input and two outputs intervals according to the table 2.
We solve the proposed model by assigning weight of 0.50 for each output slack variables in the model and weight of 1 for each input slack variable in the objective function.

The obtained efficiencies have been brought in Table 3 for each of DMUs.

Table 2: Ten DMU with one input and two outputs

<table>
<thead>
<tr>
<th>DMUj</th>
<th>$x^I_{1j}$</th>
<th>$x^U_{1j}$</th>
<th>$y^I_{1j}$</th>
<th>$y^U_{1j}$</th>
<th>$y^I_{2j}$</th>
<th>$y^U_{2j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11/5</td>
<td>12/5</td>
<td>14/5</td>
<td>15/25</td>
<td>10/75</td>
<td>11/25</td>
</tr>
<tr>
<td>2</td>
<td>34/75</td>
<td>35/25</td>
<td>17/99</td>
<td>18/23</td>
<td>5/80</td>
<td>6/12</td>
</tr>
<tr>
<td>3</td>
<td>24/5</td>
<td>25/5</td>
<td>19/75</td>
<td>20/25</td>
<td>12/40</td>
<td>13/10</td>
</tr>
<tr>
<td>4</td>
<td>21/75</td>
<td>22/25</td>
<td>11/97</td>
<td>12/12</td>
<td>20/10</td>
<td>19/95</td>
</tr>
<tr>
<td>5</td>
<td>39/25</td>
<td>40/25</td>
<td>-10/21</td>
<td>-9/80</td>
<td>24/50</td>
<td>25/02</td>
</tr>
<tr>
<td>6</td>
<td>49/5</td>
<td>50/5</td>
<td>-9</td>
<td>-7</td>
<td>26/80</td>
<td>27/10</td>
</tr>
<tr>
<td>7</td>
<td>34/5</td>
<td>35/5</td>
<td>-18/25</td>
<td>-17/75</td>
<td>5/50</td>
<td>6/25</td>
</tr>
<tr>
<td>8</td>
<td>39/99</td>
<td>40/21</td>
<td>-10/5</td>
<td>-9/5</td>
<td>21/99</td>
<td>22/06</td>
</tr>
<tr>
<td>9</td>
<td>24/75</td>
<td>25/25</td>
<td>-8</td>
<td>-6</td>
<td>18/75</td>
<td>19/05</td>
</tr>
<tr>
<td>10</td>
<td>15/5</td>
<td>16/5</td>
<td>25/50</td>
<td>26/50</td>
<td>7/75</td>
<td>8/19</td>
</tr>
</tbody>
</table>

Table 3: Efficiency results for models (3), (4), (5),(6)

<table>
<thead>
<tr>
<th>DMUj</th>
<th>$p^j_I$</th>
<th>$p^j_U$</th>
<th>$e^j_I$</th>
<th>$e^j_U$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0/955</td>
<td>1+</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0/375</td>
<td>0/428</td>
<td>0/375</td>
<td>0/428</td>
</tr>
<tr>
<td>3</td>
<td>0/770</td>
<td>1+</td>
<td>0/770</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0/993</td>
<td>1+</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0/879</td>
<td>1+</td>
<td>0/879</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>0/916</td>
<td>1+</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>0.255</td>
<td>0/289</td>
<td>0/255</td>
<td>0.298</td>
</tr>
<tr>
<td>8</td>
<td>0/612</td>
<td>0/646</td>
<td>0/612</td>
<td>0.646</td>
</tr>
<tr>
<td>9</td>
<td>0/754</td>
<td>0/711</td>
<td>0/754</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0/950</td>
<td>1+</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
As it is observed in Table 3, the efficiency intervals obtained from models (5) and (6) are smaller than or equal the efficiency intervals obtained from models (4) and (3), i.e. the efficiency interval that is obtained from models (5) and (6) is more precise and robust than the efficiency interval of models (4) and (3). And also some DMUs in the efficiency interval of $[p_l,p_u]$ are located in the best conditions, outside the pps border and become super-efficient that we have shown their value with $1^+$. But according to the presented models (models (5) and (6), all efficiency intervals of the achieved DMUs are feasible, in addition, as it is clear from the results, the obtained efficiency intervals of models (5) and (6) can be categorized as follows:

DMU_1, DMU_4, DMU_6, DMU_10 are efficient in their own the worst conditions that have lower bound equal to 1, i.e.:

$E^{++} = \{\text{DMU}_1, \text{DMU}_4, \text{DMU}_6, \text{DMU}_{10}\}$

And DMU_5, DMU_3 are also efficient in their own the best conditions that have upper bound of equal to 1, then:

$E^+ = \{\text{DMU}_5, \text{DMU}_3\}$

Also, DMU_9, DMU_8, DMU_7, DMU_2 are inefficient in their own the worst conditions, thus:

$E^- = \{\text{DMU}_9, \text{DMU}_8, \text{DMU}_7, \text{DMU}_2\}$

6. Conclusion:

In the present paper we first brought MSBM model and stated that this model is able to calculate the efficiency with negative data and then given that in some problems data are imprecise and as interval, we introduced two models of evaluating MSBM efficiency with considering these assumptions and proved that the optimal value of lower bound is smaller than that of upper bound or equal to it and also showed that the obtained efficiency interval of these two model is the smaller and more precise than that of the previous interval model.
References


