Data Envelopment Analysis as a Kaizen Tool: SBM Variations Revisited

Kaoru Tone

National Graduate Institute for Policy Studies
7-22-1 Roppongi, Minato-ku, Tokyo 106-8677, Japan
tone@grips.ac.jp

Keywords: DEA, SBM, nearest point, SBM-Max, Kaizen

Abstract. Slacks-based measure (SBM) (Tone (2001), Pastor et al. (1999)) has been widely utilized as a representative non-radial DEA model. In Tone (2010), I developed four variants of the SBM model where main concerns are to search the nearest point on the efficient frontiers of the production possibility set. However, in the worst case, a massive enumeration of facets of polyhedron associated with the production possibility set is required. In this paper, I will present a new scheme, called SBM-Max, for this purpose which requires a limited number of additional linear program solutions for each inefficient DMU. Although the point thus obtained is not always the nearest point, it is acceptable for practical purposes and from the point of computational loads. Inefficient DMUs can be improved to the efficient status with less input-reductions and less output-enlargement. Thus, this model proposes a Kaizen (improvement) tool by DEA.

1. Introduction

There are two types of models in DEA (Data Envelopment Analysis); radial and non-radial. Radial models are represented by the CCR (Charnes-Cooper-Rhodes) model. Basically they deal with proportional changes of inputs or outputs. As such, the CCR score reflects the proportional maximum input (output) reduction (expansion) rate which is common to all inputs (outputs). However, in real world businesses, not all inputs (outputs) behave in the proportional way. For example, if we employ labor, materials and capital as inputs, some of them are substitutional and do not change proportionally. Another shortcoming of the radial models is the neglect of slacks in reporting the efficiency score. In many cases, we find a lot of remaining non-radial slacks. So, if these slacks have an important role in evaluating managerial efficiency, the radial approaches may mislead the decision when we utilize the efficiency score as the only index for evaluating performance of DMUs.

In contrast, the non-radial SBM models put aside the assumption of proportionate changes in inputs and outputs, and deal with slacks directly. This may discard varying proportions of original inputs and outputs. The SBM models are designed to meet the following two conditions.

   (1) Units invariant: The measure should be invariant with respect to the units of data

   (2) Monotone: The measure should be monotone decreasing in each slack in input and output.

The original SBM model evaluates efficiency of DMUs referring to the furthest frontier point within a range. This results in the worst score for the objective DMU and the projection may go to a remote point on the efficient frontier which may be inappropriate as the reference. In Tone (2010), I developed four variants of the SBM model where main concerns are to search the nearest point on the efficient frontiers of the production possibility set.
We depict the relationship among the ordinary SBM (SBM-Min), CCR and SBM-Max models by Fig. 1. Inefficient DMU P’s projections are Q, R and S respectively by SBM-Min, CCR and SBM-Max. Mathematically, finding S belongs to a NP-hard problem, because it is a maximization problem of a convex function over a convex region. However, the projected point S indicates that we can attain an efficient status with less input reductions and less output expansions than the ordinary SBM (Min) models. We can say that the projection by the SBM-Max model represents a practical “Kaizen” (improvement) by DEA.

Referring these variations, several authors published new models. Among them, I introduce two important papers.

Fukuyama et al. (2014) developed a least distance efficiency measure with the strong/weak monotonicity of the ratio form measure under several norms including 1-norm, 2-norm and $\infty$-norm. This model utilizes mixed-integer linear programming (MILP) to identify efficiency frontiers and hence a computational difficulty arises for large-scale problems.

Hadi-Vencheh et al. (2015) developed a new SBM model to find the nearest point on the efficient frontiers. They utilize the multiplier form model to find all supporting hyperplanes. It also utilizes software which uses fractional coefficients (high precision arithmetic) to avoid loss data. Hence, computational time increases for large-scale problems.

In order to apply DEA models to actual real world problems, we need to try many instances including selection of DMUs and input/output factors before attaining the final scheme of evaluation. For this purpose, allowable computation time and easy accessible software are desirable.

The motivation and purpose of this paper is to obtain nearly closest points on the efficient frontiers within foreseeable computation loads using only popular linear programming codes.

The rest of this paper is organized as follows. Section 2 introduces the ordinary SBM-Min model briefly. Section 3 presents the new SBM-Max model. Observations on this new model are described in Section 4. Two numerical examples are exhibited in Section 5. Section 6 concludes this paper. Although we present the model in non-oriented mode, we can treat input- and output-oriented model as well. As to returns-to-scale characteristics, we present the constant returns-to-scale (CRS) case. However we can deal with the variable returns-to-scale (VRS) model as well.

2. The SBM Min model

The SBM model was introduced by Tone (2001) (see also Pastor et al. (1999)). It has three variations, i.e. input-, output- and non-oriented. The non-oriented model indicates both input- and output-oriented.

Let the set of DMUs be $J = \{1, 2, \ldots, n\}$, each DMU having $m$ inputs and $s$ outputs. We denote the vectors of inputs and outputs for DMU $j$ by $x_j = (x_{j1}, x_{j2}, \ldots, x_{jm})^T$ and $y_j = (y_{j1}, y_{j2}, \ldots, y_{js})^T$, respectively. We define input and output matrices $X$ and $Y$ by

$$X = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^{m \times n} \text{ and } Y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^{s \times n}. \quad (1)$$

We assume that all data are positive i.e. $X > 0$ and $Y > 0$.  

2.1 Production Possibility Set

The production possibility set is defined using the non-negative combination of the DMUs in the set $J$ as:

$$ P = \left\{ (x, y) \mid x \geq \sum_{j=1}^{n} \lambda_j x_j, \quad 0 \leq y \leq \sum_{j=1}^{n} \lambda_j y_j, \quad \lambda \geq 0 \right\}. $$

$\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)^T$ is called the intensity vector.

The inequalities in (2) can be transformed into equalities by introducing slacks as follows:

$$ x = \sum_{j=1}^{n} \lambda_j x_j + s^- $$
$$ y = \sum_{j=1}^{n} \lambda_j y_j - s^+ $$
$$ s^- \geq 0, \quad s^+ \geq 0, $$

where $s^- = (s_1^-, s_2^-, \ldots, s_m^-)^T \in \mathbb{R}^m$ and $s^+ = (s_1^+, s_2^+, \ldots, s_s^+)^T \in \mathbb{R}^s$ are respectively called input and output slacks.

2.2 Non-oriented SBM

Non-oriented or both-oriented SBM efficiency $\rho_o^{\text{min}}$ is defined by

$$ \text{[SBM-Min]} \quad \rho_o^{\text{min}} = \min_{\lambda, s^-, s^+} \frac{1 - \frac{1}{m} \sum_{i=1}^{m} S_i^-}{1 + \frac{1}{s} \sum_{r=1}^{s} S_r^+} $$
subject to

$$ x_{io} = \sum_{j=1}^{n} x_j \lambda_j + s_i^- \quad (i = 1, \ldots, m) $$
$$ y_{ro} = \sum_{j=1}^{n} y_j \lambda_j - s_r^+ \quad (r = 1, \ldots, s) $$
$$ \lambda_j \geq 0 \quad (\forall j), \quad s_i^- \geq 0 \quad (\forall i), \quad s_r^+ \geq 0 \quad (\forall r). $$

[Definition 1] (SBM-efficient) A DMU $o = (x_o, y_o)$ is called SBM-efficient if $\rho_o^{\text{min}} = 1$ holds. This means $s^- = 0$ and $s^+ = 0$, i.e. all input and output slacks are zero.

[SBM-Min] can be transformed into a linear program using the Charnes-Cooper transformation as follows:

$$ \text{[SBM-Min-LP]} \quad \tau^* = \min_{t, \Lambda, S^-, S^+} \frac{1}{m} \sum_{i=1}^{m} S_i^- $$
subject to

$$ 1 = t + \frac{1}{s} \sum_{r=1}^{s} S_r^+ $$
$$ tx_{io} = \sum_{j=1}^{n} x_j \Lambda_j + S_i^- \quad (i = 1, \ldots, m) $$
$$ ty_{ro} = \sum_{j=1}^{n} y_j \Lambda_j - S_r^+ \quad (r = 1, \ldots, s) $$
$$ \Lambda_j \geq 0 \quad (\forall j), \quad S_i^- \geq 0 \quad (\forall i), \quad S_r^+ \geq 0 \quad (\forall r), \quad t > 0. $$

Let an optimal solution be $(\tau^*, t^*, \Lambda^*, S^-, S^+)$. Then, we have an optimal solution of [SBM-Min] as defined by

$$ \rho_o^{\text{min}} = \tau^*, \quad \lambda^* = \Lambda^*/t^*, \quad s^+ = S^+ / t^*, \quad s^- = S^- / t^*. $$
3. The SBM Max Model

In this section, we introduce the new non-oriented SBM-Max model.

**Step 1. Solve SBM-Min**

First, we solve the ordinary SBM (SBM-Min) model as represented by the program (4) for DMU \((x_o, y_o)\) \((o = 1, \ldots, n)\). Let an optimal solution be \((\lambda^*, s^-, s^+)\).

**Step 2. Define efficient DMUs**

We define the set \(R^{\text{eff}}\) of all efficient DMUs as

\[
R^{\text{eff}} = \{j | \rho^\text{min}_j = 1, j = 1, \ldots, n\}. \tag{7}
\]

We denote these efficient DMUs as \((x^o_{eff}, y^o_{eff}), (x^o_{eff}, y^o_{eff}), \ldots, (x^o_{Neff}, y^o_{Neff})\) , where \(N^{\text{eff}}\) is the number of efficient DMUs.

**Step 3. Local reference set**

For an inefficient DMU \((x_o, y_o)\), we define the local reference set \(R^{\text{local}}_o\), i.e., efficient DMUs set for DMU \((x_o, y_o)\), by (8).

\[
R^{\text{local}}_o = \{j | \lambda^*_j > 0, j = 1, \ldots, n\}. \tag{8}
\]

**Step 4. Pseudo-Max score**

For each inefficient DMU, i.e., \(\rho^\text{min}_o < 1\), we solve the following program.

\[
\text{max} \quad 1 - \frac{1}{m} \sum_{i=1}^{m} s^o_{i-} \\
\text{subject to} \\
x_o = \sum_{j \in R^{\text{local}}_o} x_j \lambda^*_j + s^+ \\
y_o = \sum_{j \in R^{\text{local}}_o} y_j \lambda^*_j - s^-
\]

Let an optimal slacks be \((s^-, s^+)\). We solve the following program with variables \((\lambda, s^-, s^+)\).

\[
\text{min} \quad 1 - \frac{1}{m} \sum_{i=1}^{m} s^o_{i-} \\
\text{subject to} \\
x_o - s^+ = \sum_{j \in R^{\text{local}}_o} x^o_{eff} \lambda^*_j + s^+ \\
y_o + s^- = \sum_{j \in R^{\text{local}}_o} y^o_{eff} \lambda^*_j - s^-
\]

Let the optimal slacks be \((s^{**}, s^{**})\). We define the Pseudo-Max score \(\rho^\text{pseudomax}_o\) by

\[
\rho^\text{pseudomax}_o = \frac{1 - \frac{1}{m} \sum_{i=1}^{m} s^o_{i-} + s^o_{i-}^{**}}{1 + \frac{1}{s} \sum_{r=1}^{s} s^o_{r-} + s^o_{r-}^{**}}. \tag{11}
\]
Step 5. Distance and SBM-Max score
For each inefficient DMU \((x_o, y_o)\), i.e., \(\rho_o^{\text{min}} < 1\), we calculate the distance between \((x_o, y_o)\) and \((x_{o_h}^{\text{eff}}, y_{o_h}^{\text{eff}})\) \((h = 1, \ldots, \text{Neff})\) by

\[
[d_h] = \sum_{i=1}^{m} \frac{|x_{ih}^{\text{eff}} - x_{io}|}{x_{io}} + \sum_{i=1}^{s} \frac{|y_{ih}^{\text{eff}} - y_{io}|}{y_{io}}.
\] (12)

This distance is units-invariant.

Step 5.1. Reorder the distance
We renumber the efficient DMUs in the ascending order of \(d_h\), so that

\[d_1 \leq d_2 \leq \ldots \leq d_{\text{Neff}}.\] (13)

We define the set \(R_h\) by

\[R_h = \{1, \ldots, h\} \quad (h = 1, \ldots, \text{Neff}).\] (14)

Step 5.2. Find slacks and max-score for the set \(R_h\)
We evaluate the efficiency score of the inefficient DMU \((x_o, y_o)\) referring to the set \(R_h\) by solving the following program.

\[
[\text{Max-1}] \quad \max_{s, \lambda} \frac{1 - \frac{1}{m} \sum_{i=1}^{m} S_{io}}{1 + \frac{1}{s} \sum_{r=1}^{s} S_{ro}}
\]

subject to

\[
x_o = \sum_{j \in R_h} x_{j}^{\text{eff}} \lambda_j + s^-
\]

\[
y_o = \sum_{j \in R_h} y_{j}^{\text{eff}} \lambda_j - s^+
\]

\[s^-, s^+, \lambda \geq 0.\] (15)

(a) If this program is infeasible, we define \(\rho_{oh}^* = 0\). Otherwise, let an optimal slacks be \((s^-, s^+)\).

(b) If the optimal objective value is 1, i.e., \(s^- = 0\) and \(s^+ = 0\), we define \(\rho_{oh}^* = 0\). This indicates that DMU \((x_o, y_o)\) can be expressed as a non-negative combination of DMUs in \(R_h\) and hence, in view of \(\rho_o^{\text{min}} < 1\), it is inside the production possibility set.

(c) If the optimal objective value is less than 1, we again solve the following program with the variables \((\lambda, s^-, s^+)\).

\[
[\text{Max-2}] \quad \min_{s, \lambda} \frac{1 - \frac{1}{m} \sum_{i=1}^{m} S_{io}}{1 + \frac{1}{s} \sum_{r=1}^{s} S_{ro}}
\]

subject to

\[
x_o - s^- = \sum_{j \in R_h} x_{j}^{\text{eff}} \lambda_j + s^-
\]

\[
y_o + s^+ = \sum_{j \in R_h} y_{j}^{\text{eff}} \lambda_j - s^+
\]

\[s^-, s^+, \lambda \geq 0.\] (16)

Let the optimal slacks be \((s^-, s^+)\). We define \(\rho_{oh}^*\) by
\[
\rho^*_o \, \rho^*_o = \frac{1 - \frac{1}{m} \sum_{i=1}^{m} S_{io}^* + S_{io}^{**}}{1 + \frac{1}{s} \sum_{j=1}^{s} S_{jo}^* + S_{jo}^{**}}.
\]  
(17)

We assign \( \rho^*_o \) as the max-score referring to the set \( R_h \).

**Step 5.3. SBM-Max and projection**

Finally, we define the max-score \( \rho^\max_o \) of inefficient DMU \((x_o, y_o)\) by

\[
[SBM-Max] \quad \rho^\max_o = \max \left\{ \rho^\text{pseudo max}_o, \rho^*_o, \ldots, \rho^*_\text{Neff} \right\}.
\]  
(18)

We also hold the slacks \((s^{**}, s^{**})\) corresponding to the maximum \( \rho^\max_o \). The projection of DMU \((x_o, y_o)\) onto efficient frontiers is given by

\[
[\text{Projection}] \quad x^*_o = x_o - s^* - s^{**}, \quad y^*_o = y_o + s^* + s^{**}.
\]  
(19)

The projected point \((x^*_o, y^*_o)\) is efficient with respect to the efficient DMU set \( R^{\text{eff}} \). However, it does not always satisfy Pareto-Koopmans efficiency condition.

### 4. Observations

In this section, we discuss several characteristics of the algorithm in Section 3.

**4.1. Distance and choice of the set \( R_h \)**

The set \( R_h \) plays a central role in choosing referent DMUs for inefficient DMUs. Because our main concern is the projection to the nearest point on the efficient frontiers, we evaluate the distance between the DMU \((x_o, y_o)\) and efficient DMUs by (12), and choose the shortest distance DMU as the first candidate DMU. Then, we expand the referent set in the ascending order of distances. Thus, we can expect a close efficient point on the frontiers with high probability. If tie occurs in distances, we can choose any one at random.

**4.2. The role of Programs (10) and (16)**

For example, Program (16) is necessary to project the point \((x_o - s^* - s^{**}, y_o') = y_o + s^* + s^{**}\) on the efficient frontiers. Thus, \((x^*_o = x_o - s^* - s^{**}, y^*_o = y_o + s^* + s^{**})\) is the projected point on the efficient frontiers and it is expected to be close to the DMU \((x_o, y_o)\) by the selection rule of \( R_h \).

**4.3. Computational amount**

Computations needed for this algorithm for an inefficient DMU are as follows.

Let \( t_1 \) and \( t_2 \) be the CPU time for solving a LP problem, respectively, with the \((m+s)\) rows and \( n \) columns, and \((m+s)\) rows and \( \text{Neff} \) columns. Since LP solution time is proportional to the number of columns. We can estimate roughly \( t_1 = \frac{n}{\text{Neff}} \) \( t_2 \).

(1) Program (4) or (5) needs \( n \times t_1 \) CPU time.

(2) Programs (9) and (10) need at most \( 2 \times (n - \text{Neff}) \times t_2 \) CPU time.

(3) Programs (15) and (16) need at most \( 1.5 \times (n - \text{Neff}) \times \text{Neff} \times t_2 \) CPU time, because the member of \( R_h \) in (15) varies from 1 to \( \text{Neff} \).

However, if Step 5.2 (b) occurs at some set \( R_h \), we can skip the computations for the succeeding Programs (15) and (16) for \( h +1, \ldots, \text{Neff} \).

Overall, the total time for LP computation is at most

\[
T = n \times t_1 + 2 \times (n - \text{Neff}) \times t_2 + 1.5 \times (n - \text{Neff}) \times \text{Neff} \times t_2
= \left[ n + (2 + 1.5 \times \text{Neff}) \times (n - \text{Neff}) \times (\text{Neff} / n) \right] \times t_1.
\]  
(20)

Thus, the computational amount is polynomial order and we do not need other software, e.g., MILP and fractional arithmetic.
4.4. Consistency with the super-efficiency SBM measure
The SBM-Max model aims at getting to the nearest point on the efficient frontiers. This concept is in line with the super-efficiency SBM model (Tone (2002)) which solves the following program for an efficient DMU \((x_o, y_o)\) to measure the minimum ratio-scale distance from the efficient frontier excluding the DMU \((x_o, y_o)\).

\[
\delta^* = \min_{\lambda, s^-} \frac{1 + \frac{1}{m} \sum_{i=1}^{m} s_{i}^{-}}{1 - \frac{1}{s} \sum_{r=1}^{r} s_{r}^{+}}
\]

subject to
\[
\begin{align*}
x_o &= \sum_{j=1,j \neq o}^{n} x_j \lambda_j - s^- \\
y_o &= \sum_{j=1,j \neq o}^{n} y_j \lambda_j + s^+ \\
\lambda &\geq 0, s^- \geq 0, s^+ \geq 0.
\end{align*}
\]

We can solve the super-efficiency SBM model by applying LP code just once, because this problem belongs to a convex programming, i.e., minimization of a convex function over a convex region. However, SBM-Max problem cannot be solved in this manner, because it is a maximization of a convex function over a convex region.

4.5 Addition of weights to input- and output-slacks
We can assign weights \((w^- \text{ and } w^+)\) to input and output slacks in the objective function of the above SBM models corresponding to the relative importance of items as follows:

\[
[\text{Weighted-SBM}] \quad \rho^* = \min_{\lambda, s^-} \frac{1 - \frac{1}{m} \sum_{i=1}^{m} W_i^- s_{i}^-}{1 + \frac{1}{s} \sum_{r=1}^{r} W_r^+ s_{r}^+},
\]

with \(\sum_{i=1}^{m} W_i^- = m\) and \(\sum_{i=1}^{m} W_r^+ = s\). The weights should reflect the intentions of decision-makers.

We can define the input- (output-) oriented Weighted-SBM models by neglecting the denominator (numerator) of the objective function in (22).

5. Numerical examples
In this section, we show two numerical examples, the first one is illustrative and the other deals with real data. All computations are executed using a PC with Intel Core i7-3770 CPU at 3.40 GHz 16 GB (RAM) and Microsoft Excel VBA (Visual Basic for Applications). A LP soft (revised simplex method) is coded by the author. We checked the results of the first example using LINGO (LINDO Systems Inc.) and had the same figures.

5.1. An illustrative example
We deal with the same data as one in Tone (2010). Table 1 displays the data with two inputs (Doctor and Nurse) and two outputs (Outpatient and Inpatient).
Table 1: Illustrative example

<table>
<thead>
<tr>
<th>DMU</th>
<th>Doctor</th>
<th>Nurse</th>
<th>Outpatient</th>
<th>Inpatient</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>20</td>
<td>151</td>
<td>100</td>
<td>90</td>
</tr>
<tr>
<td>B</td>
<td>19</td>
<td>131</td>
<td>150</td>
<td>50</td>
</tr>
<tr>
<td>C</td>
<td>25</td>
<td>160</td>
<td>160</td>
<td>55</td>
</tr>
<tr>
<td>D</td>
<td>27</td>
<td>168</td>
<td>180</td>
<td>72</td>
</tr>
<tr>
<td>E</td>
<td>22</td>
<td>158</td>
<td>94</td>
<td>66</td>
</tr>
<tr>
<td>F</td>
<td>55</td>
<td>255</td>
<td>230</td>
<td>90</td>
</tr>
<tr>
<td>G</td>
<td>33</td>
<td>235</td>
<td>220</td>
<td>88</td>
</tr>
<tr>
<td>H</td>
<td>31</td>
<td>206</td>
<td>152</td>
<td>80</td>
</tr>
<tr>
<td>I</td>
<td>30</td>
<td>244</td>
<td>190</td>
<td>100</td>
</tr>
<tr>
<td>J</td>
<td>50</td>
<td>268</td>
<td>250</td>
<td>100</td>
</tr>
<tr>
<td>K</td>
<td>53</td>
<td>306</td>
<td>260</td>
<td>147</td>
</tr>
<tr>
<td>L</td>
<td>38</td>
<td>273</td>
<td>250</td>
<td>133</td>
</tr>
</tbody>
</table>

5.1.1. Solution of SBM-Min model
First, we solved SBM-Min model and obtained the results exhibited in Table 2.

Table 2: Results of SBM-Min model

<table>
<thead>
<tr>
<th>DMU</th>
<th>Score</th>
<th>Rank</th>
<th>Reference(Lambda)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>A 1</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>1</td>
<td>B 1</td>
</tr>
<tr>
<td>C</td>
<td>0.8265</td>
<td>8</td>
<td>B 0.449</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>1</td>
<td>D 1</td>
</tr>
<tr>
<td>E</td>
<td>0.7277</td>
<td>11</td>
<td>B 0.667</td>
</tr>
<tr>
<td>F</td>
<td>0.6857</td>
<td>12</td>
<td>A 0.092</td>
</tr>
<tr>
<td>G</td>
<td>0.8765</td>
<td>6</td>
<td>B 0.16</td>
</tr>
<tr>
<td>H</td>
<td>0.7713</td>
<td>9</td>
<td>L 0.755</td>
</tr>
<tr>
<td>I</td>
<td>0.9016</td>
<td>5</td>
<td>A 0.233</td>
</tr>
<tr>
<td>J</td>
<td>0.7653</td>
<td>10</td>
<td>B 0.152</td>
</tr>
<tr>
<td>K</td>
<td>0.8619</td>
<td>7</td>
<td>B 0.15</td>
</tr>
<tr>
<td>L</td>
<td>1</td>
<td>1</td>
<td>L 1</td>
</tr>
</tbody>
</table>

We find four efficient DMUs, i.e. \( R_{\text{eff}} = \{ A, B, D, L \} \).

5.1.2. The case of inefficient DMU I
We explain the case of inefficient DMU I, step by step.
(1) Steps 1, 2 and 3: \( \rho_{i, \text{min}} = 0.9016, R_{i, \text{local}} = \{ A, L \}, R_{i, \text{eff}} = \{ A, B, D, L \} \).
(2) Step 4: \( \rho_{i, \text{pseudo-max}} = 0.9016 \).
(3) Step 5.1: Distances from efficient DMUs are \( d_A = 1.28816, d_B = 1.54030, d_D = 0.74411 \), and \( d_L = 1.03131 \). Thus, we have \( R_1 = \{ D \}, R_2 = \{ D, L \}, R_3 = \{ D, L, A \}, \) and \( R_4 = \{ D, L, A, B \} \).
(4) Step 5.2: We solve (16) and (17), and find: \( \rho_{i, 1}^* = 0.859885, \rho_{i, 2}^* = 0.910900, \rho_{i, 3}^* = 0.921168, \) and \( \rho_{i, 4}^* = 0.920198 \).
(5) Step 5.3: From (18), we find \( \rho_{i, \text{max}} = \max \{ \rho_{i, \text{pseudo-max}}, \rho_{i, 1}^*, \rho_{i, 2}^*, \rho_{i, 3}^*, \rho_{i, 4}^* \} = 0.921168 = \rho_{i, 3}^* \) with the reference set \( R_3 = \{ A, D, L \} \). Its projection is \( \{ x_{i, 1}^* = 30, x_{i, 2}^* = 205.53, y_{i, 1}^* = 190, y_{i, 2}^* = 100 \} \) with the slacks \( (s_i^* = 0, s_j^* = 38.47, s_k^* = 0, s_l^* = 0) \). The SBM-Min model has \( \rho_{i, \text{min}} = 0.9016 \) with slacks \( (s_i^* = 0, s_j^* = 26.767, s_k^* = 0, s_l^* = 9.667) \). This indicates that SBM-Min model requires a reduction of Nurse by 26.767 and an increase of Inpatient by 9.667 to attain the efficiency status, whereas SBM-Max model requires a reduction of Nurse by 38.47 to attain the efficiency status.
5.1.3. Comparisons among SBM-Max, Pseudo-Max and SBM-Min scores

Table 3 compares results with SBM-Max, SBM-Pseudo and SBM-Min scores. Inefficient DMUs increased their efficiency from SBM-Min to SBM-Max.

<table>
<thead>
<tr>
<th>DMU</th>
<th>SBM-Max</th>
<th>Rank</th>
<th>Pseudo</th>
<th>Rank</th>
<th>SBM-Min</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>0.87507</td>
<td>8</td>
<td>0.855</td>
<td>8</td>
<td>0.8265</td>
<td>8</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>E</td>
<td>0.7682</td>
<td>11</td>
<td>0.7391</td>
<td>11</td>
<td>0.7277</td>
<td>11</td>
</tr>
<tr>
<td>F</td>
<td>0.72648</td>
<td>12</td>
<td>0.6868</td>
<td>12</td>
<td>0.6857</td>
<td>12</td>
</tr>
<tr>
<td>G</td>
<td>0.93688</td>
<td>5</td>
<td>0.9052</td>
<td>5</td>
<td>0.8765</td>
<td>6</td>
</tr>
<tr>
<td>H</td>
<td>0.80918</td>
<td>10</td>
<td>0.7714</td>
<td>10</td>
<td>0.7714</td>
<td>9</td>
</tr>
<tr>
<td>I</td>
<td>0.92117</td>
<td>6</td>
<td>0.9016</td>
<td>6</td>
<td>0.9016</td>
<td>5</td>
</tr>
<tr>
<td>J</td>
<td>0.81032</td>
<td>9</td>
<td>0.7898</td>
<td>9</td>
<td>0.7653</td>
<td>10</td>
</tr>
<tr>
<td>K</td>
<td>0.88894</td>
<td>7</td>
<td>0.8622</td>
<td>7</td>
<td>0.8619</td>
<td>7</td>
</tr>
<tr>
<td>L</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Average</td>
<td>0.8947</td>
<td>0.8759</td>
<td>0.8681</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min</td>
<td>0.7265</td>
<td>0.6868</td>
<td>0.6857</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>St Dev</td>
<td>0.0982</td>
<td>0.1114</td>
<td>0.115</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4 exhibits $\rho_0^{\max}, \rho_1^*, \cdots, \rho_4^*$ for inefficient DMUs. Shaded portions indicate the max. The SBM-Max scores are found at several stage of $R_h$.

5.1.4. Comparisons of average differences between SBM-Max and SBM-Min

Fig. 2 shows the average of percentage deviations, $|\text{Data - Projection}| \times 100 / \text{Data}$. It is observed large differences exist in SBM-Min, while small differences in SBM-Max.
Table 5 reports data and projections along with deviations (%) in the case of SBM-Max.

Table 5: Data and projection by SBM-Max

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>20</td>
<td>20</td>
<td>0</td>
<td>151</td>
<td>151</td>
<td>0</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>90</td>
<td>90</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>1</td>
<td>19</td>
<td>19</td>
<td>0</td>
<td>131</td>
<td>131</td>
<td>0</td>
<td>150</td>
<td>150</td>
<td>0</td>
<td>50</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0.8751</td>
<td>8</td>
<td>25</td>
<td>25</td>
<td>0</td>
<td>160</td>
<td>155.6</td>
<td>-2.78</td>
<td>160</td>
<td>166.7</td>
<td>4.17</td>
<td>55</td>
<td>66.7</td>
<td>21.21</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>1</td>
<td>27</td>
<td>27</td>
<td>0</td>
<td>168</td>
<td>168</td>
<td>0</td>
<td>180</td>
<td>180</td>
<td>0</td>
<td>72</td>
<td>72</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>0.7682</td>
<td>11</td>
<td>22</td>
<td>20.9</td>
<td>-4.88</td>
<td>158</td>
<td>158</td>
<td>0</td>
<td>94</td>
<td>104.6</td>
<td>11.32</td>
<td>66</td>
<td>94.2</td>
<td>42.69</td>
</tr>
<tr>
<td>F</td>
<td>0.7265</td>
<td>12</td>
<td>55</td>
<td>34.5</td>
<td>-37.27</td>
<td>255</td>
<td>214.7</td>
<td>-15.82</td>
<td>230</td>
<td>230</td>
<td>0</td>
<td>90</td>
<td>92</td>
<td>2.22</td>
</tr>
<tr>
<td>G</td>
<td>0.9369</td>
<td>5</td>
<td>33</td>
<td>33</td>
<td>0</td>
<td>235</td>
<td>205.3</td>
<td>-12.62</td>
<td>220</td>
<td>220</td>
<td>0</td>
<td>88</td>
<td>88</td>
<td>0</td>
</tr>
<tr>
<td>H</td>
<td>0.8092</td>
<td>10</td>
<td>31</td>
<td>30</td>
<td>-3.23</td>
<td>206</td>
<td>186.7</td>
<td>-9.39</td>
<td>152</td>
<td>200</td>
<td>31.58</td>
<td>80</td>
<td>80</td>
<td>0</td>
</tr>
<tr>
<td>I</td>
<td>0.9212</td>
<td>6</td>
<td>30</td>
<td>30</td>
<td>0</td>
<td>244</td>
<td>205.5</td>
<td>-15.77</td>
<td>190</td>
<td>190</td>
<td>0</td>
<td>100</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>J</td>
<td>0.8103</td>
<td>9</td>
<td>50</td>
<td>43.1</td>
<td>-13.86</td>
<td>268</td>
<td>268</td>
<td>0</td>
<td>250</td>
<td>287.1</td>
<td>14.86</td>
<td>100</td>
<td>115</td>
<td>14.86</td>
</tr>
<tr>
<td>K</td>
<td>0.888</td>
<td>7</td>
<td>53</td>
<td>46.5</td>
<td>-12.26</td>
<td>306</td>
<td>306</td>
<td>0</td>
<td>260</td>
<td>289.1</td>
<td>11.2</td>
<td>147</td>
<td>147</td>
<td>0</td>
</tr>
<tr>
<td>L</td>
<td>1</td>
<td>1</td>
<td>38</td>
<td>38</td>
<td>0</td>
<td>273</td>
<td>273</td>
<td>0</td>
<td>250</td>
<td>250</td>
<td>0</td>
<td>133</td>
<td>133</td>
<td>0</td>
</tr>
</tbody>
</table>

5.2. Japanese municipal hospitals

The data were collected from the Annual Databook of Local Public Enterprise published by the Ministry of Internal Affairs and Communications Japanese Government, 2005.

5.2.1. Data

Number of DMUs: 707 hospitals ($n = 707$).
Number of inputs: 5. (1) No. of beds (Bed), (2) Expenses for outsourcing (Outsource), (3) No. of doctors (Doctor), (4) No. of nurses (Nurse) and (5) Expenses for other medical materials (Material). ($m = 5$)
Number of outputs: 4. (1) Revenue from operation per day (Operation), (2) Revenue from first consultation per day (1st time), (3) Revenue from return to clinic per day (Follow-up) and (4) Revenue from hospitalization per day (Hotel). ($s = 4$)
Table 6 exhibits statistics of the dataset.

Table 6: Statistics of dataset ($n = 707$)

<table>
<thead>
<tr>
<th></th>
<th>Bed</th>
<th>Outsource</th>
<th>Doctor</th>
<th>Nurse</th>
<th>Material</th>
<th>Operation</th>
<th>1st time</th>
<th>Follow-up</th>
<th>Hotel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max</td>
<td>1063</td>
<td>2231247</td>
<td>215.562</td>
<td>955.464</td>
<td>3395791</td>
<td>1.7E+07</td>
<td>1432079</td>
<td>3359160</td>
<td>1.8E+07</td>
</tr>
<tr>
<td>Min</td>
<td>25</td>
<td>7767</td>
<td>0.98</td>
<td>11</td>
<td>9197</td>
<td>8979</td>
<td>2706</td>
<td>13636</td>
<td>109650</td>
</tr>
<tr>
<td>Average</td>
<td>255.924</td>
<td>312686</td>
<td>33.2783</td>
<td>175.709</td>
<td>491909</td>
<td>2128506</td>
<td>211538</td>
<td>415872</td>
<td>3263845</td>
</tr>
<tr>
<td>SD</td>
<td>191.764</td>
<td>334184</td>
<td>34.1647</td>
<td>149.678</td>
<td>598474</td>
<td>2533050</td>
<td>212439</td>
<td>327749</td>
<td>3020809</td>
</tr>
</tbody>
</table>

5.2.2. SBM scores

The SBM-Min model found that 66 hospitals among 707 are efficient ($N_{eff} = 66$). Table 7 compares three scores. We found large differences between Max and Min models.

Table 7: Comparisons of three scores

<table>
<thead>
<tr>
<th></th>
<th>SBM-Max</th>
<th>Pseudo</th>
<th>SBM-Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>0.7835</td>
<td>0.6997</td>
<td>0.4515</td>
</tr>
<tr>
<td>Max</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Min</td>
<td>0.1889</td>
<td>0.0394</td>
<td>0.0118</td>
</tr>
<tr>
<td>St Dev</td>
<td>0.1339</td>
<td>0.211</td>
<td>0.229</td>
</tr>
</tbody>
</table>

Fig. 3 (SBM-Max) and Fig. 4 (SBM-Min) exhibit respectively scores of 707 hospitals in ascending order where we can observe big differences.
Table 8 shows the average of percentage deviations, $\frac{\text{Data} - \text{Projection}}{\text{Data}} \times 100$ / Data. It is observed that large differences exist in SBM-Min, while small differences in SBM-Max.

Table 8: Average deviation (%)

<table>
<thead>
<tr>
<th>DMU</th>
<th>SBM-Max</th>
<th>SBM-Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bed</td>
<td>13.1644</td>
<td>3.4014</td>
</tr>
<tr>
<td>Outsource</td>
<td>20.0291</td>
<td>24.8771</td>
</tr>
<tr>
<td>Doctor</td>
<td>12.0707</td>
<td>9.9367</td>
</tr>
<tr>
<td>Nurse</td>
<td>9.6586</td>
<td>5.7022</td>
</tr>
<tr>
<td>Material</td>
<td>10.7101</td>
<td>8.144</td>
</tr>
<tr>
<td>Operation</td>
<td>12.7155</td>
<td>48.2925</td>
</tr>
<tr>
<td>1st time</td>
<td>9.4884</td>
<td>407.243</td>
</tr>
<tr>
<td>Follow-up</td>
<td>13.3537</td>
<td>192.406</td>
</tr>
<tr>
<td>Hotel</td>
<td>17.4181</td>
<td>1.1889</td>
</tr>
</tbody>
</table>
Fig. 5 illustrates deviations graphically. Large differences found in SBM-Min, while balanced deviations in SBM-Max.

![Graph showing deviations between SBM-Min and SBM-Max](image)

**Figure 5: Average deviations (%) of inputs and outputs (cut at 150%).**

### 5.2.3. Computational time

The computational time increases as the number of efficient DMUs ($N_{eff}$) increases, because number of facets increases accordingly and we need to solve additional $N_{eff}$ LPs. In this example we have:

1. CPU time for SBM-Min and SBM-Pseudo = 12 seconds.
2. CPU time for SBM-Max = 179 seconds.

SBM-Max needs about 15 times of SBM-Min and SBM-Pseudo. This number is reasonable and consistent with the formula (20).

### 6. Conclusions

In this paper, we have developed the SBM-Max model which attempts to find nearly closest reference point on the efficient frontiers so that slacks are minimized, while the scores are maximized. Sacrificing the rigorous solutions, the proposed model utilizes a standard LP code and finds approximate solutions in allowable (polynomial) times. Inefficient DMUs can be improved to the efficient status with less input-reductions and less output-enlargement. Thus, this model proposes a Kaizen (improvement) tool by DEA.

Many applications of the SBM-Min models have been developed over the world. Also many DEA models are developed based on this model. Above all, Network SBM (NSBM) (Tone and Tsutsui (2009)), Dynamic SBM (DSBM) (Tone and Tsutsui (2010)), Dynamic and Network SBM (DNSBM) (Tone and Tsutsui (2014)), and Malmquist SBM (Cooper at al. (2007)) are representative. Revisions of these models based on the SBM-Max model are imperative future research subjects.
References