



# A directional distance based super-efficiency DEA model handling negative data

Ruiyue Lin<sup>1,2</sup> and Zhiping Chen<sup>1\*</sup>

<sup>1</sup>Department of Computing Science, School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, Shaanxi Province, People's Republic of China; and <sup>2</sup>School of Mathematics and Information Science, Wenzhou University, Higher Education Zone, Wenzhou 325035, Zhejiang Province, People's Republic of China

This paper develops a new radial super-efficiency data envelopment analysis (DEA) model, which allows input–output variables to take both negative and positive values. Compared with existing DEA models capable of dealing with negative data, the proposed model can rank the efficient DMUs and is feasible no matter whether the input–output data are non-negative or not. It successfully addresses the infeasibility issue of both the conventional radial super-efficiency DEA model and the Nerlove–Luenberger super-efficiency DEA model under the assumption of variable returns to scale. Moreover, it can project each DMU onto the super-efficiency frontier along a suitable direction and never leads to worse target inputs or outputs than the original ones for inefficient DMUs. Additional advantages of the proposed model include monotonicity, units invariance and output translation invariance. Two numerical examples demonstrate the practicality and superiority of the new model.

*Journal of the Operational Research Society* (2017) **68**(11), 1312–1322. doi:10.1057/s41274-016-0137-8;

published online 9 January 2017

**Keywords:** data envelopment analysis; negative data; directional distance function; super-efficiency; infeasibility

## 1. Introduction

Data envelopment analysis (DEA) is a non-parameter technique for measuring the relative efficiency of a set of peer decision-making units (DMUs) with multiple inputs and outputs (Charnes *et al.*, 1978). A weakness of traditional DEA models is that it assumes that all the inputs and outputs are non-negative. However, negative values, especially those of outputs, could exist in many situations. For example, the expected return is generally treated as an output measure for estimating the efficiency of mutual fund, which might be negative for some mutual funds. Likewise, the profit, which is generally chosen as an output for measuring the efficiency of projects, might be negative for some projects. Hence, it is necessary to improve the DEA model to expand its application.

There are several approaches to deal with negative data in DEA models. The simplest method is to treat negative inputs (outputs) as outputs (inputs). If inputs (outputs) are all non-positive, their absolute values can be treated as non-negative outputs (inputs), so that non-positive inputs (outputs) will decrease (increase) when those corresponding non-negative outputs (inputs) expand (Scheel, 2001). However, this method

is not applicable if there is some input or output with both positive and negative values. Another approach to handle negative data is to utilize the “translation invariance” property. A DEA model is translation invariant if the translated input–output data yield the same results as the original data. The variable returns to scale (VRS) additive DEA models (Charnes *et al.*, 1984) are translation invariant (Ali and Seiford, 1990; Lovell and Pastor, 1995; Pastor, 1996), but they yield the “furthest” target on the production frontier for inefficient DMUs (Portela *et al.*, 2004) and cannot provide any measure of efficiency. The output-oriented BCC model (Banker *et al.*, 1984) is input translation invariant, and the input-oriented BCC model is output translation invariant. These two kinds of BCC models cannot be applied to the situation where negative values exist in both inputs and outputs. Based on a modified directional distance function (DDF) (Chambers *et al.*, 1996), Portela *et al.* (2004) develop a range directional measure (RDM) model, which can deal with inputs and/or outputs taking positive values for some DMUs and negative values for the others. However, the RDM model may be unbounded when the evaluated DMU has the maximum values for all the outputs and the minimum values for all the inputs (Cheng *et al.*, 2013). Inspired by the RDM model, Sharp *et al.* (2007) introduce a modified slack-based measure, which can deal with negative inputs and/or negative outputs. Emrouznejad *et al.* (2010) propose a semi-oriented radial measure (SORM) to handle negative input–output data.

\*Correspondence: Zhiping Chen, Department of Computing Science, School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, Shaanxi Province, People's Republic of China.  
E-mail: zchen@mail.xjtu.edu.cn

Kerstens and Woestyne (2011) recommend a generalized Farrell proportional distance function that handles negative data and maintains a proportional interpretation under mild conditions. Cheng *et al* (2013) find that the SORM model might lead to worse target inputs or outputs than the original ones for inefficient DMUs. They develop a variant of the traditional radial model where original values are replaced with their absolute values. Some imprecisions in Cheng *et al* (2013) are corrected by Kerstens and Woestyne (2014).

A limitation of the above DEA models is that they cannot further discriminate efficient DMUs, all of which have an efficiency score of unity. Andersen and Petersen (1993) develop a super-efficiency DEA model, see also Banker *et al* (1989), which can rank efficient DMUs. The input-oriented (output-oriented) super-efficiency DEA model excludes the DMU under evaluation from the reference set so that efficient DMUs may have efficiency scores larger (smaller) than or equal to one. The original super-efficiency DEA model is introduced under the condition of constant returns to scale (CRS) and is feasible if all inputs and outputs of DMUs are positive. However, the infeasibility issue might occur in VRS super-efficiency DEA models (Seiford and Zhu, 1999).

Many modified VRS radial super-efficiency DEA models (Chen, 2005; Ray, 2008; Cook *et al*, 2009; Lee *et al*, 2011) have been proposed to address the infeasibility issue. Among them, the VRS Nerlove–Luenberger super-efficiency DEA model (Ray, 2008) is based on the DDF and is very often feasible under the non-negative data set. However, this model fails in two exceptions (see Ray, 2008 for details). By choosing proper directions, Chen *et al* (2013) and Lin and Chen (2015) propose two DDF-based VRS super-efficiency DEA models to eliminate the infeasibility in two exceptions. The model in Chen *et al* (2013) may be infeasible if zero data exist in outputs (Lin and Chen, 2015). All these modified super-efficiency DEA models are proposed for the non-negative data. In the situations where there exist negative inputs or outputs, the infeasibility issue still exists. Based on the RDM model (Portela *et al*, 2004), Hadi-Vencheh and Esmailzadeh (2013) propose two super-efficiency models in the presence of negative data, called the super RDM+ model and the super RDM—model, respectively. Both models can rank efficient DMUs, but they are still infeasible in some cases. We will illustrate this point in Section 2.

In this paper, we propose a novel DDF-based VRS radial super-efficiency DEA model which is feasible and is able to handle negative data. There are at least five contributions in this article.

1. By choosing a proper direction for the DDF, we propose an alternative VRS radial super-efficiency DEA model. The proposed model not only successfully addresses the infeasibility problem in VRS radial super-efficiency DEA models, but also extends the application of the super-efficiency measure to negative data.

2. The proposed model projects each DMU onto the super-efficiency frontier along a suitable direction and provides improved targets for inefficient DMUs.
3. The proposed model yields a bounded measure of super-efficiency.
4. In the situation where outputs are all non-negative, the proposed model always generates reference points with non-negative outputs.
5. The proposed model is monotonous, units invariant and output translation invariant.

The rest of the paper is organized as follows. Section 2 presents the DDF and some existing directions, whose limitations are illustrated through a numerical example. Section 3 proposes a modified DDF, and based on it, we develop a new VRS radial super-efficiency DEA model capable of dealing with negative data, whose useful properties are also investigated in this section. In Section 4, the proposed model is applied to the numerical example in Section 2 and a data set from the literature, respectively, in order to demonstrate its properties and merits. Conclusions are presented in the last section.

## 2. DDF-based super-efficiency and directions

Assume that there are  $n$  DMUs, each DMU has  $m$  inputs and  $s$  outputs, and each of inputs and outputs has at least one nonzero value. For each DMU  $j$  ( $j = 1, \dots, n$ ), let  $x_{ij}$  ( $i = 1, \dots, m$ ) denote the  $i$ th input and  $y_{rj}$  ( $r = 1, \dots, s$ ) denote the  $r$ th output. Under the standard assumptions of convexity and free disposability of inputs and outputs (Chen *et al*, 2013), the production possibility set (PPS) for a target DMU  $p$  ( $p \in \{1, \dots, n\}$ ) with respect to super-efficiency is spanned by  $(x_{ij}, y_{rj}), j = 1, \dots, n, j \neq p$ , as follows.

$$T_p = \left\{ (x_i, y_r) \mid \begin{aligned} &x_i \geq \sum_{j=1, j \neq p}^n \lambda_j x_{ij}, \quad i = 1, \dots, m; \\ &y_r \leq \sum_{j=1, j \neq p}^n \lambda_j y_{rj}, \quad r = 1, \dots, s; \\ &\sum_{j=1, j \neq p}^n \lambda_j = 1, \quad \lambda_j \geq 0, \quad j = 1, \dots, n, \quad j \neq p \end{aligned} \right\}$$

Choosing a direction vector  $(g^x, g^y)$ , the directional distance function (DDF) for DMU $_p$  with respect to  $T_p$  is defined as:

$$D(x_{ip}, y_{rp}; g^x, g^y) = \max \beta_p : (x_{ip} - \beta_p g^x, y_{rp} + \beta_p g^y) \in T_p.$$

Then, the following general DDF-based super-efficiency DEA model can be established.

$$\max \beta_p \tag{1}$$

$$\text{s.t.} \quad \sum_{j=1, j \neq p}^n \lambda_j x_{ij} \leq x_{ip} - \beta_p g^x, \quad i = 1, \dots, m, \tag{2}$$

$$\sum_{j=1, j \neq p}^n \lambda_j y_{rj} \geq y_{rp} + \beta_p g^y, \quad r = 1, \dots, s, \quad (3)$$

$$\sum_{j=1, j \neq p}^n \lambda_j = 1, \quad \lambda_j \geq 0, \quad j = 1, \dots, n, \quad j \neq p. \quad (4)$$

Denote the optimum value of model (1)–(4) as  $\beta_p^o$ . The super-efficiency score of the evaluated DMU<sub>p</sub> can be determined as  $1 - \beta_p^o$  (Ray, 2008). The smaller the value of  $\beta_p^o$ , the more efficient the DMU<sub>p</sub>. For any efficient DMU<sub>p</sub>,  $1 - \beta_p^o$  is no less than 1.0.

The direction vector  $(g^x, g^y)$  should be non-negative and nonzero and can be chosen in an arbitrary way (Chen *et al*, 2013; Ray, 2008). Bricc and Kerstens (2009a) point out that model (1)–(4) cannot guarantee the feasibility if the direction is a constant vector and the output direction vector is nonzero. Hence,  $g^x$  and  $g^y$  are often considered as the function of  $x_{ip}$  and  $y_{rp}$ . If all input and output data are non-negative, the standard DDF for the DMU<sub>p</sub> is adopted by choosing  $(x_{ip}, y_{rp})$  as  $(g^x, g^y)$  (Chambers *et al*, 1996), and the VRS Nerlove–Luenberger super-efficiency DEA model (Ray, 2008) (called the NL model for short) is obtained. The NL model is very often feasible for non-negative data, but it fails in the following two exceptions (Ray, 2008):

1. When a super-efficiency score is greater than 2.0, the NL model will yield a reference point with negative outputs. In applications where the outputs should be non-negative, such as the performance evaluation of airlines (Ray, 2008), a reference point with negative outputs results in a conceptual problem.
2. If a zero input exists in the evaluated DMU and all other DMUs in the reference set have positive values in that input, the NL model becomes infeasible.

In the context of the definition of a Luenberger productivity indicator, Bricc and Kerstens (2009b) avoid reference points with negative outputs by adding a constraint that the projected output should remain positive. However, this method still cannot solve the infeasibility issue in the second exception. To eliminate the infeasibility in both the exceptions, Lin and Chen (2015) recently have put forward a modified DDF-based super-efficiency DEA model (called the LC model for short)

by choosing  $(x_{ip} + \max_{j=1, \dots, n, j \neq p} \{x_{ij}\}, y_{rp})$  as  $(g^x, g^y)$ . The LC model successfully addresses the infeasibility issue in conventional VRS radial super-efficiency DEA models and the NL model under non-negative data.

In the presence of negative data, both the NL and LC models might be infeasible. This is because their related direction vectors,  $(x_{ip}, y_{rp})$  and  $(x_{ip} + \max_{j=1, \dots, n, j \neq p} \{x_{ij}\}, y_{rp})$ , might be negative, which could guide the DMU<sub>p</sub> to be further away from the super-efficiency frontier and thus lead to infeasibility. To illustrate this problem, we consider a simple example with six DMUs. Each DMU has three inputs and two outputs. The concrete data set of this example is shown in columns 2 to 6 of Table 1. Input X3 and output Y1 are negative for DMU 2 and DMU 1, respectively. The super-efficiency results yielded by the NL model and the LC model are shown in columns 7 and 8 of Table 1, respectively. Obviously, both the NL model and the LC model become infeasible for some DMUs.

In order to rank DMUs in the presence of negative data, Hadi-Vencheh and Esmailzadeh (2013) propose the super RDM+ model and the super RDM– model, respectively. Let

$$P_{ip}^- = x_{ip} - \min_j \{x_{ij}\}, \quad P_{rp}^+ = \max_j \{y_{ij}\} - y_{rp} \quad (5)$$

(Portela *et al*, 2004). By employing  $(P_{ip}^-, P_{rp}^+)$  and  $(\frac{1}{P_{ip}^-}, \frac{1}{P_{rp}^+})$  as the direction vector  $(g^x, g^y)$  in model (1)–(4), respectively, the dual of the super RDM+ model and the dual of the super RDM– model can be obtained. If there exists some  $i$  or  $r$  such that  $P_{ip}^- = 0$  or  $P_{rp}^+ = 0$ , the direction  $(\frac{1}{P_{ip}^-}, \frac{1}{P_{rp}^+})$  would be meaningless. To solve this problem, we can use 0 to replace  $\frac{1}{P_{ip}^-}$  and  $\frac{1}{P_{rp}^+}$ , as Portela *et al* (2004), Hadi-Vencheh and Esmailzadeh (2013) proposed. However, these two models still fail in some cases. Let us consider DMU 6 in the above example. According to (5), we can find  $P_{16}^- = 9, P_{26}^- = 0, P_{36}^- = 2, P_{16}^+ = 0, P_{26}^+ = 3$ . Then for both the dual super RDM+ model and the dual super RDM– model, the constraint (3) with respect to the output Y1 of DMU 6 is expressed as

$$-8\lambda_1 + 4\lambda_3 + \lambda_4 + 3\lambda_5 \geq 6 + 0 \cdot \beta = 6. \quad (6)$$

According to (4), we have  $-8\lambda_1 + 4\lambda_3 + \lambda_4 + 3\lambda_5 \leq 4$ . This contradicts with (6). Therefore, both the dual super RDM+

**Table 1** Input–output data and evaluation results

DMU <sub>p</sub>	X1	X2	X3	Y1	Y2	Super-efficiency			
						NL	LC	Dual super RDM+	Dual super RDM–
1	0.8	0	0	–8	0	Infeasible	1.102	2.500	1.960
2	0	1	–2	0	3	Infeasible	Infeasible	Infeasible	Infeasible
3	4	0	0	4	0	1.025	1.015	1.038	1.235
4	4	0	0	1	0	0.460	1.000	1.000	1.000
5	2	0	0	3	0	1.619	1.118	1.619	3.943
6	9	0	0	6	0	1.333	1.333	Infeasible	Infeasible

model and the dual super RDM– model are infeasible for DMU 6. According to the duality theory of linear programs, the super RDM+ and the super RDM– models are infeasible or have an unbounded optimal value for DMU 6. Actually, these two models fail for DMUs 2 and 6, as one can see from the last two columns of Table 1.

### 3. Modified DDF-based super-efficiency model

Considering that negative values might exist in the input-output data, we need to choose a new direction vector which is always non-negative and nonzero, independent of inputs and outputs being non-negative or not. To this end, we introduce the following constants:

$$a_i = k \cdot \max_{j=1, \dots, n} \{ |x_{ij}| \}, \quad i = 1, \dots, m, \quad (7)$$

$$b_r = \min_{j=1, \dots, n} \{ y_{rj} \}, \quad r = 1, \dots, s, \quad (8)$$

where  $k$  is a constant satisfying  $k \geq 3$ . Clearly, we have  $x_{ip} + a_i > 0$  and  $y_{rp} - b_r \geq 0$  for all  $i = 1, \dots, m$ , and  $r = 1, \dots, s$ , respectively. Therefore, we can choose  $(x_{ip} + a_i, y_{rp} - b_r)$  as  $(g^x, g^y)$  and then obtain the following VRS radial super-efficiency DEA model:

$$\max \beta_p \quad (9)$$

$$\text{s.t.} \quad \sum_{j=1, j \neq p}^n \lambda_j x_{ij} \leq (1 - \beta_p)x_{ip} - a_i \beta_p, \quad i = 1, \dots, m, \quad (10)$$

$$\sum_{j=1, j \neq p}^n \lambda_j y_{rj} \geq (1 + \beta_p)y_{rp} - b_r \beta_p, \quad r = 1, \dots, s, \quad (11)$$

$$\sum_{j=1, j \neq p}^n \lambda_j = 1, \quad \lambda_j \geq 0, \quad j = 1, \dots, n, \quad j \neq p. \quad (12)$$

For the sake of distinction, we denote the optimum value of model (9)–(12) as  $\beta_p^*$ . The following proposition shows the feasibility of model (9)–(12).

**Proposition 1** *Model (9)–(12) is feasible, and we have  $0 \leq \beta_p^* < 1$  for  $(x_{ip}, y_{rp}) \in T_p$  and  $-1 \leq \beta_p^* < 0$  for  $(x_{ip}, y_{rp}) \notin T_p$ .*

**Proof** Since  $x_{ip} + a_i > 0$ , constraints (10) can be rewritten as

$$\beta_p \leq \frac{x_{ip} - \sum_{j=1, j \neq p}^n \lambda_j x_{ij}}{x_{ip} + a_i}, \quad i = 1, \dots, m. \quad (13)$$

For each  $p \in \{1, \dots, n\}$ , let  $J_p = \{r | y_{rp} - b_r > 0, r = 1, \dots, s\}$  and  $O_p = \{r | y_{rp} - b_r = 0, r = 1, \dots, s\}$ . As  $y_{rp} - b_r \geq 0$ , we have  $J_p \cup O_p = \{1, \dots, s\}$ .

According to (8), (11) and (12), we have

$$\sum_{j=1, j \neq p}^n \lambda_j y_{rj} \geq \min_{\substack{j=1, \dots, n \\ j \neq p}} \{ y_{rj} \} \geq \min_{j=1, \dots, n} \{ y_{rj} \} = b_r = y_{rp}, \quad r \in O_p. \quad (14)$$

It is easy to see from (14) that the inequalities in (11) naturally hold for all  $r \in O_p$ . Hence, constraints (11) are equivalent to

$$\beta_p \leq \frac{\sum_{j=1, j \neq p}^n \lambda_j y_{rj} - y_{rp}}{y_{rp} - b_r}, \quad r \in J_p. \quad (15)$$

With (13) and (15), we consider the following two complementary cases for each DMU  $p \in \{1, \dots, n\}$ .

**Case I**  $(x_{ip}, y_{rp}) \in T_p$ : In this case, we have  $x_{ip} \geq \sum_{j=1, j \neq p}^n \lambda_j x_{ij}$  and  $y_{rp} \leq \sum_{j=1, j \neq p}^n \lambda_j y_{rj}$  for  $i = 1, \dots, m$  and  $r = 1, \dots, s$ , respectively. So,

$$\frac{x_{ip} - \sum_{j=1, j \neq p}^n \lambda_j x_{ij}}{x_{ip} + a_i} \geq 0, \quad i = 1, \dots, m, \quad (16)$$

and

$$\frac{\sum_{j=1, j \neq p}^n \lambda_j y_{rj} - y_{rp}}{y_{rp} - b_r} \geq 0, \quad r \in J_p. \quad (17)$$

Inequalities (13), (15), (16) and (17) mean that  $\beta_p = 0$  is a feasible solution of model (9)–(12). Then as the optimum value of problem (9)–(12),  $\beta_p^* \geq 0$  always holds for  $p = 1, \dots, n$ . In addition, we have from (7), (12) and (13) that

$$\begin{aligned} \beta_p &\leq \frac{x_{ip} - \sum_{j=1, j \neq p}^n \lambda_j x_{ij}}{x_{ip} + a_i} \leq \frac{x_{ip} + \max_{j=1, \dots, n, j \neq p} \{ |x_{ij}| \}}{x_{ip} + a_i} \\ &\leq \frac{x_{ip} + \max_{j=1, \dots, n} \{ |x_{ij}| \}}{x_{ip} + a_i} < 1. \end{aligned}$$

Thus in this case, we have  $0 \leq \beta_p^* < 1$ .

**Case II**  $(x_{ip}, y_{rp}) \notin T_p$ : We know from (14) that the  $r$  satisfying  $\sum_{j=1, j \neq p}^n \lambda_j y_{rj} < y_{rp}$  does not belong to  $O_p$ . Hence

in this case, there must exist some  $i \in \{1, \dots, m\}$  such that  $\sum_{j=1, j \neq p}^n \lambda_j x_{ij} > x_{ip}$  or some  $r \in J_p$  such that  $\sum_{j=1, j \neq p}^n \lambda_j y_{rj} < y_{rp}$ .

We can then easily see from (13) and (15) that model (9)–(12) is still feasible in this case and its optimal

value,  $\beta_p^*$ , is negative. Indeed, since  $k \geq 3$ , we know from (7) that

$$\max_{j=1, \dots, n} \{ |x_{ij}| \} \leq 2x_{ip} + a_i. \tag{18}$$

From (12) and (18), we have

$$\begin{aligned} \frac{x_{ip} - \sum_{\substack{j=1 \\ j \neq p}}^n \lambda_j x_{ij}}{x_{ip} + a_i} &\geq \frac{x_{ip} - \max_{j=1, \dots, n} \{ |x_{ij}| \}}{x_{ip} + a_i} \\ &\geq \frac{x_{ip} - 2x_{ip} - a_i}{x_{ip} + a_i} = -1, \quad i = 1, \dots, m. \end{aligned} \tag{19}$$

On the other hand, we obtain from (8) and (12) that

$$\begin{aligned} \frac{\sum_{\substack{j=1 \\ j \neq p}}^n \lambda_j y_{rj} - y_{rp}}{y_{rp} - b_r} &\geq \frac{\min_{j=1, \dots, n} \{ y_{rj} \} - y_{rp}}{y_{rp} - b_r} \\ &= \frac{b_r - y_{rp}}{y_{rp} - b_r} = -1, \quad r \in J_p. \end{aligned} \tag{20}$$

Since we maximize  $\beta_p$  in model (9)–(12),  $\beta_p^* \geq -1$ ,  $p = 1, \dots, n$ , always holds due to (13), (15), (19) and (20). Therefore,  $-1 \leq \beta_p^* < 0$  for  $(x_{ip}, y_{rp}) \notin T_p$ .  $\square$

According to Proposition 1, no matter whether inputs and outputs are non-negative or not, for each DMU whose input–output bundle belongs to  $T_p$ , model (9)–(12) expands its outputs and reduces its inputs simultaneously to reach the super-efficiency frontier formed by the rest of DMUs; for each DMU whose input–output bundle does not belong to  $T_p$ , model (9)–(12) reduces (expands) at least one of its outputs (inputs) to reach the super-efficiency frontier formed by the remaining DMUs.

Inspired by the criterion for identifying efficiency in conventional DEA models such as the CCR model (Charnes *et al*, 1978), we judge whether the evaluated DMU is efficient or not under model (9)–(12) by the following criterion: *The DMU  $p$  is inefficient if  $\beta_p^* > 0$  or  $\beta_p^* = 0$  with positive slacks yielded by model (9)–(12); otherwise, it is efficient.* Let  $\lambda_j^*$  denote the optimal maximum slack solution (Cooper *et al*, 2007) of model (9)–(12). Then, the inputs and outputs of the projection on the super-efficiency frontier with respect to the DMU  $p$  can be expressed as

$$\hat{x}_{ip} = \sum_{j=1, j \neq p}^n \lambda_j^* x_{ij}, \quad i = 1, \dots, m, \tag{21}$$

$$\hat{y}_{rp} = \sum_{j=1, j \neq p}^n \lambda_j^* y_{rj}, \quad r = 1, \dots, s. \tag{22}$$

Like traditional DEA models, model (9)–(12) might provide multiple projections for the evaluated DMU because there might exist multiple optimal solutions for  $\lambda_j^*$ . From Proposition 1 and constraints (10) and (11), we know that model (9)–(12) projects the DMU  $p$  onto the super-efficiency frontier formed by the rest of DMUs along the direction  $(x_{ip} + a_i, y_{rp} - b_r)$ , without the actual data transformation. Concretely, for inefficient DMUs, it preserves the proportionate improvement property of the traditional DEA model and any projection it provides can be seen as an improved target for the evaluated DMU since  $\hat{x}_{ip} = \sum_{\substack{j=1 \\ j \neq p}}^n \lambda_j^* x_{ij} \leq x_{ip} - \beta_p^* (x_{ip} + a_i) \leq x_{ip}$ ,  $\hat{y}_{rp} = \sum_{\substack{j=1 \\ j \neq p}}^n \lambda_j^* y_{rj} \geq y_{rp} + \beta_p^* (y_{rp} - b_r) \geq y_{rp}$ ,  $\forall i, r$ , hold; for efficient DMUs, at least one of the outputs (or inputs) of projections is not larger (or not smaller) than that of the corresponding DMU due to the positive  $\beta_p^*$  or  $\beta_p^* = 0$  with all zero slacks. Thus, model (9)–(12) generalizes current VRS radial super-efficiency DEA models suitable for non-negative data to the situation with partially or fully negative data.

Similarly to the determination of the super-efficiency score in Ray (2008), the DDF-based super-efficiency score for DMU  $p$  is determined by  $1 - \beta_p^*$  under model (9)–(12). From Proposition 1, we have the following corollary about the boundedness of the super-efficiency score determined by model (9)–(12).

**Corollary 1**  $(1 - \beta_p^*) \in (0, 2]$ .

We know from Ray (2008) that under model (9)–(12), the outputs of the reference point for DMU  $p$  are

$$\tilde{y}_{rp} = (1 + \beta_p^*)y_{rp} - b_r \beta_p^*, \quad r = 1, \dots, s. \tag{23}$$

According to Proposition 1, we have

$$\begin{aligned} \tilde{y}_{rp} &= y_{rp} + \beta_p^* (y_{rp} - b_r) \geq y_{rp}, \quad (x_{ip}, y_{rp}) \in T_p, \\ \tilde{y}_{rp} &\geq y_{rp} - (y_{rp} - b_r) = b_r, \quad (x_{ip}, y_{rp}) \notin T_p. \end{aligned}$$

Therefore, we have the following conclusion for  $\tilde{y}_{rp}$ :

**Corollary 2** *For the data set with non-negative outputs,  $\tilde{y}_{rp} \geq 0$  holds for any DMU  $p$ ,  $p \in \{1, \dots, n\}$ .*

Corollary 2 shows that the conceptual problem described in Ray (2008) does not occur under our model. From Proposition 1 and Corollary 2, we know that the proposed model eliminates the infeasibility issue of the conventional VRS super-efficiency DEA model and the NL model and meanwhile extends the applicability of the VRS radial super-efficiency DEA model to the situation with negative input–output data.

Furthermore, our model has the following three useful properties.

(i) *Monotonicity*

Suppose that the inputs of DMU  $p$  are reduced to  $x_{ip} - \Delta x_{ip}$  and the outputs of DMU  $p$  are increased to  $y_{rp} + \Delta y_{rp}$ , here

$\Delta x_{ip} \geq 0$  and  $\Delta y_{rp} \geq 0$  for all  $i = 1, \dots, m$ , and  $r = 1, \dots, s$ , respectively. Notice that here the input (output) data of the DMU  $p$  are decreased (increased). According to the principle, we introduced the constants  $a_i$  and  $b_r$  in (7) and (8), and these two constants should be determined by considering all possible values of inputs and outputs. Therefore,  $a_i$  and  $b_r$  in this situation should be determined by

$$a_i = k \cdot \max\{|x_{ij}|, \forall j, |x_{ip} - \Delta x_{ip}|\}, i = 1, \dots, m, \quad (24)$$

$$b_r = \min\{y_{rj}, \forall j, y_{rp} + \Delta y_{rp}\} = \min_{j=1, \dots, n} \{y_{rj}\}, r = 1, \dots, s. \quad (25)$$

Equality (25) holds due to the non-negativity of  $\Delta y_{rp}$ . With  $a_i$  and  $b_r$  in (24) and (25), we have the following conclusion:

**Proposition 2** *The optimal value of model (9)–(12) does not increase if inputs (outputs) of the DMU  $p$  are reduced (increased).*

**Proof** After the specified input reduction and output expansion, the direction vector becomes  $(x_{ip} - \Delta x_{ip} + a_i, y_{rp} + \Delta y_r - b_r)$ . With  $a_i$  and  $b_r$  in (24) and (25), we have  $x_{ip} - \Delta x_{ip} + a_i > 0$ ,  $i = 1, \dots, m$ , and  $y_{rp} + \Delta y_r - b_r \geq 0$ ,  $r = 1, \dots, s$ . Then, the corresponding model (9)–(12) for the DMU  $p$  should be rewritten as

$$\max \beta_p \quad (26)$$

s.t.

$$\sum_{j=1, j \neq p}^n \lambda_j x_{ij} \leq (1 - \beta_p)(x_{ip} - \Delta x_{ip}) - a_i \beta_p, i = 1, \dots, m, \quad (27)$$

$$\sum_{j=1, j \neq p}^n \lambda_j y_{rj} \geq (1 + \beta_p)(y_{rp} + \Delta y_{rp}) - b_r \beta_p, r = 1, \dots, s, \quad (28)$$

$$\sum_{j=1, j \neq p}^n \lambda_j = 1, \lambda_j \geq 0, j = 1, \dots, n, j \neq p. \quad (29)$$

Denote the optimal solution of model (26)–(29) as  $(\beta'_p, \lambda'_j)$ . By using the same derivation as that for (13) and (15), constraints (27) and (28) are equivalent to

$$\beta_p \leq \frac{x_{ip} - \Delta x_{ip} - \sum_{j=1, j \neq p}^n \lambda_j x_{ij}}{x_{ip} - \Delta x_{ip} + a_i}, i = 1, \dots, m, \quad (30)$$

and

$$\beta_p \leq \frac{\sum_{j=1, j \neq p}^n \lambda_j y_{rj} - y_{rp} - \Delta y_{rp}}{y_{rp} + \Delta y_{rp} - b_r}, \forall r \in J'_p, \quad (31)$$

respectively, where  $J'_p = \{r | y_{rp} + \Delta y_{rp} - b_r > 0, r = 1, \dots, s\}$ . We have from (29) and (24) that

$$\frac{x_{ip} - \Delta x_{ip} - \sum_{j=1, j \neq p}^n \lambda_j x_{ij}}{x_{ip} - \Delta x_{ip} + a_i} \leq \frac{x_{ip} - \Delta x_{ip} + \max_{j=1, \dots, n, j \neq p} \{|x_{ij}|\}}{x_{ip} - \Delta x_{ip} + a_i} < 1. \quad (32)$$

Therefore, we have  $\beta'_p < 1$ . Since  $k \geq 3$ , we know from (24) that

$$\begin{aligned} \max_{\substack{j=1, \dots, n, \\ j \neq p}} \{|x_{ij}|\} &\leq \max\{|x_{ij}|, \forall j, |x_{ip} - \Delta x_{ip}|\} \\ &\leq 2(x_{ip} - \Delta x_{ip}) + a_i. \end{aligned}$$

As  $x_{ip} - \Delta x_{ip} + a_i > 0$ ,  $i = 1, \dots, m$ , we have

$$\begin{aligned} \frac{x_{ip} - \Delta x_{ip} - \sum_{j=1, j \neq p}^n \lambda_j x_{ij}}{x_{ip} - \Delta x_{ip} + a_i} &\geq \frac{x_{ip} - \Delta x_{ip} - \max_{j=1, \dots, n, j \neq p} \{|x_{ij}|\}}{x_{ip} - \Delta x_{ip} + a_i} \geq \\ \frac{x_{ip} - \Delta x_{ip} - 2(x_{ip} - \Delta x_{ip}) - a_i}{x_{ip} - \Delta x_{ip} + a_i} &= -1, i = 1, \dots, m. \end{aligned} \quad (33)$$

On the other hand, we have from (31) that

$$\begin{aligned} \frac{\sum_{j=1, j \neq p}^n \lambda_j y_{rj} - y_{rp} - \Delta y_{rp}}{y_{rp} + \Delta y_{rp} - b_r} &\geq \frac{\min_{j=1, \dots, n} \{y_{rp}\} - y_{rp} - \Delta y_{rp}}{y_{rp} + \Delta y_{rp} - b_r} \\ &= -1, \forall r \in J'_p. \end{aligned} \quad (34)$$

From (30), (31), (33) and (34), we have  $\beta'_p \geq -1$  for  $p = 1, \dots, n$ , since  $\beta_p$  is maximized in model (26)–(29). Then,  $1 - \beta'_p$  and  $1 + \beta'_p$  are non-negative, and we have

$$\begin{aligned} \sum_{j=1, j \neq p}^n \lambda'_j x_{ij} &\leq (1 - \beta'_p)(x_{ip} - \Delta x_{ip}) - a_i \beta'_p \\ &\leq (1 - \beta'_p)x_{ip} - a_i \beta'_p, i = 1, \dots, m. \end{aligned} \quad (35)$$

$$\begin{aligned} \sum_{j=1, j \neq p}^n \lambda'_j y_{rj} &\geq (1 + \beta'_p)(y_{rp} + \Delta y_{rp}) - b_r \beta'_p \\ &\geq (1 + \beta'_p)y_{rp} - b_r \beta'_p, r = 1, \dots, s. \end{aligned} \quad (36)$$

As chosen constants, the values of  $a_i$  and  $b_r$  in model (26)–(29) should be the same as those in model (9)–(12). Thus according to (35) and (36),  $(\lambda'_j, \beta'_p)$  is a feasible solution vector for the model (9)–(12). Therefore,  $\beta_p^* \geq \beta'_p$  holds since we maximize  $\beta_p$  in model (9)–(12).  $\square$

(ii) Units invariance

Assume that the  $i_0$ th ( $i_0 \in \{1, \dots, m\}$ ) input and the  $r_0$ th ( $r_0 \in \{1, \dots, s\}$ ) output of all DMUs are scaled by the factors  $c$ ,  $c > 0$  and  $c \neq 1$ , and  $d$ ,  $d > 0$  and  $d \neq 1$ , respectively. After this scaling,  $a_{i_0}$  and  $b_{r_0}$  should be determined by  $k \cdot \max_{j=1, \dots, n} \{c \cdot x_{i_0j}\} = c \cdot k \cdot \max_{j=1, \dots, n} \{x_{i_0j}\}$  and  $\min_{j=1, \dots, n} \{d \cdot y_{r_0j}\} = d \cdot \min_{j=1, \dots, n} \{y_{r_0j}\}$ , respectively, according to (7) and (8). Thus, the corresponding constraints in (10) and (11) become

$$\begin{aligned} \sum_{j=1, j \neq p}^n \lambda_j \cdot c \cdot x_{i_0j} &\leq (1 - \beta_p) \cdot c \cdot x_{i_0p} - c \cdot k \\ &\cdot \max_{j=1, \dots, n} \{x_{i_0j}\} \cdot \beta_p, \quad i = 1, \dots, m, \\ \sum_{j=1, j \neq p}^n \lambda_j \cdot d \cdot y_{r_0j} &\geq (1 + \beta_p) \cdot d \cdot y_{r_0p} - d \\ &\cdot \min_{j=1, \dots, n} \{y_{r_0j}\} \cdot \beta_p, \quad r = 1, \dots, s. \end{aligned}$$

These constraints are equivalent to the original constraints in (10) and (11), respectively. Therefore, the optimal solution of our model does not change after the above scaling. This means that our model is unit invariant.

(iii) Translation invariance

• *Output translation* Assume that all the outputs of DMUs are translated to  $y'_{rj} = y_{rj} + \tau_r$ ,  $r = 1, \dots, s$ ,  $j = 1, \dots, n$ , where  $\tau_r$ ,  $r = 1, \dots, s$ , are arbitrary constants. After this translation,  $b_r$  should be calculated by

$$\begin{aligned} b_r &= \min_{j=1, \dots, n} \{y'_{rj}\} = \min_{j=1, \dots, n} \{y_{rj} + \tau_r\} = \min_{j=1, \dots, n} \{y_{rj}\} \\ &+ \tau_r, \quad r = 1, \dots, s, \end{aligned} \tag{37}$$

according to (8). Therefore, the constraints in (11) become

$$\begin{aligned} \sum_{j=1, j \neq p}^n \lambda_j (y_{rj} + \tau_r) &\geq y_{rp} + \tau_r + \beta_p (y_{rp} + \tau_r \\ &- \min_{j=1, \dots, n} \{y_{rj}\} - \tau_r), \quad r = 1, \dots, s, \end{aligned}$$

which is equivalent to (11) due to (12). Hence, model (9)–(12) is output translation invariant.

• *Input translation* If all the inputs of DMUs are translated to  $x'_{ij} = x_{ij} + \eta_i$ ,  $i = 1, \dots, m$ ,  $j = 1, \dots, n$ , where  $\eta_i$ ,  $i = 1, \dots, m$ , are arbitrary constants. After this translation,  $a_i$  should be calculated by

$$a_i = k \cdot \max_{j=1, \dots, n} \{x'_{ij}\} = k \cdot \max_{j=1, \dots, n} \{x_{ij} + \eta_i\}, \quad i = 1, \dots, m. \tag{38}$$

With (38), the constraints in (10) become

$$\begin{aligned} \sum_{j=1, j \neq p}^n \lambda_j (x_{ij} + \eta_i) &\leq x_{ip} + \eta_i - \beta_p (x_{ip} + \eta_i \\ &+ k \cdot \max_{j=1, \dots, n} \{x_{ij} + \eta_i\}), \quad i = 1, \dots, m. \end{aligned}$$

These constraints are not equivalent to the original constraints in (10).

To ensure the input translation invariance, we should translate  $a_i$  by the opposite amount of the corresponding input translation, ie,

$$a_i = k \cdot \max_{j=1, \dots, n} \{x_{ij}\} - \eta_i, \quad i = 1, \dots, m. \tag{39}$$

With (39), the constraints in (10) become

$$\begin{aligned} \sum_{j=1, j \neq p}^n \lambda_j (x_{ij} + \eta_i) &\leq x_{ip} + \eta_i - \beta_p (x_{ip} + \eta_i \\ &+ k \cdot \max_{j=1, \dots, n} \{x_{ij}\} - \eta_i), \quad i = 1, \dots, m, \end{aligned}$$

which are the same as those original constraints in (10) due to (12). Because of the adjustment in (39), we call the above input translation invariance with parameter adjustment the *generalized input translation invariance* for distinction. Although our model does not satisfy the traditional input translation invariance, it is generalized input translation invariant.

### 4. Numerical examples

In this section, two numerical examples are used to show the applicability and merits of the proposed model.

**Example 1** To show the properties of the proposed model and meanwhile to compare it with the NL model (Ray, 2008), the LC model (Lin and Chen, 2015), the super RDM+ model and the super RDM− model (Hadi-Vencheh and Esmailzadeh, 2013), we apply the proposed model to the data set in Table 1, situated in Section 2 above. In this paper, we set  $k = 3$ . Then according to (7) and (8), we have  $a_1 = 27$ ,  $a_2 = 3$ ,  $a_3 = 6$ ,  $b_1 = -8$  and  $b_2 = 0$  for the data set in Table 1. The proposed model is feasible for each of six DMUs. The resulting optimal value is shown in the second column of Table 2. Note that for DMU 4, some resulting slacks are not equal to zero. Thus, except DMU 4, all other DMUs are efficient. According to the super-efficiency scores, these six DMUs are ranked as column 4 of Table 2 shows. The last five columns of Table 2 show the inputs and outputs of the projection of each DMU on the super-efficiency frontier. For this data set (as well as the data set in Example 2), our model yields a unique projection for each DMU. It is easy to see from the projection results that: Inefficient DMU 4

**Table 2** Results for six DMUs

DMUp	$\beta_p^*$	Super-efficiency score	Rank	Projection on super-efficiency frontier				
				X1	X2	X3	Y1	Y2
1	-0.036	1.036	4	1.787	0.107	-0.213	2.680	0.320
2	-1.000	2.000	1	2.000	0.000	0.000	3.000	0.000
3	-0.006	1.006	5	4.175	0.000	0.000	3.932	0.000
4	0.000	1.000	6	2.000	0.000	0.000	3.000	0.000
5	-0.043	1.043	3	3.234	0.128	-0.255	2.532	0.383
6	-0.143	1.143	2	4.000	0.000	0.000	4.000	0.000

should generate lower target inputs and higher target outputs than the original values in order to reach the super-efficiency frontier; other DMUs should reduce (expand) at least one of its outputs (inputs) to reach the super-efficiency frontier.

To examine the monotonicity of the proposed model, we assume that the first input of DMU 1 is decreased from 1 to -99 and its first output is increased from -8 to 192 in the following way:  $x_{1,1} = 1 - l$ , and  $y_{1,1} = -8 + 2 \cdot l$ , here  $l$  increases from 0 to 100 with the step size being 1. As the values of  $x_{1,1}$  and  $y_{1,1}$  are changed in this situation, then according to (24) and (25), we have

$$\begin{aligned}
 a_1 &= 3 \cdot \max\{|1 - l|, 0, 4, 4, 2, 9, l = 0, \dots, 100\} = 297, \\
 a_2 &= 3 \cdot \max\{0, 1, 0, 0, 0, 0\} = 3, \\
 a_3 &= 3 \cdot \max\{0, |-2|, 0, 0, 0, 0\} = 6, \\
 b_1 &= \min\{-8, 0, 4, 1, 3, 6\} = -8, \\
 b_2 &= \min\{0, 3, 0, 0, 0, 0\} = 0,
 \end{aligned}$$

for the changing data set. When  $l$  increases from 0 to 100, the proposed model is always feasible and the optimal value  $\beta_1^*$  decreases gradually from -0.00329 to -0.93000. Figure 1 plots the change of  $\beta_1^*$  with respect to  $l$ . It is easy to see from Figure 1 that  $\beta_1^*$  monotonically decreases with the increase of  $l$ . This confirms the theoretical analysis about the monotonicity of model (9)–(12).

For the data set in Table 1, we assume that all the inputs are scaled down by 0.001 and two outputs are scaled down by 0.01. Then according to (7) and (8), we have  $a_1 = 0.027, a_2 = 0.003, a_3 = 0.006, b_1 = -0.08$  and  $b_2 = 0$  for the scaled data. By solving problem (9)–(12) with the scaled data, we find that all the resulting super-efficiency scores are still equal to those in column 3 of Table 2. This confirms the theoretical result about the unit invariance of the proposed model.

Assume the input and output values of all DMUs are translated to

$$\begin{aligned}
 x'_{1j} &= x_{1j}; \quad x'_{2j} = x_{2j}; \quad x'_{3j} = x_{3j} + 4; \\
 y'_{1j} &= y_{1j} + 9, \quad y'_{2j} = y_{2j}, \quad j = 1, \dots, 6.
 \end{aligned} \tag{40}$$

Then according to (39) and (37), we have

$$a_1 = 27, a_2 = 3, a_3 = 6 - 4 = 2, b_1 = -8 + 9 = 1, b_2 = 0. \tag{41}$$

With  $a_i$  and  $b_r$  in (41), we solve problem (9)–(12) under the translated data and find that all the resulting super-efficiency scores are still equal to those in column 3 of Table 2. This confirms the usual output translation invariance and the generalized input translation invariance of the proposed model.

It is worth noting that by the translation in (40), the input-output values of every DMU become non-negative. The first input is zero for DMU 2, and it is positive for other DMUs. By solving the NL model under these translated data, we find that it is infeasible for DMU 2 and the resulting super-efficiency score for DMU 1 is 2.5, which results in a negative referenced output for DMU 1. Therefore, the infeasibility issue in two exceptions of the NL model, mentioned in Section 2, occurs under the translated data. In contrast, the proposed model is feasible under the translated data for all DMUs, and we have from (23) that the outputs of referenced points for individual DMUs are  $\tilde{y}'_{11} = 11.680, \tilde{y}'_{12} = 10.183, \tilde{y}'_{13} = 12.932, \tilde{y}'_{14} = 10.903, \tilde{y}'_{15} = 11.532, \tilde{y}'_{16} = 13.000, \tilde{y}'_{21} = 0.320, \tilde{y}'_{22} = 0.000, \tilde{y}'_{23} = 0.000, \tilde{y}'_{24} = 0.000, \tilde{y}'_{25} = 0.383, \tilde{y}'_{26} = 0.000$ , which are non-negative. This confirms the conclusion of Corollary 2. From Proposition 1, Corollary 2 and the above analysis, we can safely say that for non-negative data, the proposed model is feasible and ensures the non-negativity of the referenced outputs for all DMUs. So, for the non-negative data set, the infeasibility issue of the NL model (Ray, 2008) does not occur for model (9)–(12).

**Example 2** In order to further show the practicability of our model, we consider the data set in Sharp *et al* (2007). This data set has 13 DMUs with two inputs and three outputs. The detailed data are shown in columns 2–6 of Table 3. It is easy to see from Table 3 that just one input (cost) and one output (saleable) are non-negative and other data are non-positive. For this example, we have  $a_1 = 32.4, a_2 = 6.96, b_1 = 0.49, b_2 = -1.42$  and  $b_3 = -3.79$  according to (7) and (8). By solving the proposed model for all DMUs, we obtain the resulting super-efficiency scores, which are shown in column 7 of Table 3.



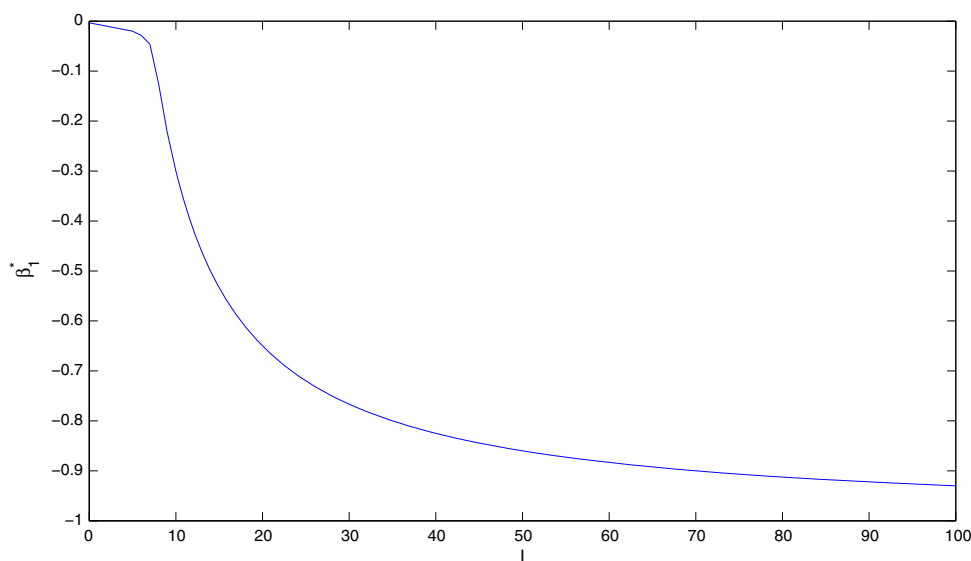


Figure 1 The change of the optimal value of model (9)–(12) for DMU 1.

Table 3 Input–output data and super-efficiency results

DMUp	Input		Output			Super-efficiency score	Rank
	Cost	Effluent	Saleable	CO <sub>2</sub>	Methane		
1	1.03	−0.05	0.56	−0.09	−0.44	0.998	7
2	1.75	−0.17	0.74	−0.24	−0.31	0.986	10
3	1.44	−0.56	1.37	−0.35	−0.21	1.041	3
4	10.80	−0.22	5.61	−0.98	−3.79	0.919	13
5	1.30	−0.07	0.49	−1.08	−0.34	0.995	8
6	1.98	−0.10	1.61	−0.44	−0.34	0.992	9
7	0.97	−0.17	0.82	−0.08	−0.43	1.011	5
8	9.82	−2.32	5.61	−1.42	−1.94	1.402	2
9	1.59	0.00	0.52	0.00	−0.37	1.000	6
10	5.96	−0.15	2.14	−0.52	−0.18	0.983	11
11	1.29	−0.11	0.57	0.00	−0.24	1.029	4
12	2.38	−0.25	0.57	−0.67	−0.43	0.969	12
13	10.30	−0.16	9.56	−0.58	0.00	1.540	1

It is easy to see that the proposed model is feasible for all DMUs. Compared with DEA models handling negative data (Portela *et al*, 2004; Sharp *et al*, 2007; Emrouznejad *et al*, 2010; Cheng *et al*, 2013), an advantage of the proposed model is that it can differentiate the performance of efficient DMUs. This superiority comes from an inherent characteristic of super-efficiency DEA models. According to the resulting super-efficiency scores, all DMUs are ranked as column 8 of Table 3 shows.

Table 4 shows the target input–output values of inefficient DMUs, determined by the proposed model. It is obvious that under the proposed model, each inefficient DMU should reduce its inputs and expand its outputs in order to tend to the super-efficiency frontier. Therefore, the proposed model can

provide improved target inputs and outputs for all inefficient DMUs.

From the theoretical analyses and the above two examples, we can conclude that the proposed model can deal with the data set with negative values and can provide improved targets for inefficient DMUs. In addition, our model is unit invariant, output translation invariant, generalized input translation invariant and monotonous. For the non-negative data set, the proposed model fully eliminates the infeasibility issue of the NL model. Therefore, the proposed model successfully addresses the infeasibility problem occurring in conventional VRS radial super-efficiency DEA models and the NL model. More importantly, different from current DEA models handling negative data, the proposed model can rank efficient DMUs.

**Table 4** Improved targets for inefficient DMUs

DMUp	Input 1		Input 2		Output 1		Output 2		Output 3	
	Original	Target	Original	Target	Original	Target	Original	Target	Original	Target
1	1.03	0.970	-0.05	-0.170	0.56	0.820	-0.09	-0.080	-0.44	-0.430
2	1.75	1.284	-0.17	-0.263	0.74	0.866	-0.24	-0.124	-0.31	-0.262
4	10.80	7.308	-0.22	-0.765	5.61	6.024	-0.98	-0.690	-3.79	-0.495
5	1.30	1.148	-0.07	-0.137	0.49	0.681	-1.08	-0.036	-0.34	-0.324
6	1.98	1.709	-0.10	-0.548	1.61	1.619	-0.44	-0.357	-0.34	-0.204
9	1.59	1.290	0.00	-0.110	0.52	0.570	0.00	0.000	-0.37	-0.240
10	5.96	5.305	-0.15	-0.385	2.14	4.943	-0.52	-0.450	-0.18	-0.118
12	2.38	1.314	-0.25	-0.456	0.57	1.223	-0.67	-0.278	-0.43	-0.269

## 5. Conclusions

Super-efficiency model in the presence of negative data is a rather neglected issue in the DEA field. The existing super-efficiency models capable of handling negative data might be infeasible in some cases. By choosing appropriate direction variables in the DDF, this paper develops a DDF-based VRS radial super-efficiency DEA model for dealing with negative data. Compared with existing related models, the proposed model is feasible no matter whether the input–output data are non-negative or not, and meanwhile, it can rank all DMUs. It can project each DMU onto the super-efficiency frontier along a suitable direction and can provide improved targets for inefficient DMUs. It possesses good properties such as monotonicity, units invariance, output translation invariance and generalized input translation invariance. Moreover, it successfully eliminates the infeasibility issue occurring in two exceptions of the NL model. In summary, the proposed model not only overcomes the infeasibility issue of VRS super-efficiency DEA models, but also extends current VRS radial super-efficiency DEA models suitable for non-negative data to the situation with partially or fully negative data.

The new super-efficiency DEA model is developed by utilizing the directional distance function, and the resulting super-efficiency scores are not equal to those yielded by traditional radial super-efficiency DEA models if the latter models are feasible and the input–output data are non-negative. As for the future research, we will propose an alternative radial super-efficiency DEA model which not only overcomes the above drawback, but also keeps all advantages of the proposed super-efficiency DEA model.

*Acknowledgments*—This work is supported by the National Natural Science Foundation of China (Grant Nos. 11301395, 11201344, 11571270 and 71371152) and the Natural Science Foundation of Zhejiang province, China (Grant No. LY17G010004). The authors are grateful to two anonymous reviewers and the editor for their detailed and constructive comments, which have helped us to improve the paper significantly.

## References

Ali AI and Seiford LM (1990). Translation-invariance in data envelopment analysis. *Operations Research Letters* **9**(6):403–405.

- Andersen P and Petersen NC (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science* **39**(10):1261–1264.
- Banker RD, Charnes A and Cooper WW (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* **30**(9):1078–1092.
- Banker RD, Das S and Datar SM (1989). Analysis of cost variances for management control in hospitals. *Research in Governmental and Nonprofit Accounting* **5**:268–291.
- Briec W and Kerstens K (2009a). Infeasibilities and directional distance functions with application to the determinateness of the Luenberger productivity indicator. *Journal of Optimization Theory and Applications* **141**(1):55–73.
- Briec W and Kerstens K (2009b). The Luenberger productivity indicator: An economic specification leading to infeasibilities. *Economic Modelling* **26**(3):597–600.
- Chambers RG, Chung Y and Färe R (1996). Benefit and distance functions. *Journal of Economic Theory* **70**(2):407–419.
- Charnes A, Cooper WW and Rhodes E (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research* **2**(6):429–444.
- Charnes A, Clark T, Cooper WW and Golany B (1984). A developmental study of data envelopment analysis in measuring the efficiency of maintenance units in US. Air forces. *Annals of Operational Research* **2**(1):95–112.
- Chen Y (2005). Measuring super-efficiency in DEA in the presence of infeasibility. *European Journal of Operational Research* **161**(2):545–551.
- Chen Y, Du J and Huo J (2013). Super-efficiency based on a modified directional distance function. *OMEGA* **41**(3):621–625.
- Cheng G, Zervopoulos P and Qian Z (2013). A variant of radial measure capable of dealing with negative inputs and outputs in data envelopment analysis. *European Journal of Operational Research* **225**(1):100–105.
- Cook WD, Liang L, Zha Y and Zhu J (2009). A modified super-efficiency DEA model for infeasibility. *Journal of the Operational Research Society* **60**(2):276–281.
- Cooper WW, Seiford LM and Tone K (2007). *Data Envelopment Analysis* (2nd edition). Springer, New York.
- Emrouznejad A, Anouze AL and Thanassoulis E (2010). A semi-oriented radial measure for measuring the efficiency of decision making units with negative data, using DEA. *European Journal of Operational Research* **200**(1):297–304.
- Hadi-Vencheh A and Esmaeilzadeh A (2013). A new super-efficiency model in the presence of negative data. *Journal of the Operational Research Society* **64**(3):396–401.
- Kerstens K and Van de Woestyne I (2011). Negative data in DEA: a simple proportional distance function approach. *Journal of the Operational Research Society* **62**(7):1413–1419.
- Kerstens K and Van de Woestyne I (2014). A note on a variant of radial measure capable of dealing with negative inputs and outputs

- in DEA. *European Journal of Operational Research* **234**(1):341–342.
- Lee HS, Chu CW and Zhu J (2011). Super-efficiency DEA in the presence of infeasibility. *European Journal of Operational Research* **212**(1):141–147.
- Lin R and Chen Z (2015). Super-efficiency measurement under variable return to scale: an approach based on a new directional distance function. *Journal of the Operational Research Society* **66**(9):1506–1510.
- Lovell CAK and Pastor JT (1995). Units invariant and translation invariant DEA models. *Operations Research Letters* **18**(3):147–151.
- Pastor JT (1996). Translation invariance in data envelopment analysis: a generalization. *Annals of Operations Research* **66**(2):91–102.
- Portela MCAS, Thanassoulis E and Simpson G (2004). Negative data in DEA: a directional distance approach applied to bank branches. *Journal of the Operational Research Society* **55**(10):1111–1121.
- Ray SC (2008). The directional distance function and measurement of super-efficiency: an application to airlines data. *Journal of the Operational Research Society* **59**(6):788–797.
- Scheel H (2001). Undesirable outputs in efficiency valuations. *European Journal of Operational Research* **132**(2):400–410.
- Seiford LM and Zhu J (1999). Infeasibility of super-efficiency data envelopment analysis models. *INFOR* **37**(2):174–187.
- Sharp JA, Meng W and Liu W (2007). A modified slacks-based measure model for data envelopment analysis with ‘natural’ negative outputs and inputs. *Journal of the Operational Research Society* **58**(12):1672–1677.

*Received 28 July 2015;  
accepted 4 October 2016*