2021-09-24

A Range Adjusted Measure of Super-efficiency in Integer-valued Data Envelopment Analysis with Undesirable Outputs

Chunhua Chen

Haohua Liu

Lijun Tang *Plymouth Business School*

Jianwei Ren

Let us know how access to this document benefits you

Recommended Citation

A Range Adjusted Measure of Super-efficiency in Integer-valued Data Envelopment Analysis with Undesirable Outputs

Chunhua CHEN
School of Business Administration, Jiangxi University of Finance and Economics, Nanchang 330013, China
Inner Mongolia Branch of Agricultural Bank of China, Hohhot 010010, China
Email: nmgchenchunhua@126.com

Haohua LIU
School of Business Administration, Jiangxi University of Finance and Economics, Nanchang 330013, China
Email: jxliuhaohua@126.com

Lijun TANG
Plymouth Business School, University of Plymouth, Plymouth PL4 8AA, United Kingdom
Email: lijun.tang@plymouth.ac.uk

Jianwei REN *
Transportation Institute, Inner Mongolia University, Hohhot 010070, China
School of Mathematical Sciences, Inner Mongolia University, Hohhot 010010, China
Email: Jianwei Ren (renjianwei309@126.com)

Abstract DEA (Data Envelopment Analysis) models can be divided into two groups: radial DEA and non-radial DEA, and the latter has higher discriminatory power than the former. The Range Adjusted Measure (RAM) is an effective and widely used non-radial DEA approach. However, to the best of our knowledge, there is no literature on the integer-valued super-efficiency RAM-DEA model, especially when undesirable outputs are included. We first propose an integer-valued RAM-DEA model with undesirable outputs and then extend this model to an integer-valued super-efficiency RAM-DEA model with undesirable outputs. Compared with other DEA models, the two novel models have many advantages: (1) they are non-oriented and non-radial DEA models, which enable decision makers to simultaneously and non-proportionally improve inputs and outputs; (2) they can handle integer-valued variables and undesirable outputs, so the results obtained are more reliable; (3) the results can be easily obtained as it is based on linear programming; (4) the integer-valued super-efficiency RAM-DEA model with undesirable outputs can be used to accurately rank efficient DMUs. The proposed models are applied to evaluate the efficiency of China's regional transportation systems (RTSs) considering the number of transport accidents (an undesirable output). The results help decision makers improve the performance of inefficient RTSs and analyze the strengths of efficient RTSs.

Keywords DEA, Range Adjusted Measure (RAM), super-efficiency, undesirable outputs, transport.

1. Introduction

Data envelopment analysis (DEA) is generally regarded as an effective nonparametric technique to evaluate the relative efficiency (performance) of decision making units (DMUs) [1,2]. DEA models can be divided into two groups: radial DEA and non-radial DEA [3, 4]. Radial DEA models, i.e., the CCR model by Charnes et al. [5] and the BCC model by Banke et al. [6], only allow proportional reductions of inputs or increases of outputs to improve the performance of DMUs, while non-radial DEA models enable non-proportional changes of inputs and outputs [7]. Thus, non-radial DEA models have higher discriminatory power than radial DEA models [8,9].

There are several types of non-radial DEA models including additive DEA [10], Slacks-Based Measure (SBM) [11], and Range Adjusted Measure (RAM) [12]. Additive DEA models, however, have a weakness: they cannot generate efficiency scores for DMUs [13]. To address this weakness, RAM-DEA and SBM-DEA are developed based on the additive DEA. While both of them have all the advantages of additive DEA, they are also able to generate efficiency scores. Therefore, they have been applied to various situations [14]. In this paper, we apply the RAM-DEA because integer-valued RAM-DEA models are linear programming while integer-valued SBM-DEA models are nonlinear programming even after the Charnes-Cooper transformation [15,16]. In the context that many inputs and outputs for example, the number of employees, machines, gold medals, transport accidents and so on,, can only be integer values in reality, integer-valued variables have to be taken into account. As real-valued DEA models cannot deal with integer-valued variables, integer-valued DEA models have been developed to address the issue [17-19]. Lozano and Villa were the first to propose a mixed integer linear

18 programming (MILP) DEA model to restrict variables to integers [17], but this model is not consistent with the
19 minimum extrapolation principle. To tackle this problem, Kuosmanen and Matin developed a novel axiomatic
20 foundation for integer-valued DEA models that assumes subsets of input and output variables to be integer-valued [19].
21 As far as we know, there is no literature on integer-valued RAM-DEA models. All the RAM-DEA models discussed
22 below are real-valued DEA models.
23   Cooper et al. presented the first RAM-DEA model and analyzed its advantages at the Georgia Productivity Workshop in
24 1996, and their paper was published in 1999 [12]. Aida et al. tested the robustness of the RAM-DEA model and applied this
25 model to estimate the performance of water supply services in Japan [20]. Steinmann and Zweifel [21] as well as Cooper et al.
26 [22] also studied the properties of the RAM-DEA model. Sueyoshi and Sekitani proposed the measurement of returns to scale
27 using a RAM-DEA model [23]. Based on the RAM-DEA, Cooper et al. developed a Bounded Adjusted Measure (BAM) DEA
28 model that considered lower bounds for inputs and upper bounds for outputs [24]. Chiu et al. proposed a context-dependent
29 RAM-DEA model to evaluate the relative attractiveness and progress efficiency of Taiwanese commercial banks [25].
30 Avkiran and McCrystal presented a dynamic network RAM-DEA model [26]. Tavassoli et al. developed a RAM-SCSC-
31 DEADA approach to rank Iranian airlines, where the SCSC indicated the strong complementary slackness condition and the
32 DEADA represented the DEA discriminant analysis [27]. Liu et al. proposed a closest RAM-DEA model to evaluate the
33 performance of coal-fired power plants in China [28]. Li et al. presented a virtual frontier dynamic RAM-DEA model to
34 measure the energy efficiency of airlines [29]. Wanke and Barros also applied a virtual frontier dynamic RAM-DEA model to
35 measure the performance of Latin American airlines [30]. Li et al. developed an input-shared network RAM-DEA model to
36 evaluate the efficiency of several airlines [31].
37   The aforementioned RAM-DEA models did not take undesirable outputs into account. However, in reality, many
38 production activities generate undesirable outputs that should be reduced as much as possible. Many researches have proved
39 that it is necessary to consider undesirable outputs when evaluating the efficiency of DMUs with undesirable outputs [32,33].
40 Sueyoshi and Goto first proposed a RAM-DEA model with undesirable outputs (RAM-UDEA) [34,35]. Ramli and Munisamy
41 applied a RAM-UDEA model to evaluate the eco-efficiency of the manufacturing industry in Malaysia [36]. Wang and Yu
42 developed a RAM-UDEA model to estimate the industrial energy and environmental performance of Chinese cities [37].
43 Wang et al. presented a RAM-UDEA model to measure the innovation and environment performance of China's
44 manufacturing industry [38]. Cui and Li proposed a network RAM-UDEA model with weak-G disposability to evaluate the
45 environmental performance of several airlines [39]. Meng et al. applied a RAM-UDEA model to estimate the low-carbon
46 economy performance of China's provinces [40]. Kang et al. presented a RAM-UDEA model with natural and managerial
47 disposability to measure the energy and environmental performance of the manufacturing industry in China [41]. Miao et al.
48 studied the joint decomposition of RAM-UDEA and Luenberger productivity indicators and applied this approach to
49 evaluate the atmospheric environmental performance in China [42]. Eftekharian et al. proposed a RAM-UDEA model to
50 measure the environmental efficiency of industrial industries in Iran [43]. Wang and Yuan proposed a RAM-UDEA model to
51 measure the energy and CO_2 (carbon dioxide) emission efficiencies of cigarette companies in China [44]. Chen et al. used the
52 RAM-UDEA approach to compare congestion effects of 46 countries along the Belt and Road [45]. Yuan et al. presented a
53 RAM-UDEA model to assess the inclusive and sustainable industrial development in China [46].
54   However, conventional radial and non-radial DEA models cannot be used to distinguish better performers from efficient
55 DMUs because the efficiency scores of all efficient DMUs must be equal to 1. To differentiate the efficient DMUs, it is
56 necessary to measure super-efficiency. Super-efficiency DEA models can be developed by removing the DMU under
57 evaluation from the reference set [47, 48] allowing their efficiency (super-efficiency) scores greater than one. In this way,
58 decision makers can rank the efficient DMUs based on their super-efficiency scores. Andersen and Petersen proposed the
59 first radial super-efficiency DEA model [49]. Radial super-efficiency DEA models can only proportionally change inputs and
60 outputs and therefore their discriminatory ability is weaker than non-radial super-efficiency DEA models [50,51]. In the field of
61 non-radial super-efficiency DEA, Du et al. [52], Guo et al. [53], Yu and Hsu [54], among others studied super-efficiency additive
62 DEA models, and Tone [55], Tran et al. [56], Chen et al. [57], and many other scholars researched super-efficiency SBM-DEA
63 models. Many studies show that super-efficiency additive DEA models are always feasible while conventional radial super-
64 efficiency DEA models would be infeasible under the assumption of VRS (Variable Returns to Scale) [52,54].
65   To the best of our knowledge, however, there is no literature on the super-efficiency RAM-DEA model, especially
66 when both integer-valued variables and undesirable outputs are included. Therefore, in this paper we propose an integer-
67 valued super-efficiency RAM-DEA model with undesirable outputs (Super-RAM-IUDEA) and apply the model to evaluate
68 the efficiency of China's regional transportation systems considering the number of transport accidents (an undesirable
69 output). The conceptual model of our research is shown in Figure 1. To propose the Super-RAM-IUDEA we first develop an
70 integer-valued RAM-DEA model with undesirable outputs (RAM-IUDEA). The main advantages of the two novel models
71 are: (1) they are non-oriented and non-radial DEA models, which enable decision makers to simultaneously and non-
72 proportionally improve inputs and outputs; (2) they can handle integer-valued variables and undesirable outputs, and as such

73 the results obtained from our model are more reliable; (3) the results can be easily obtained because the model is based on
74 linear programming; (4) the Super-RAM-IUDEA model can be used to accurately rank efficient DMUs.
75   The remaining of this paper is organized as follows. The RAM-IUDEA and Super-RAM-IUDEA models are presented
76 in Section 2. The proposed models are applied to evaluate the efficiency (super-efficiency) of China's regional transportation
77 systems in Section 3. A concluding comment and the limitation of this research are summarized in Section 4. We also
78 highlight the contributions of this paper in the last section.
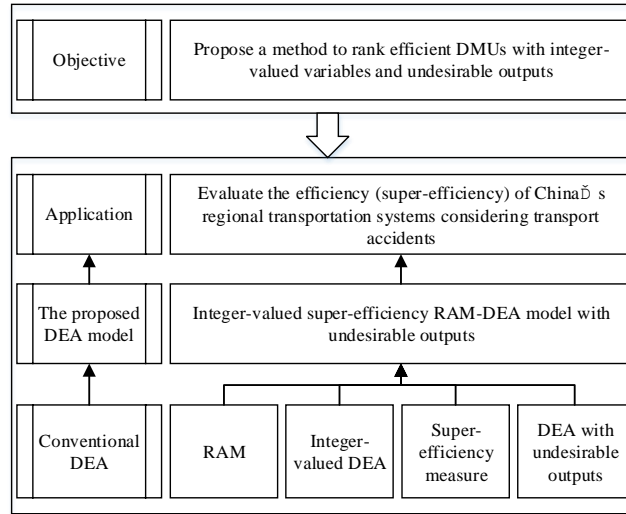
Figure 1. The conceptual model of this research.

79
80
81

2. Methodology

82
83   Consider that we have q DMUs (DMU_i, $i = 1, 2, ..., q$) and each DMU uses r inputs to produce m desirable outputs
84 as well as t undesirable outputs. In addition, some of these variables can only be integers. Let x_{ji} ($j = 1, 2, ..., g$),
85 x_{ji} ($j = g+1, g+2, ..., r$), y_{pi} ($p = 1, 2, ..., o$), y_{pi} ($p = o+1, o+2, ..., m$), z_{hi} ($h = 1, 2, ..., e$), and z_{hi} ($h = e+1, e+2, ..., t$)
86 denote DMU_i's integer-valued inputs, real-valued inputs, integer-valued desirable outputs, real-valued desirable outputs,
87 integer-valued undesirable outputs, and real-valued undesirable outputs, respectively.

2.1 Classic Additive DEA Model and Range Adjusted Measure

88
89   Charnes et al. first proposed model (1) to evaluate the performance of DMU_k (the DMU under evaluation) in 1985 [10].
90 The model was named "additive DEA model" because its objective is to maximize the sum of all slacks. This model
91 simultaneously deals with the slacks for inputs and desirable outputs, so it can identify all inefficiencies in inputs and desirable
92 outputs [10].

$$\max = s_j^- + s_p^+$$

$$s.t. \quad \sum_{i=1}^{q} x_{ji} \lambda_i = x_{jk} - s_j^-, \ j = 1, 2, ..., r$$

$$\sum_{i=1}^{q} y_{pi} \lambda_i = y_{pk} + s_p^+, \ p = 1, 2, ..., m$$

$$\sum_{i=1}^{q} \lambda_i = 1$$

$$\lambda_i, s_j^-, s_p^+ \geq 0 \ for \ all \ i, j, p \qquad (1)$$

93

94 where s_j^- indicates the slack related to the j th input; s_p^+ denotes the slack related to the p th desirable output; and λ_i is
95 the weight for DMU_i. Note that the original additive DEA model proposed by Charnes et al. is under the CRS (Constant
96 Returns to Scale) condition while model (1) assumes VRS (Variable Returns to Scale) [10]. The original CRS additive DEA

97 model can be obtained by deleting $\sum_{i=1}^{q} \lambda_i = 1$ from model (1). According to model (1), DMU_k is efficient if and only if all

98 slacks equal to 0 ($s_j^{-*} = s_p^{+*} = 0$). We let the superscript * indicate the optima obtained from DEA models. Obviously,

99 model (1) cannot generate efficiency scores for DMUs. In addition, it treats integer-valued variables as real-valued variables

100 and cannot deal with undesirable outputs. Therefore, this model is not good enough for our problem.

101 Cooper et al. developed the RAM-DEA model (model 2) based on the additive DEA model, which maximized the

102 weighted sum of all slacks [12]. Different from the additive DEA model, the RAM-DEA model can generate efficiency scores.

103 $\theta_k *$ is the efficiency score of DMU_k. Moreover, the RAM-DEA model has five properties that should be satisfied. (1)

104 $0 \le \theta_k * \le 1$, which means that the efficiency scores obtained from model (2) are between zero and one. (2) $\theta_k * = 1$

105 ($s_j^{-*} = s_p^{+*} = 0$) if and only if DMU_k is fully Pareto-Koopmans efficient. (3) $\theta_k *$ is invariant to the units of inputs and

106 outputs and invariant to alternative optima. (4) $\theta_k *$ is invariant to transformation ($\theta_k *$ is not affected by a change in the

107 origin of coordinate values). (5) $\theta_k *$ has strong monotonicity [20-22]. However, model (2) cannot handle integer-valued

108 variables or undesirable outputs. Therefore, we update this model and extend the updated RAM-DEA model to a super-

109 efficiency RAM-DEA model.

$$\min \theta_k = 1 - \frac{1}{r+m}\left(\sum_{j=1}^{r} \frac{s_j^-}{R_j^-} + \sum_{p=1}^{m} \frac{s_p^+}{R_p^+}\right)$$

$$s.t. \quad \sum_{i=1}^{q} x_{ji}\lambda_i = x_{jk} - s_j^-, \; j = 1,2,...,r$$

110

$$\sum_{i=1}^{q} y_{pi}\lambda_i = y_{pk} + s_p^+, \; p = 1,2,...,m$$

$$\sum_{i=1}^{q} \lambda_i = 1$$

$$\lambda_i, s_j^-, s_p^+ \ge 0 \; for \; all \; i, j, p \qquad (2)$$

111 where $R_j^- = \max_i \{x_{ji}\} - \min_i \{x_{ji}\}$ is the range of the j th input and $R_p^+ = \max_i \{y_{pi}\} - \min_i \{y_{pi}\}$ is the range of the P th

112 desirable output. The other mathematical notations are the same as those in model (1).

2.2 Integer-Valued RAM-DEA Model with Undesirable Outputs

114 Suppose $x_{ji} \in X^I$ ($j = 1,2,....,g$), $y_{pi} \in Y^I$ ($p = 1,2,....,o$), $z_{hi} \in Z^I$ ($h = 1,2,...,e$), $x_{ji} \in X^{NI}$ ($j = g+1, g+2,....,r$),

115 $y_{pi} \in Y^{NI}$ ($p = o+1, o+2,....,m$), and $z_{hi} \in Z^{NI}$ ($h = e+1, e+2,...,t$), where $X^I \in \cancel{c}_+^g$, $Y^I \in \cancel{c}_+^o$, $Z^I \in \cancel{c}_+^e$ indicate the

116 subsets of corresponding integer-valued variables while X^{NI}, Y^{NI}, Z^{NI} denote the subsets of corresponding real-valued

117 variables. Then, the DEA production possibility set (PPS) assuming VRS for our problem can be represented by

118

$$T_{IUDEA} = \left\{ \begin{pmatrix} X^I, Y^I, Z^I \\ X^{NI}, Y^{NI}, Z^{NI} \end{pmatrix} \middle| \begin{array}{l} (X^I, Y^I, Z^I) \in \cancel{c}_+^{g+o+e}; \begin{pmatrix} X^I \\ X^{NI} \end{pmatrix} \ge \sum_{i=1}^{q} \begin{pmatrix} X_i^I \\ X_i^{NI} \end{pmatrix} \lambda_i; \\ \begin{pmatrix} Y^I \\ Y^{NI} \end{pmatrix} \le \sum_{i=1}^{q} \begin{pmatrix} Y_i^I \\ Y_i^{NI} \end{pmatrix} \lambda_i; \\ \begin{pmatrix} Z^I \\ Z^{NI} \end{pmatrix} \ge \sum_{i=1}^{q} \begin{pmatrix} Z_i^I \\ Z_i^{NI} \end{pmatrix} \lambda_i; \sum_{i=1}^{q} \lambda_i = 1; \lambda_i \ge 0; \forall i \end{array} \right\} \qquad (3)$$

119 The efficiency score of DMU_k relative to the T_{IUDEA} reference technology can be computed by solving the proposed

120 integer-valued RAM-DEA model with undesirable outputs (RAM-IUDEA, model 4). In the model, λ_i is the weight of

121 DMU_i; $R_j^- = \max_i \{x_{ji}\} - \min_i \{x_{ji}\}$ is the range of the j th input; $R_p^+ = \max_i \{y_{pi}\} - \min_i \{y_{pi}\}$ is the range of the P th

122 desirable output; $R_h^- = \max_i \{z_{hi}\} - \min_i \{z_{hi}\}$ is the range of the h th undesirable output; s_j^{I-}, s_j^-, s_p^{I+}, s_p^+, ς_h^{I-}, and

123 ς_h^- indicate the slacks related to integer-valued inputs, real-valued inputs, integer-valued desirable outputs, real-valued

124 desirable outputs, integer-valued undesirable outputs, and real-valued undesirable outputs, respectively; $\cancel{c}$ is the set of

125 integer values and therefore $s_j^{I-} *$, $s_p^{I+} *$, and $\varsigma_h^{I-} *$ must be integers.

4

Model (4) has all the advantages of the classic RAM-DEA model and is able to identify all inefficiencies in integer-valued and real-valued inputs, integer-valued and real-valued desirable outputs, and integer-valued and real-valued undesirable outputs. Note that we use the slacks-based measure (SBM) approach ($\sum_{i=1}^{q} z_{hi}\lambda_i \leq z_{hk} - \varsigma_h^{I-}, h=1,2,...,e; \sum_{i=1}^{q} z_{hi}\lambda_i = z_{hk} - \varsigma_h^{-}, h=e+1,e+2,...,t$, where ς_h^{I-} and ς_h^{-} are the slacks related to integer-valued undesirable outputs and real-valued undesirable outputs, respectively.) and formulate RAM-DEA models to handle undesirable outputs [58]. There are several approaches to deal with undesirable outputs, such as the assumption of weak disposability proposed by Färe et al. [59], direction distance function (DDF) suggested by Chung et al. [60,61], linear or non-linear monotonic decreasing transformation [62,63], and treating undesirable outputs as inputs [13]. Compared with these approaches, the SBM approach is simple and easy to understand [58]. The constraints related to integer-valued variables are inequalities because $\sum_{i=1}^{q} x_{ji}\lambda_i *$, $\sum_{i=1}^{q} y_{pi}\lambda_i *$, and $\sum_{i=1}^{q} z_{hi}\lambda_i *$ may not be integers (where $0 \leq \lambda_i * \leq 1$) while $x_{jk} - s_j^{I-}*$, $y_{pk} + s_p^{I+}*$, and $z_{hk} - \varsigma_h^{I-}*$ must be integers. This method can effectively deal with integer-valued variables when measuring efficiency, so it has been widely used [52,64].

$$\min \theta_k = 1 - \frac{\sum_{j=1}^{g} \frac{s_j^{I-}}{R_j^-} + \sum_{j=g+1}^{r} \frac{s_j^-}{R_j^-} + \sum_{p=1}^{o} \frac{s_p^{I+}}{R_p^+} + \sum_{p=o+1}^{m} \frac{s_p^+}{R_p^+} + \sum_{h=1}^{e} \frac{\varsigma_h^{I-}}{R_h^-} + \sum_{h=e+1}^{t} \frac{\varsigma_h^-}{R_h^-}}{r+m+t}$$

$$s.t. \quad \sum_{i=1}^{q} x_{ji}\lambda_i \leq x_{jk} - s_j^{I-}, \ j=1,2,...,g$$

$$\sum_{i=1}^{q} x_{ji}\lambda_i = x_{jk} - s_j^-, \ j=g+1,g+2,...,r$$

$$\sum_{i=1}^{q} y_{pi}\lambda_i \geq y_{pk} + s_p^{I+}, \ p=1,2,...,o$$

$$\sum_{i=1}^{q} y_{pi}\lambda_i = y_{pk} + s_p^+, \ p=o+1,o+2...,m$$

$$\sum_{i=1}^{q} z_{hi}\lambda_i \leq z_{hk} - \varsigma_h^{I-}, \ h=1,2,...,e$$

$$\sum_{i=1}^{q} z_{hi}\lambda_i = z_{hk} - \varsigma_h^-, \ h=e+1,e+2,...,t$$

$$\sum_{i=1}^{q} \lambda_i = 1$$

$$\lambda_i, s_j^{I-}, s_j^-, s_p^{I+}, s_p^+, \varsigma_h^{I-}, \varsigma_h^- \geq 0 \ for \ all \ i,j,p,h$$

$$s_j^{I-} \in \phi, \ j=1,2,...,g$$

$$s_p^{I+} \in \phi, \ p=1,2,...,o$$

$$\varsigma_h^{I-} \in \phi, \ h=1,2,...,e \tag{4}$$

Model (4) is the first RAM-DEA model that can deal with undesirable outputs and integer-valued variables. The main advantages of the novel model are: (1) it is a non-oriented and non-radial DEA model, which enable decision makers to simultaneously and non-proportionally improve inputs and outputs; (2) it can handle integer-valued variables and undesirable outputs, and as such the results obtained from our model are more reliable; (3) the results can be easily obtained because the model is based on linear programming.

Proposition 1. DMU_k is efficient if and only if $\theta_k *=1$ and all slacks equal to 0 ($s_j^{I-}*=s_j^-*=s_p^{I+}*=s_p^+*=\varsigma_h^{I-}*=\varsigma_h^-*=0$).

Proposition 2. DMU_k is inefficient if and only if $\theta_k *<1$.

Proposition 3. The greater $\theta_k *$ is, the better the performance of DMU_k would be.

Proof:

(1) If $s_j^{I-}*=s_j^-*=s_p^{I+}*=s_p^+*=\varsigma_h^{I-}*=\varsigma_h^-*=0$, then we have $\sum_{i=1}^{q} x_{ji}\lambda_i *+\varepsilon_j^{I-}=x_{jk}$ for $j=1,2,...,g$, $\sum_{i=1}^{q} x_{ji}\lambda_i *=x_{jk}$ for $j=g+1,g+2,...,r$, $\sum_{i=1}^{q} y_{pi}\lambda_i *-\varepsilon_p^{I+}=y_{pk}$ for $p=1,2,...,o$, $\sum_{i=1}^{q} y_{pi}\lambda_i *=y_{pk}$ for $p=o+1,o+2...,m$,

$\sum_{i=1}^{q} z_{hi}\lambda_i * + \varepsilon_h^{I-} = z_{hk}$ for $h=1,2,...,e$, and $\sum_{i=1}^{q} z_{hi}\lambda_i * = z_{hk}$ for $h=e+1,e+2,...,t$, which means that DMU_k is on the efficient frontier (the production possibility frontier). Therefore, DMU_k is efficient. ε_j^{I-} , ε_p^{I+} , ε_h^{I-} are used to make $\sum_{i=1}^{q} x_{ji}\lambda_i * + \varepsilon_j^{I-}$, $\sum_{i=1}^{q} y_{pi}\lambda_i - \varepsilon_p^{I+}$, $\sum_{i=1}^{q} z_{hi}\lambda_i * + \varepsilon_h^{I-}$ integers. This method makes the efficient frontier of DMUs with integer-valued variables more reliable [17-19]. Note that $\theta_k * = 1$ if and only if all slacks equal to 0. Therefore, DMU_k is efficient if $\theta_k * = 1$ and all slacks equal to 0.

$s_j^{I-} * \neq 0$, $s_j^- * \neq 0$, $s_p^{I+} * \neq 0$, $s_p^+ * \neq 0$, $\varsigma_h^{I-} * \neq 0$, and $\varsigma_h^- * \neq 0$, which imply $\theta_k * \neq 1$, indicate that DMU_k can move to the efficient frontier by reducing integer-valued inputs, cutting down on real-valued inputs, increasing integer-valued desirable outputs, increasing real-valued desirable outputs, reducing integer-valued undesirable outputs, and reducing real-valued undesirable outputs, respectively. Thus, DMU_k is inefficient. We have proved that DMU_k is efficient only if $\theta_k * = 1$ and all slacks equal to 0. Proposition 1 has been proved.

(2) There must be $\theta_k * \leq 1$ for model (4) because we have $0 \leq s_j^{I-} * \leq x_{jk} - \sum_{i=1}^{q} x_{ji}\lambda_i * \leq R_j^-$ for $j=1,2,...,g$,

$0 \leq s_j^- * = x_{jk} - \sum_{i=1}^{q} x_{ji}\lambda_i * \leq R_j^-$ for $j=g+1,g+2,...,r$, $0 \leq s_p^{I+} * \leq \sum_{i=1}^{q} y_{pi}\lambda_i * - y_{pk} \leq R_p^+$ for $p=1,2,...,o$,

$0 \leq s_p^+ * = \sum_{i=1}^{q} y_{pi}\lambda_i * - y_{pk} \leq R_p^+$ for $p=o+1,o+2...,m$, $0 \leq \varsigma_h^{I-} * \leq z_{hk} - \sum_{i=1}^{q} z_{hi}\lambda_i * \leq R_h^-$ for $h=1,2,...,e$, and

$0 \leq \varsigma_h^- * = z_{hk} - \sum_{i=1}^{q} z_{hi}\lambda_i * \leq R_h^-$ for $h=e+1,e+2,...,t$. Proposition 2 is valid because we also have Proposition 1.

(3) The greater $\theta_k *$ (the less the second part of the right side of the objective function in model 4) is, the shorter the distance between DMU_k and the efficient frontier would be, which indicates that there are less waste on inputs, less shortage in desirable outputs, and less undesirable outputs in DMU_k . Therefore, the greater $\theta_k *$ is, the better the performance of DMU_k would be.

To achieve efficiency, decision makers can set targets to eliminate slacks. The targets for DMU_k 's integer-valued inputs, real-valued inputs, integer-valued desirable outputs, real-valued desirable outputs, integer-valued undesirable outputs, and real-valued undesirable outputs can be calculated using equations (5), (6), (7), (8), (9), and (10), respectively. In this approach, integer-valued targets for integer-valued variables can be obtained because the slacks for integer-valued variables are integers ($s_j^{I-} * \in \mathcal{C}$, $s_p^{I+} * \in \mathcal{C}$, $\varsigma_h^{I-} * \in \mathcal{C}$).

$$\overline{x}_{jk} = x_{jk} - s_j^{I-} *, j=1,2,...,g \qquad (5)$$

$$\overline{x}_{jk} = x_{jk} - s_j^- *, j=g+1,g+2,...,r \qquad (6)$$

$$\overline{y}_{pk} = y_{pk} + s_p^{I+} *, p=1,2,...,o \qquad (7)$$

$$\overline{y}_{pk} = y_{pk} + s_p^+ *, p=o+1,o+2...,m \qquad (8)$$

$$\overline{z}_{hk} = z_{hk} - \varsigma_h^{I-} *, h=1,2,...,e \qquad (9)$$

$$\overline{z}_{hk} = z_{hk} - \varsigma_h^- *, h=e+1,e+2,...,t \qquad (10)$$

2.3 Integer-Valued Super-Efficiency RAM-DEA Model with Undesirable Outputs

However, as the efficiency scores of all efficient DMUs must equal to one, it is impossible to differentiate efficient DMUs using model (4). To develop an integer-valued super-efficiency RAM-DEA model with undesirable outputs (Super-RAM-IUDEA), we cannot just delete the evaluated DMU from the reference set because the model may be infeasible for efficient DMUs if we do so. In this context, we develop model (11) to calculate the super-efficiency score $\varphi_k *$ of efficient DMU_k . In the model, ξ_j^{I-} , ξ_j^- , ξ_p^{I+} , ξ_p^+ , τ_h^{I-} , and τ_h^- denote DMU_k 's integer-valued inputs savings, real-valued inputs savings, integer-valued desirable outputs surpluses, real-valued desirable outputs surpluses, integer-valued undesirable outputs savings, and real-valued undesirable outputs savings, respectively. The other mathematical notations are the same as those in model (4). Model (11) has all the advantages of model (4) and can be used to accurately rank efficient DMUs.

6

$$\min \varphi_k = 1 + \frac{\sum\limits_{j=1}^{g} \dfrac{\xi_j^{I-}}{R_j^-} + \sum\limits_{j=g+1}^{r} \dfrac{\xi_j^-}{R_j^-} + \sum\limits_{p=1}^{o} \dfrac{\xi_p^{I+}}{R_p^+} + \sum\limits_{p=o+1}^{m} \dfrac{\xi_p^+}{R_p^+} + \sum\limits_{h=1}^{e} \dfrac{\tau_h^{I-}}{R_h} + \sum\limits_{h=e+1}^{t} \dfrac{\tau_h^-}{R_h}}{r+m+t}$$

$$s.t. \quad \sum_{i=1,i\neq k}^{q} x_{ji}\lambda_i \le x_{jk} + \xi_j^{I-}, \ j=1,2,...,g$$

$$\sum_{i=1,i\neq k}^{q} x_{ji}\lambda_i \le x_{jk} + \xi_j^-, \ j=g+1,g+2,...,r$$

$$\sum_{i=1,i\neq k}^{q} y_{pi}\lambda_i \ge y_{pk} - \xi_p^{I+}, \ p=1,2,...,o$$

$$\sum_{i=1,i\neq k}^{q} y_{pi}\lambda_i \ge y_{pk} - \xi_p^+, \ p=o+1,o+2...,m$$

$$\sum_{i=1,i\neq k}^{q} z_{hi}\lambda_i \le z_{hk} + \tau_h^{I-}, \ h=1,2,...,e$$

$$\sum_{i=1,i\neq k}^{q} z_{hi}\lambda_i \le z_{hk} + \tau_h^-, \ h=e+1,e+2,...,t$$

$$\sum_{i=1,i\neq k}^{q} \lambda_i = 1$$

$$\lambda_i, \xi_j^{I-}, \xi_j^-, \xi_p^{I+}, \xi_p^+, \tau_h^{I-}, \tau_h^- \ge 0 \ for \ all \ i,j,p,h$$

$$\xi_j^{I-} \in \phi, \ j=1,2,...,g$$

$$\xi_p^{I+} \in \phi, \ p=1,2,...,o$$

$$\tau_h^{I-} \in \phi, \ h=1,2,...,e \tag{11}$$

Proposition 4. The Super-RAM-IUDEA model (model 11) is always feasible.

Proposition 5. $\varphi_k^* \ge 1$, and $\varphi_k^* > 1$ only if DMU_k is efficient.

Proposition 6. The greater φ_k^* is, the better the performance of DMU_k would be.

Proof:

(1) We prove the Proposition 4 based on the idea of Du et al. [52]. For any $\lambda_i^{\dagger}$ ($\lambda_i^{\dagger} \ge 0$ and $\sum\limits_{i=1,i\neq k}^{q} \lambda_i^{\dagger} = 1$),

let $\xi_j^{I-\dagger} = \max\left\{\sum\limits_{i=1,i\neq k}^{q} [x_{ji}\lambda_i^{\dagger}]+1, x_{jk}\right\} - x_{jk} \ge 0$ for $j=1,2,...,g$, $\xi_j^{-\dagger} = \max\left\{\sum\limits_{i=1,i\neq k}^{q} x_{ji}\lambda_i^{\dagger}, x_{jk}\right\} - x_{jk} \ge 0$ for $j=g+1,g+2,...,r$,

$\xi_p^{I+\dagger} = y_{pk} - \min\left\{[\sum\limits_{i=1,i\neq k}^{q} y_{pi}\lambda_i^{\dagger}], y_{pk}\right\} \ge 0$ for $p=1,2,...,o$, $\xi_p^{+\dagger} = y_{pk} - \min\left\{\sum\limits_{i=1,i\neq k}^{q} y_{pi}\lambda_i^{\dagger}, y_{pk}\right\} \ge 0$ for $p=o+1,o+2...,m$,

$\tau_h^{I-\dagger} = \max\left\{\sum\limits_{i=1,i\neq k}^{q} [z_{hi}\lambda_i^{\dagger}]+1, z_{hk}\right\} - z_{hk} \ge 0$ for $h=1,2,...,e$, and $\tau_h^{-\dagger} = \max\left\{\sum\limits_{i=1,i\neq k}^{q} z_{hi}\lambda_i^{\dagger}, z_{hk}\right\} - z_{hk} \ge 0$ for $h=e+1,e+2,...,t$. Then

$\left\{\lambda_i^{\dagger}, \xi_j^{I-\dagger}, \xi_j^{-\dagger}, \xi_p^{I+\dagger}, \xi_p^{+\dagger}, \tau_h^{I-\dagger}, \tau_h^{-\dagger}\right\}$ is a feasible solution to model (11). $[\bullet]$ denotes the greatest integer less than $\bullet$.

Therefore, the Super-RAM-IUDEA model (model 11) is always feasible.

(2) $\varphi_k^* \ge 1$ because the second part of the right side of the objective function in model (11) is non-negative in value. If

DMU_k is inefficient, we can obtain $\sum\limits_{i=1}^{q} x_{ji}\lambda_i^* < x_{jk}$ for $j=1,2,...,r$, $\sum\limits_{i=1}^{q} y_{pi}\lambda_i^* > y_{pk}$ for $p=1,2,...,m$, and $\sum\limits_{i=1}^{q} z_{hi}\lambda_i^* < z_{hk}$

for $h=1,2,...,t$ from model (4). Thus, $\xi_j^{I-*}=\xi_j^{-*}=\xi_p^{I+*}=\xi_p^{+*}=\tau_h^{I-*}=\tau_h^{-*}=0$ must be the optimal solution for model (11),

which implies $\varphi_k^*=1$. Note that inefficient DMUs do not affect the efficient frontier whether they are in the reference set or

not. We have proved that $\varphi_k^* > 1$ only if DMU_k is efficient.

(3) The greater φ_k^* (the greater the second part of the right side of the objective function in model 11) is, the greater

the distance between DMU_k and the super-efficient frontier is, which indicates that there are more inputs savings, more

desirable outputs surpluses, and more undesirable outputs savings in DMU_k . Therefore, the greater φ_k^* is, the better the

performance of DMU_k would be.

3. Application

The transportation industry in China has experienced rapid growth since China's Reform and Opening in 1978. The amount of value-added based on the current prices produced by the transport industry was approximately 4033.7 billion RMB in 2018, which accounted for about 4% of China's GDP (Gross Domestic Product) [65]. However, transport accidents do not disappear although many policies and measures have been implemented to control this problem [66]. As shown in Figure 2, the number of transport accidents reached 244937 in 2018, an increase of 20% on the previous year. These transport accidents directly caused property damages amount to about RMB 1.4 billion and killed approximately 63194 people [65].

In this section, we apply our novel RAM-DEA models to evaluate the performance of China's regional transportation systems (RTSs) considering several integer-valued variables including the number of transport accidents (an undesirable output). As far as we know, limited attention has been paid to this topic [67-70]. While the number of accidents can only take integer values, these limited studies treated it as a real-valued variable. Besides the integer-valued undesirable output (the number of transport accidents), we also include one integer-valued input (the number of labor), one real-valued input (the amount of fixed capital investment, billion RMB), and one real-valued desirable output (the amount of value-added based on the current prices, billion RMB) in the model. Many scholars took "labor", "fixed capital investment", and "value-added" into account when measuring the efficiency of China's RTSs because the three variables were vital, but they treated all the variables as real-valued variables. However, as we stated above, the number of labor can only take integer values in reality.
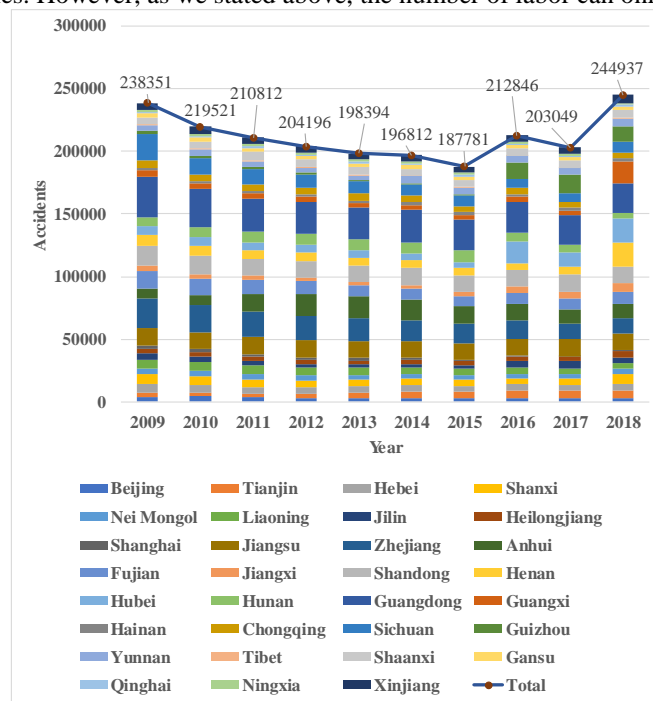
Figure 2. The number of transport accidents in China from 2009 to 2018.

3.1 Data

Let x_1, x_2, y_1, and z_1 denote the "labor", "fixed capital investment", "value-added", and "transport accidents", respectively. As shown in Table 1, the data about the RTSs of 31 provinces in mainland China in 2017 are collected from the National Bureau of Statistics of China [65]. As the data related to all the variables in 2019 and the data about "fixed capital investment" in 2018 are not available, we use the dataset in 2017. Note that Nei Mongol is also called Inner Mongolia.

The differences between China's RTSs are huge. Tibet, the least developed and sparsely populated area in China, has the fewest "labor", the least "value-added", and the fewest "transport accidents" in its RTS. Ningxia, one of the least developed areas in China, has the least "fixed capital investment". Guangdong, the most developed and populous area in China, has the most "labor", the most "value-added", and also the most "transport accidents" in its RTS. Sichuan, whose GDP ranks the 6th and population ranks the 4th in mainland China, has the most "fixed capital investment".

8

243

Table 1. Inputs and outputs.

Province	x_1	x_2	y_1	z_1	Province	x_1	x_2	y_1	z_1
Beijing	477259	112.91	120.84	3223	Guangdong	641733	375.96	358.09	23900
Tianjin	103586	53.70	78.04	5564	Guangxi	148008	200.56	95.57	3843
Hebei	199378	213.55	249.79	4848	Hainan	56464	48.60	24.89	2022
Shanxi	203859	42.55	105.21	4988	Chongqing	231369	195.48	93.95	4573
Nei Mongol	176748	118.02	105.00	3385	Sichuan	308100	449.26	159.58	6947
Liaoning	310178	60.20	131.00	4793	Guizhou	102403	233.43	107.02	14711
Jilin	123422	121.17	60.31	5485	Yunnan	148453	374.17	36.66	5371
Heilongjiang	217001	120.05	80.13	3623	Tibet	6354	58.17	3.41	329
Shanghai	369045	96.03	134.45	709	Shaanxi	227772	189.12	83.26	5819
Jiangsu	372686	289.10	309.77	13226	Gansu	112716	95.66	29.35	2755
Zhejiang	235280	296.75	193.82	12782	Qinghai	42461	73.05	10.37	1128
Anhui	199510	166.78	87.54	11506	Ningxia	32282	33.01	19.93	1569
Fujian	180441	280.89	188.97	8693	Xinjiang	151746	197.85	66.82	5017
Jiangxi	176305	73.46	86.63	5356	Max.	641733.00	449.26	358.09	23900.00
Shandong	388358	395.50	326.80	13403	Min.	6354.00	33.01	3.41	329.00
Henan	377919	249.85	216.29	6361	Ave.	219247.23	184.49	124.36	6549.97
Hubei	284346	293.99	142.00	11661	Std. Dev.	138189.40	116.82	91.18	5078.41
Hunan	191482	210.44	149.60	5459	-	-	-	-	-

244 **3.2 Efficiency**
245   The RAM-IUDEA model (model 4) is used to measure the efficiency of China's RTSs, and the results are shown in
246 Table 2. Ten RTSs are efficient DMUs, i.e., Tianjin, Hebei, Shanxi, Liaoning, Shanghai, Jiangsu, Shandong, Guangdong,
247 Tibet, and Ningxia. Note that the RTSs of Tibet, Ningxia, and Guangdong are regarded as efficient DMUs while the
248 performance of Sichuan's RTS is the worst. The four RTSs have the maximum or minimum values of the selected variables
249 as we mentioned in Subsection 3.1.

250

Table 2. Efficiency and rankings.

Province	Efficiency	Ranking	Province	Efficiency	Ranking
Beijing	0.8511	24	Hunan	0.9074	17
Tianjin	1.0000	1	Guangdong	1.0000	1
Hebei	1.0000	1	Guangxi	0.8985	19
Shanxi	1.0000	1	Hainan	0.9808	11
Nei Mongol	0.9502	13	Chongqing	0.8595	23
Liaoning	1.0000	1	Sichuan	0.7114	31
Jilin	0.9064	18	Guizhou	0.7918	28
Heilongjiang	0.9080	16	Yunnan	0.7243	30
Shanghai	1.0000	1	Tibet	1.0000	1
Jiangsu	1.0000	1	Shaanxi	0.8418	25
Zhejiang	0.8009	27	Gansu	0.9267	15
Anhui	0.8099	26	Qinghai	0.9711	12
Fujian	0.8709	20	Ningxia	1.0000	1
Jiangxi	0.9462	14	Xinjiang	0.8600	22
Shandong	1.0000	1	Max.	1.0000	-
Henan	0.8614	21	Min.	0.7114	-
Hubei	0.7480	29	Ave.	0.9073	-

251   The inputs waste, desirable outputs surpluses, and undesirable outputs savings for inefficient China's RTSs are shown
252 in Table 3. s_1^{I-}, s_2^-, s_1^+, and ς_1^{I-} respectively indicate the labor inputs waste, the fixed capital investment inputs waste, the
253 value-added outputs shortage, and the transport accidents outputs excess. Decision makers can further calculate the targets (to
254 eliminate slacks) for DMUs based on the results and equations (5)-(10). For example, Beijing's RTS has the largest number of
255 labor to cut, Yunnan's RTS has the largest amount of fixed capital investment to reduce, Qinghai's RTS has the largest amount
256 of value-added to increase, and Guizhou's RTS has the largest number of transport accidents to reduce.
257   The results show that our novel RAM-IUDEA model is able to generate the integer-valued slacks (and targets) for integer-
258 valued variables. This advantage of our model is important for managers because it makes it simple for them to make and
259 implement an efficiency improvement plan. Moreover, integer-valued DEA models are more reliable than real-valued DEA
260 models as we explained in Section 1 and Section 2.

261
262
263
264
265

Table 3. Inputs waste, desirable outputs shortage, and undesirable outputs excess for inefficient RTSs.

Province	s_1^{I-}	s_2^-	s_1^+	ς_1^{I-}
Beijing	371627	0.623	-	215
Nei Mongol	82632	18.173	-	603
Jilin	61786	56.443	-	3340
Heilongjiang	140954	39.759	-	1195
Zhejiang	76597	127.139	-	8733
Anhui	118085	80.649	-	8973
Fujian	25281	115.095	-	4713
Jiangxi	72746	-	-	2378
Henan	202895	62.608	-	1991
Hubei	163330	165.089	-	8351
Hunan	64934	75.585	-	2040
Guangxi	60740	108.139	-	1195
Hainan	20574	11.696	-	382
Chongqing	145278	104.332	-	1948
Sichuan	174301	306.559	-	3386
Guizhou	6816	132.003	-	11900
Yunnan	104006	328.022	-	3563
Shaanxi	149458	106.353	-	3347
Gansu	73591	55.237	-	1052
Qinghai	19400	31.092	3.685	-
Xinjiang	85377	128.011	-	2779
Max.	371627	328.022	3.685	11900
Min.	6816	0.623	3.685	215
Ave.	105733.714	102.630	3.685	3604.200

3.3 Super-Efficiency

In order to differentiate efficient China's RTSs the Super-RAM-IUDEA model (model 11) is applied. As shown in Table 4, the performance of Hebei's RTS is the best, followed by Shanghai, Guangdong, Tibet, Ningxia, Shandong, Shanxi, Jiangsu, Liaoning, and Tianjin.

Table 4. Super-efficiency and rankings.

Province	Efficiency	Ranking
Tianjin	1.0058	10
Hebei	1.0753	1
Shanxi	1.0102	7
Liaoning	1.0081	9
Shanghai	1.0342	2
Jiangsu	1.0101	8
Shandong	1.0114	6
Guangdong	1.0243	3
Tibet	1.0229	4
Ningxia	1.0135	5

Table 5 shows the inputs savings, desirable outputs surpluses, and undesirable outputs savings for efficient RTSs. ξ_1^{I-}, ξ_1^-, ξ_1^+, and τ_1^{I-} represent the labor inputs savings, fixed capital investment inputs savings, value-added outputs surpluses, and transport accidents outputs savings, respectively. The results not only show the rankings of efficient RTSs but also enable decision makers to identify the strengths of efficient RTSs. Moreover, the results are reliable because the models can generate integer values for integer-valued variables and real values for real-valued variables.

Table 5. Inputs savings, desirable outputs surpluses, and undesirable outputs savings for efficient RTSs.

Province	ξ_1^{I-}	ξ_2^-	ξ_1^+	τ_1^{I-}
Tianjin	-	9.718	-	-
Hebei	-	-	100.860	400
Shanxi	-	16.957	-	-
Liaoning	-	-	11.548	-
Shanghai	-	-	-	3229
Jiangsu	-	-	14.294	-
Shandong	-	-	16.229	-
Guangdong	-	-	34.418	-
Tibet	32325	-	-	963
Ningxia	-	22.521	-	-

278   Finally, we get the performance rankings of China's RTSs. For the sake of clarity, Figure 3 shows the efficiency (or
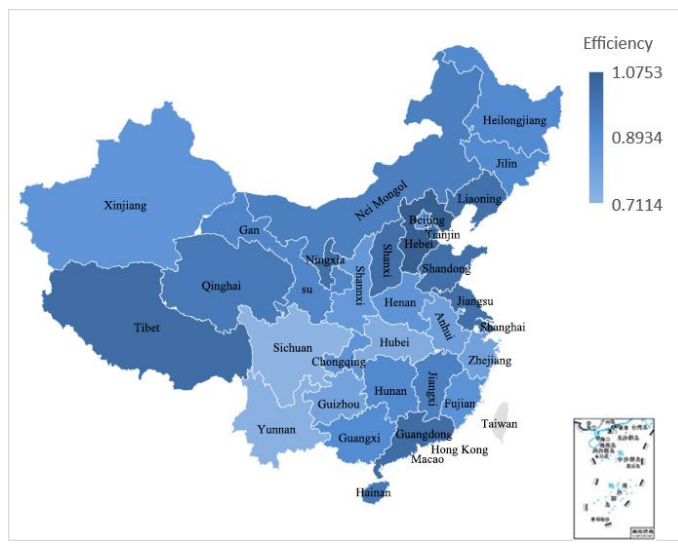279 super-efficiency) of them instead of their rankings.

Figure 3. The efficiency of China's RTSs.

280
281

3.4 Discussion

283   Decision makers should take undesirable outputs into account when measuring the performance of DMUs with undesirable
284 outputs. Otherwise, the results may be problematic. If we do not consider the "transport accidents" when measuring the
285 efficiency of China's RTSs, the results would be inaccurate as shown in Table 6. For example, the RTS of Shanghai is regarded
286 as an inefficient DMU if we apply the integer-valued super-efficiency RAM-DEA model without undesirable outputs (Super-
287 RAM-IDEA). Shanghai is the most densely populated area and the amount of value-added by Shanghai's RTS ranks the 11th in
288 mainland China. But the number of transport accidents in Shanghai is only 709, which is less than the other provinces except
289 Tibet. Therefore, the ranking of Shanghai's RTS should not be only the 18th (the results obtained from the Super-RAM-IDEA
290 model).

Table 6. Rankings obtained from our Super-RAM-IUDEA model and the Super-RAM-IDEA model.

Province	Super-RAM-IUDEA	Super-RAM-IDEA	Province	Super-RAM-IUDEA	Super-RAM-IDEA
Beijing	24	28	Hubei	29	29
Tianjin	10	9	Hunan	17	16
Hebei	1	1	Guangdong	3	2
Shanxi	7	7	Guangxi	19	21
Nei Mongol	13	13	Hainan	11	10
Liaoning	9	8	Chongqing	23	26
Jilin	18	14	Sichuan	31	31
Heilongjiang	16	20	Guizhou	28	19
Shanghai	2	18	Yunnan	30	30
Jiangsu	8	5	Tibet	4	4
Zhejiang	27	23	Shaanxi	25	27
Anhui	26	22	Gansu	15	15
Fujian	20	17	Qinghai	12	12
Jiangxi	14	11	Ningxia	5	3
Shandong	6	6	Xinjiang	22	24
Henan	21	25	-	-	-

292   In addition, decision makers cannot simply treat integer-valued variables as reals and round the real-valued solutions to
293 integers. If we apply the real-valued super-efficiency RAM-DEA model with undesirable outputs (Super-RAM-UDEA) to
294 measure the super-efficiency of DMUs with integer-valued variables, the resulting inputs savings, desirable outputs
295 surpluses, and undesirable outputs savings for efficient DMUs may be wrong. Similarly, if we apply the real-valued RAM-
296 DEA model with undesirable output (RAM-UDEA) to measure the efficiency of DMUs with integer-valued variables, the
297 resulting inputs waste, desirable outputs shortage, and undesirable outputs excess for inefficient DMUs may be also
298 erroneous. The results obtained from RAM-UDEA and Super-RAM-UDEA models are shown in Table 7 and Table 8, where
299 "N" indicates that the real-valued solutions for the RTS obtained from the RAM-UDEA or Super-RAM-UDEA models cannot
300 be rounded to our integer-valued solutions.

Table 7. Inputs waste, desirable outputs shortage, and undesirable outputs excess for inefficient RTSs resulting from the RAM-UDEA model.

Province	s_1^{I-}	s_2^{-}	s_1^{+}	ς_1^{I-}
Beijing	371620.800	0.642	-	214.498
Nei Mongol	82624.630	18.193	-	602.459
Jilin	61785.890	56.444	-	3339.971
Heilongjiang	140956.800	39.757	-	1195.235
Zhejiang	76589.210	127.161	-	8732.424
Anhui	118079.100	80.667	-	8972.529
Fujian	25275.910	115.110	-	4712.610
Jiangxi	72760.360	-	-	2377.760
Henan	202893.700	62.612	-	1990.884
Hubei	163325.600	165.102	-	8350.646
Hunan	64936.800	75.583	-	2040.231
Guangxi	60739.740	108.140	-	1194.980
Hainan	20576.340	11.694	-	382.245
Chongqing	145278.400	104.332	-	1948.090
Sichuan	174299.900	306.564	-	3385.864
Guizhou	6811.200	132.017	-	11899.640
Yunnan	104009.200	328.020	-	3563.343
Shaanxi	149452.500	106.368	-	3346.585
Gansu	73586.160	55.251	-	1051.622
Qinghai	19400.170	31.092	3.685	-
Xinjiang	85377.460	128.011	-	2779.105

Table 8. Inputs savings, desirable outputs surpluses, and undesirable outputs savings for efficient RTSs resulting from the Super-RAM-UDEA model.

Province	ξ_1^{I-}	ξ_2^{-}	ξ_1^{+}	τ_1^{I-}
Tianjin	-	9.718	-	-
Hebei	-	-	100.864	399.774
Shanxi	-	16.957	-	-
Liaoning	-	-	11.548	-
Shanghai	-	-	-	3228.896
Jiangsu	-	-	14.294	-
Shandong	-	-	16.229	-
Guangdong	-	-	34.418	-
Tibet	32324.190	-	-	962.888
Ningxia	-	22.521	-	-

303   The results show that there are huge differences between the results for China's RTSs efficiency measurement obtained
304 from our integer-valued models and those resulting from real-valued models. As to the inefficient RTSs, about 76% of
305 RTSs' real-valued solutions obtained from the RAM-UDEA cannot be rounded to the integer-valued solutions. For instance,
306 the labor inputs slack in the RTS of Heilongjiang should be 140954 (as shown in Table 3), but the result obtained from the
307 Super-RAM-UDEA model is 140956.800 (as shown in Table 7). The labor inputs savings in the RTS of Tibet should be
308 32325 (as shown in Table 5), but the result obtained from the Super-RAM-UDEA model is 32324.190 (as shown in Table 8).
309 The reason for the differences is that the PPS (as shown in formula 3) for DMUs could be changed if decision makers do not
310 consider their integer-valued variables when measuring efficiency. Therefore, our integer-valued models are better than
311 traditional real-valued models.
312   As stated above, the results may be problematic if decision makers did not take undesirable outputs into account when
313 measuring the performance of DMUs with undesirable outputs or restrict variables to integers when measuring the performance
314 of DMUs with integer-valued variables. The proposed RAM-IUDEA and Super-RAM-IUDEA models can be used to deal with
315 integer-valued variables and undesirable outputs, which can help decision makers accurately measure the efficiency and super-
316 efficiency of DMUs.
317

318 **4. Conclusion**
319   Non-radial DEA models, i.e., additive DEA, SBM, and RAM, have greater discriminatory ability than radial DEA
320 models. However, integer-valued additive DEA models cannot generate efficiency scores for DMUs and integer-valued SBM-
321 DEA models are nonlinear programming even after Charnes-Cooper transformation. Therefore, we develop an integer-valued
322 super-efficiency RAM-DEA model that is linear programming and can generate efficiency scores. Moreover, the model can
323 handle undesirable outputs that can significantly affect the DEA efficiency measurement of DMUs.

The main contributions of this study are as follows. (1) An integer-valued super-efficiency RAM-DEA model with undesirable outputs is proposed, which can be used to accurately rank efficient DMUs with integer-valued variables and undesirable outputs. (2) The proposed model is applied to evaluate the efficiency of China's regional transportation systems (RTSs) considering the number of transport accidents (an undesirable output). The results help decision makers improve the performance of inefficient RTSs and analyze the strengths of efficient RTSs. The empirical study shows that it is necessary to consider integer-valued variables and undesirable outputs when measuring efficiency.

We also propose an integer-valued RAM-DEA model with undesirable outputs besides our integer-valued super-efficiency RAM-DEA model. However, the objective function of this model is to maximize the slacks of the selected variables so that the resulting targets for inefficient DMUs are the farthest. In the future, we will research the RAM-DEA approach to find the closest targets for DMUs.

The integer-valued super-efficiency RAM-DEA model is proposed under the assumption of strong disposability of undesirable outputs. This assumption is based on our belief that the undesirable output (the number of transport accidents) in our case study can be disposed of. Although the disposal of the undesirable output involves some costs, it is possible to be done with an acceptable increase in the costs of production. However, there are some undesirable outputs which may be weakly disposable, e.g., the amount of carbon emissions. In the future, it is necessary to develop an integer-valued super-efficiency RAM-DEA model considering both strongly and weakly disposable undesirable outputs at the same time.

Acknowledgment

References

[1] Zelenyuk V. Aggregation of inputs and outputs prior to data envelopment analysis under big data. European Journal of Operational Research, 2020, 282(1): 172-187.

[2] Yu M, Chen L. Evaluation of efficiency and technological bias of tourist hotels by a meta-frontier DEA model. Journal of the Operational Research Society, 2019, 12(4): 1-15.

[3] Guner S, Coskun E. Estimating the operational and service efficiency of bus transit routes using a non-radial DEA approach. EURO Journal on Transportation and Logistics, 2019, 8(3): 249-268.

[4] Chen P, Yu M, Shih J, et al. A reassessment of the Global Food Security Index by using a hierarchical data envelopment analysis approach. European Journal of Operational Research, 2019, 272(2): 687-698.

[5] Charnes A, Cooper W W, Rhodes E, et al. Measuring the efficiency of decision making units. European Journal of Operational Research, 1978, 2(6): 429-444.

[6] Banker R D, Charnes A, Cooper W W, et al. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management Science, 1984, 30(9): 1078-1092.

[7] Caldas P, Ferreira D C, Dollery B, et al. Are there scale economies in urban waste and wastewater municipal services? A non-radial input-oriented model applied to the Portuguese local government. Journal of Cleaner Production, 2019, 219: 531-539.

[8] Chen X, Wu G, Li D, et al. Efficiency measure on the truck restriction policy in China: A non-radial data envelopment model. Transportation Research Part A-Policy and Practice, 2019, 129: 140-154.

[9] Zhou Z, Wu H, Song P, et al. Measuring the resource and environmental efficiency of industrial water consumption in China: A non-radial directional distance function. Journal of Cleaner Production, 2019, 240:1-11.

[10] Charnes A, Cooper W W, Golany B, et al. Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions. Journal of Econometrics, 1985, 30(1-2): 91-107.

[11] Tone K. A slacks-based measure of efficiency in data envelopment analysis. European Journal of Operational Research, 2001, 130(3): 498-509.

[12] Cooper W W, Park K S, Pastor J T, et al. RAM: A range adjusted measure of inefficiency for use with additive models, and relations to other models and measures in DEA. Journal of Productivity Analysis, 1999, 11(1): 5-42.

[13] Yu S. Benchmarking and performance evaluation towards the sustainable development of regions in Taiwan: A minimum distance-based measure with undesirable outputs in additive DEA. Social Indicators Research, 2019, 144(3): 1323-1348.

[14] Arteaga F J, Tavana M, Caprio D D, et al. A dynamic multi-stage slacks-based measure data envelopment analysis model with knowledge accumulation and technological evolution. European Journal of Operational Research, 2019, 278(2): 448-462.

[15] Khezrimotlagh D, Salleh S, Mohsenpour Z, et al. A new robust mixed integer-valued model in DEA. Applied Mathematical Modelling, 2013, 37(24): 9885-9897.

[16] Taleb M, Ramli R, Khalid R, et al. Developing a two-stage approach of super efficiency slack-based measure in the presence of non-discretionary factors and mixed integer-valued data envelopment analysis. Expert Systems with Applications, 2018, 103: 14-24.

[17] Lozano S, Villa G. Data envelopment analysis of integer-valued inputs and outputs. Computers & Operations Research, 2006, 33(10): 3004-3014.

[18] Chen C H, Liu H H, Liu B, et al. Measuring the performance of pallet rental companies: Integer-valued DEA models with generalized reference sets. IEEE Access, 2020, 8: 3374-3386.

[19] Kuosmanen T, Matin R K. Theory of integer-valued data envelopment analysis. European Journal of Operational Research, 2009, 192(2):658-667.

[20] Aida K, Cooper W W, Pastor J T, et al. Evaluating water supply services in Japan with RAM: A range-adjusted measure of inefficiency. Omega-International Journal of Management Science, 1998, 26(2): 207-232.

[21] Steinmann L, Zweifel P. The range adjusted measure (RAM) in DEA: Comment. Journal of Productivity Analysis, 2001, 15(2): 139-144.

[22] Cooper W W, Park K S, Pastor J T, et al. The range adjusted measure (RAM) in DEA: A response to the comment by Steinmann and Zweifel. Journal of Productivity Analysis, 2001, 15(2): 145-152.

[23] Sueyoshi T, Sekitani K. Measurement of returns to scale using a non-radial DEA model: A range-adjusted measure approach. European Journal of Operational Research, 2007, 176(3): 1918-1946.

[24] Cooper W W, Pastor J T, Borras F, et al. BAM: A bounded adjusted measure of efficiency for use with bounded additive models. Journal of Productivity Analysis, 2011, 35(2): 85-94.

[25] Chiu C, Chiu Y, Fang C, et al. The performance of commercial banks based on a context‑dependent range‑adjusted measure model. International Transactions in Operational Research, 2014, 21(5): 761-775.

[26] Avkiran N K, McCrystal A. Dynamic network range-adjusted measure vs. dynamic network slacks-based measure. Journal of The Operations Research Society of Japan, 2014, 57(1): 1-14.

[27] Tavassoli M, Badizadeh T, Saen R F, et al. Performance assessment of airlines using range-adjusted measure, strong complementary slackness condition, and discriminant analysis. Journal of Air Transport Management, 2016, 54: 42-46.

[28] Liu X, Zhu Q, Chu J, et al. Environmental performance and benchmarking information for coal-fired power plants in China: A DEA approach. Computational Economics, 2019, 54(4): 1287-1302.

[29] Li Y, Wang Y, Cui Q, et al. Energy efficiency measures for airlines: An application of virtual frontier dynamic range adjusted measure. Journal of Renewable and Sustainable Energy, 2016, 8(1): 1-14.

[30] Wanke P, Barros C P. Efficiency in Latin American airlines: A two-stage approach combining virtual frontier dynamic DEA and simplex regression. Journal of Air Transport Management, 2016, 54: 93-103.

[31] Li Y, Cui Q. Airline efficiency with optimal employee allocation: An Input-shared Network Range Adjusted Measure. Journal of Air Transport Management, 2018, 73: 150-162.

[32] Walheer B. Output, input, and undesirable output interconnections in data envelopment analysis: Convexity and returns-to-scale. Annals of Operations Research, 2020, 284(1): 447-467.

[33] Wang H, Pan C, Wang Q, et al. Assessing sustainability performance of global supply chains: An input-output modeling approach. European Journal of Operational Research, 2020, Available: https://www.sciencedirect.com/science/article/pii/S0377221720300941.

[34] Sueyoshi T, Goto M. Should the US clean air act include CO2 emission control?: Examination by data envelopment analysis. Energy Policy, 2010, 38(10): 5902-5911.

[35] Sueyoshi T, Goto M. Methodological comparison between two unified (operational and environmental) efficiency measurements for environmental assessment. European Journal of Operational Research, 2011, 210(3): 684-693.

[36] Ramli N A, Munisamy S. Eco-efficiency in greenhouse emissions among manufacturing industries: A range adjusted measure. Economic Model, 2015, 47: 219-227.

[37] Wang K, Yu X. Industrial energy and environment efficiency of Chinese cities: an analysis based on range-adjusted measure. International Journal of Information Technology and Decision Making, 2017, 16(4): 1023-1042.

[38] Wang W, Yu B, Yan X, et al. Estimation of innovation's green performance: A range-adjusted measure approach to assess the unified efficiency of China's manufacturing industry. Journal of Cleaner Production, 2017, 149: 919-924.

[39] Cui Q, Li Y. Airline environmental efficiency measures considering materials balance principles: An application of a network range-adjusted measure with weak-G disposability. Journal of Environmental Planning and Management, 2018, 61(13): 2298-2318.

[40] Meng M, Fu Y, Wang L, et al. Low-carbon economy efficiency analysis of China's provinces based on a range-adjusted measure and data envelopment analysis model. Journal of Cleaner Production, 2018, 199: 643-650.

[41] Kang Y, Xie B, Wang J, et al. Environmental assessment and investment strategy for China's manufacturing industry: A non-radial DEA based analysis. Journal of Cleaner Production, 2018, 175: 501-511.

[42] Miao Z, Chen X, Baležentis T, et al. Atmospheric environmental productivity across the provinces of China: Joint decomposition of range adjusted measure and Luenberger productivity indicator. Energy Policy, 2019, 132: 665-677.

435 [43] Eftekharian P, Homayonifar M, Kazemi M, et al. Environmental efficiency of energy intensive industries in Iran: Range
436   Adjusted Measure. Journal of Energy Planning and Policy Research, 2019, 5(16): 33-66.
437 [44] Wang Q, Yuan Q. Energy-saving and emission reduction potential of the tobacco industry: A case study of China's 18
438   cigarette enterprises. Journal of Cleaner Production, 2020, Available: https://doi.org/10.1016/j.jclepro.2019.118429.
439 [45] Chen Z, Zhao W, Li F, Wang W. Congestion assessment for the Belt and Road countries considering carbon emission
440   reduction. Journal of Cleaner Production, 2020, Available: https://doi.org/10.1016/j.jclepro.2019.118405.
441 [46] Yuan Q Q, Cheng C F C, Wang J Y, et al. Inclusive and sustainable industrial development in China: An efficiency-
442   based analysis for current status and improving potentials. Applied Energy, 2020, Available:
443   https://doi.org/10.1016/j.apenergy.2020.114876.
444 [47] Ren J, Chen C, Gao B, et al. Performance evaluation of pallet rental companies: A non-oriented super-efficiency integer-
445   valued DEA model. IEEE Access, 2019, 7: 151628-151637.
446 [48] Lin R, Yang W, Huang H, et al. A modified slacks-based super-efficiency measure in the presence of negative data.
447   Computers & Industrial Engineering, 2019, 135: 39-52.
448 [49] Andersen P, Petersen N C. A procedure for ranking efficient units in data envelopment analysis. Management Science,
449   1993, 39(10): 1261-1264.
450 [50] Gokgoz F, Erkul E. Investigating the energy efficiencies of European countries with super efficiency model and super
451   SBM approaches. Energy Efficiency, 2019, 12(3): 601-618.
452 [51] Ma X, Wang C, Yu Y, et al. Ecological efficiency in China and its influencing factors—a super-efficient SBM
453   metafrontier-Malmquist-Tobit model study. Environmental Science and Pollution Research, 2018, 25(21): 20880-20898.
454 [52] Du J, Chen C, Chen Y, et al. Additive super-efficiency in integer-valued data envelopment analysis. European Journal of
455   Operational Research, 2012, 218(1): 186-192.
456 [53] Guo I, Lee H, Lee D, et al. An integrated model for slack-based measure of super-efficiency in additive DEA. Omega-
457   International Journal of Management Science, 2017, 67: 160-167.
458 [54] Yu S, Hsu C. A unified extension of super-efficiency in additive data envelopment analysis with integer-valued inputs
459   and outputs: An application to a municipal bus system. Annals of Operations Research, 2020, 287(1): 515-535.
460 [55] Tone K. A slacks-based measure of super-efficiency in data envelopment analysis. European Journal of Operational
461   Research, 2001, 143(1): 32-41.
462 [56] Tran T H, Mao Y, Nathanail C P, et al. Integrating slacks-based measure of efficiency and super-efficiency in data
463   envelopment analysis. Omega-International Journal of Management Science, 2019, 85: 156-165.
464 [57] Chen F, Zhao T, Wang J, et al. The evaluation of energy–environmental efficiency of China's industrial sector: Based
465   on Super-SBM model. Clean Technologies and Environmental Policy, 2019, 21(7): 1397-1414.
466 [58] Chen C M , Du J , Huo J , et al. Undesirable factors in integer-valued DEA: Evaluating the operational efficiencies of
467   city bus systems considering safety records. Decision Support Systems, 2012, 54(1):330-335.
468 [59] Färe R, Grosskopf S, Lovell C A, et al. Multilateral productivity comparisons when some outputs are undesirable: A
469   nonparametric approach. The Review of Economics and Statistics, 1989, 71(1): 90-98.
470 [60] Chung Y, Fare R, Grosskopf S, et al. Productivity and undesirable outputs: A Directional Distance Function Approach.
471   Journal of Environmental Management, 1997, 51(3): 229-240.
472 [61] Layer K, Johnson A L, Sickles R C, et al. Direction selection in stochastic directional distance functions. European
473   Journal of Operational Research, 2020, 280(1): 351-364.
474 [62] Seiford L M, Zhu J. Modeling undesirable factors in efficiency evaluation. European Journal of Operational Research,
475   2002, 142(1): 16-20.
476 [63] Halkos G, Petrou K N. Treating undesirable outputs in DEA: A critical review. Economic Analysis and Policy, 2019,
477   62: 97-104.
478 [64] Du J, Liang L, Zhu J, et al. A slacks-based measure of super-efficiency in data envelopment analysis: A comment.
479   European Journal of Operational Research, 2010, 204(3): 694-697.
480 [65] National Bureau of Statistics of China. China statistical yearbook, 2020, Available:
481   http://data.stats.gov.cn/easyquery.htm?cn=E0103.
482 [66] Liu X, Turla T, Zhang Z, et al. Accident-cause-specific risk analysis of rail transport of hazardous materials.
483   Transportation Research Record, 2018, 2672(10): 176-187.
484 [67] Chiu Y, Huang C, Ma C, et al. Assessment of China transit and economic efficiencies in a modified value-chains DEA
485   model. European Journal of Operational Research, 2011, 209(2): 95-103.
486 [68] Pal D, Mitra S K. An application of the directional distance function with the number of accidents as an undesirable
487   output to measure the technical efficiency of state road transport in India. Transportation Research Part A-policy and
488   Practice, 2016, 93: 1-12.
489 [69] Tian N, Tang S, Che A, et al. Measuring regional transport sustainability using super-efficiency SBM-DEA with
490   weighting preference. Journal of Cleaner Production, 2020, 242: 1-11.

491 [70] Stefaniec A, Hosseini K, Xie J, et al. Sustainability assessment of inland transportation in China: A triple bottom line-
492 based network DEA approach. Transportation Research Part D-Transport and Environment, 2020, Available:
493 https://doi.org/10.1016/j.trd.2020.102258.
494