Available online at www.sciencedirect.com
ScienceDirect
Omega 35 (2007) 578–587
omega
The International Journal of Management Science
www.elsevier.com/locate/omega

# Eco-efficiency analysis of paper mills along the Huai River: An extended DEA approach☆

Zhongsheng Hua*, Yiwen Bian, Liang Liang

*Department of Information Management & Decision Science, School of Management, University of Science and Technology of China, Hefei, Anhui 230026, People's Republic of China*

Received 24 June 2005; accepted 18 November 2005
Available online 9 January 2006

**Abstract**

The objective of this paper is to estimate the ecological efficiency of paper mills along the Huai River in China. The main characteristic of the ecological efficiency evaluation problem is that an undesirable output of biochemical oxygen demand (BOD) and a non-discretionary input (BOD emission quota) should be considered simultaneously. By analyzing the impacts of the non-discretionary input on decision-making units' (DMUs) desirable and undesirable outputs, a non-radial output-oriented DEA model is proposed. In the proposed model, we describe a new approach of defining reference set that requires reference units operate in a similar environment on average. We employ the model to provide efficient inputs/outputs targets for DMU managers to improve DMUs' efficiencies. Based on the developed model, impacts of the non-discretionary input on DMUs' returns are also analyzed. We illustrate the proposed model, using real data, for 32 paper mills along the Huai River in China.


*Keywords:* Data envelopment analysis; Non-radial model; Undesirable output; Non-discretionary input

## 1. Introduction

The Huai River is one of the seven largest rivers in China, which is about mid-way between the Yellow River and the Yangzi River. Like them it runs from west to east, and its location is illustrated in Fig. 1.

Originating in the Tongbai Mountains in Henan province, it flows through Henan, Anhui, Shandong and Jiangsu provinces and into the Hongze Lake in Jiangsu province. With more than 500 branches, the Huai River basin is home to some 200 million people, less than 38 million of whom are urban residents. This heavily agricultural area, with one-eighth of the national farmland, produces nearly a quarter of the country's marketed grain, cotton, and oilseeds. Other agricultural undertakings, such as animal husbandry and fish and other aquatic production, have also flourished. However, this productivity has not translated into high incomes, for basin residents produce less than two-thirds of the national average per capita GNP.

A growing proportion of income is generated in basin's industry, dominated by agricultural product processing, including paper making, brewing, tanneries, chemical production and tobacco and food processing. Driven by sustained, rapid growth of industrial, agricultural and municipal pollutants, water quality of the River has deteriorated at an accelerating rate over the past 20 years. The main pollutants in the River are

☆ This manuscript was processed by Area Editor Prof. B. Lev.

* Corresponding author. Tel.: +86 551 3607792; fax: +86 551 360002.
*E-mail address:* zshua@ustc.edu.cn (Z. Hua).

0305-0483/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.omega.2005.11.001

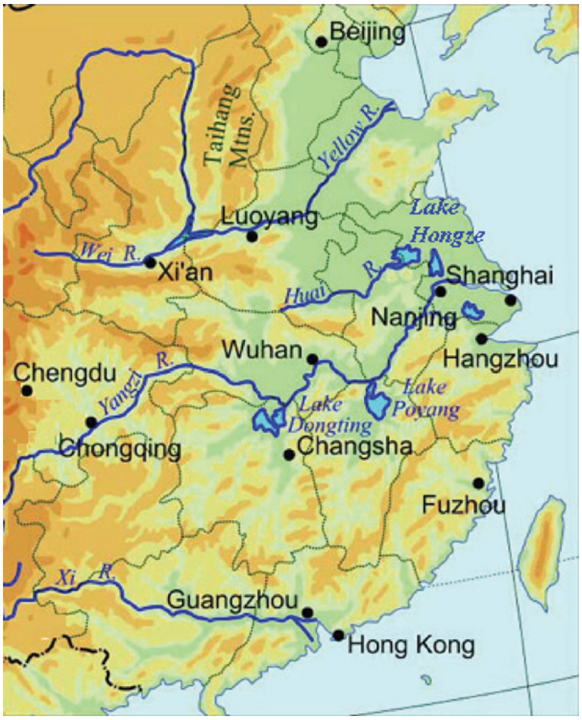


Fig. 1. Location of the Huai River.

biochemical/chemical oxygen demand (BOD/COD), nitrogen (N) and ammonia. For example, emission amount of BOD/COD in 2003 was 1,232,000 ton, which was more than three times of permitted emission amount (368,000 ton); emitted amount of N and ammonia in 2004 was 130,000 ton, which was more than four times of permitted emission amount (26,600 ton) [1]. The polluted water of the River is badly endangering the health of the basin residents.

To deal with the increasingly serious environmental problem of the Huai River, State Environmental Protection Administration of China (EPAC) had passed an Act limiting the total quantities of main pollutants emission (e.g., BOD/COD, N, etc.) from industrial enterprises (e.g., paper mills) along the River [1]. According to the EPAC's Act, an annual emission quota of BOD is allocated to each paper mill along the River. Every paper mill should not emit more BOD than the allocated quota or it would be forced to be closed (especially for paper mills with middle- or small-scale sizes). For example, during 1994–1996, 1111 paper mills along the River were forced to be closed. For paper mills, emission of BOD is an undesirable output, and the corresponding emission quota is a non-discretionary input allocated by EPAC. The non-discretionary input, on the one hand, restricts pollutants emission; while on the other hand, it also affects the production of desirable outputs (paper products). For example, in 1997, total amount of paper products in Shandong province was 3,422,500 ton, but in 1998 it was reduced to 2,790,800 ton [2]. When we evaluate efficiencies of paper mills along the River, undesirable outputs and non-discretionary inputs should be considered simultaneously.

The rest of paper is organized as follows: in Section 2, we make a brief literature review on DEA models with non-discretionary inputs or in the presence of undesirable outputs. In Section 3, we propose a DEA-based non-radial model for efficiency evaluation incorporating non-discretionary inputs and undesirable outputs. Based on the model, we present an approach to analyze decision-making units' (DMUs') returns in Section 4. We illustrate our approach, using real data set, for 32 paper mills in Section 5. Section 6 concludes the paper.

## 2. Literature review

There are many research results on performance evaluation considering non-discretionary inputs or in the presence of undesirable outputs.

The literature on the measurement of ecological efficiency has grown substantially since Färe et al. [3] firstly introduced a non-linear programming problem for efficiency evaluation in the presence of undesirable outputs. Scheel [4] presented some radial measures which assume that any change of output level will involve both desirable and undesirable outputs. Seiford and Zhu [5] developed a radial DEA model, in the presence of undesirable outputs, to improve the efficiency via increasing desirable outputs and decreasing undesirable outputs. Vencheh et al. [6] developed a DEA-based model for efficiency evaluation incorporating undesirable inputs and undesirable outputs, simultaneously. For more related researches, one can see [7–9], etc.

In current literature of ecological efficiency evaluation, researchers focused on incorporating undesirable outputs into their models, without considering impacts of non-discretionary inputs on DMUs' performance.

Since Banker and Morey [10] presented the first DEA model for controlling non-discretionary inputs, there have been many modified DEA models developed to dispose the non-discretionary inputs in performance evaluation problems (e.g., [11–14]).

In [10], Bank and Morey defined non-discretionary inputs as those inputs that were exogenously fixed and beyond the discretionary control of DMUs' managers. In an empirical illustration example about fast food restaurants, they took four factors (i.e., the age of the store, the advertising expenditures allocated to the store by the headquarters, store's location: urban or rural area, and whether it had a drive-in window) as non-discretionary inputs of a fast food restaurant store.

Ruggiero [11] proposed a DEA model that allowed environmental variables (non-discretionary inputs), e.g., parental educational backgrounds (percentage of adults with college education), to be included in evaluating educational production efficiencies of schools in New York State districts. He extended [12] the model proposed in [11] for allowing multiple non-discretionary inputs included in the DEA model, and further extended [13] this work again by considering the correlation between non-discretionary factors and technical efficiency. In all his researches, socio-economic environment of students is treated as a non-discretionary input.

Although we cannot find any formal definition for non-discretionary input in DEA literature, we can find some characteristics of non-discretionary input from the above examples and also from other relative literatures (e.g., [15,16]). Non-discretionary inputs comprises those factors that reflect different background peculiarities and potential political restrictions (both organizational and managerial), which may affect the performance of each DMU and thereby influence the results of each evaluated unit, either favorably or unfavorably. A main distinction between discretionary and non-discretionary inputs may be that, the former is consumed in producing the outputs, while the later is usually a kind of environmental constraints. There are two different ways in disposing of non-discretionary input in DEA analyses. One is to include them in DEA models [10–13]; the other is to exclude them from the DEA models [14].

In a case study about estimating relative efficiency in the public school districts of Connecticut, Ray [14] separated the community social–economic status variables expected to affect achievement from the school inputs. He defined social–economic factors, e.g., parental education measured by the percentage of the population 25 years or older with 4 or more years of college education (COLL), percentage of students in the district from ethnic minority groups (MIN) and percentage of children in the community from single-parent families (SPRNT), as non-discretionary inputs of public schools. DEA was performed with the school discretionary inputs only. Efficiency measures obtained from DEA were subsequently related to the non-discretionary inputs in a regression model. His findings from DEA and regression analyses suggested that differences of productivity of schools can be ascribed to a large extent to differences in their non-discretionary inputs. Therefore, his result is essentially consistent with the idea behind the reference unit selection when non-discretionary inputs are included in DEA model.

When non-discretionary inputs are included in DEA models, their impacts were usually reflected by the way of reference units selecting. In the BM model [10], the reference set may include units that operate in the similar or more unfavorable environment compared with the assessed unit. In the Ruggiero model [11–13], non-discretionary inputs are treated as "jumps", i.e., discontinuity between the production possibility sets in different environments. The composite reference group may only include those units that operate in similar or more difficult environments.

Their ways of reference unit selection are not suitable for the performance analysis problem with undesirable outputs. This is simply because that the BM and the Ruggiero models assume positive impacts of non-discretionary inputs on (desirable) outputs, while this assumption is not true in the presence of undesirable outputs. For given levels of discretionary inputs, a DMU with more non-discretionary inputs will emit more pollutants, i.e., non-discretionary inputs have negative impacts on eco-efficiency.

Furthermore, since desirable outputs and undesirable outputs may not be improved equiproportionally, potential problems arise with radial performance measure [17]. With radial measure, even an efficient DMU may still have slacks in some outputs.

To do performance analysis when impacts of non-discretionary inputs on DMUs' desirable outputs and undesirable outputs should be considered simultaneously, we develop a DEA-based non-radial linear programming model. In the proposed model, we examine the eco-efficiency evaluation problem in the context of the constrained reference set, and then estimate possible augmentation of desirable outputs and possible reduction of undesirable outputs, keeping non-discretionary inputs constant.

## 3. The model

This section introduces the DEA-based non-radial linear programming model for performance analysis incorporating undesirable outputs and non-discretionary inputs simultaneously. Suppose we have $n$ independent homogeneous decision making units, denoted by $\mathrm{DMU}_j$ $(j = 1, 2, \ldots, n)$. For given $p$ non-discretionary inputs $Z_j = (z_{1j}, z_{2j}, \ldots, z_{pj})^{\mathrm{T}}$, each DMU consumes $m$ discretionary inputs $X_j = (x_{1j}, x_{2j}, \ldots, x_{mj})^{\mathrm{T}}$ to produce $s$ outputs $Y_j = (y_{1j}, y_{2j}, \ldots, y_{sj})^{\mathrm{T}}$. The outputs corresponding to indices $1, 2, \ldots, k$ are desirable and the outputs corresponding to indices $k+1, k+2, \ldots, s$ are undesirable. We like to produce desirable outputs as much as possible and undesirable outputs as little as possible for given levels of non-discretionary inputs. Let $Y \in \Re_{+}^{s\times n}$ be the matrix, consisting of non-negative elements, containing the observed desirable and undesirable outputs for DMUs.

The data matrix $Y$ can be represented as follows:

$$Y = \begin{pmatrix} Y^{\mathrm{g}} \\ Y^{\mathrm{b}} \end{pmatrix},$$

where a $k \times n$ matrix $Y^{\mathrm{g}}$ stands for desirable (good) outputs and a $(s-k)\times n$ matrix $Y^{\mathrm{b}}$ stands for undesirable (bad) outputs. Similarly, for $\mathrm{DMU}_j$, $(j = 1, 2, \ldots, n)$, the vector $y_j$ can be decomposed into two parts as

$$y_j = \begin{pmatrix} y_j^{\mathrm{g}} \\ y_j^{\mathrm{b}} \end{pmatrix},$$

where sub-vectors $y_j^{\mathrm{g}}$ and $y_j^{\mathrm{b}}$ represent the desirable and undesirable output-values of $\mathrm{DMU}_j$, respectively. Following Seiford and Zhu [5], we multiply each undesirable output by $(-1)$ and find a proper translation vector $v$ to convert negative data to non-negative data, then

$$Y = \begin{pmatrix} Y^{\mathrm{g}} \\ Y^{\mathrm{b}'} \end{pmatrix},$$

where $y_j^{\mathrm{b}'} = -y_j^{\mathrm{b}} + v \geqslant 0$. According to [5], the output-oriented radial BCC model is as follows:

$$\begin{aligned}
\max \quad & h_0 \\
\text{s.t.} \quad & \sum_{j=1}^{n} \eta_j x_{ij} + s_i^- = x_{i0}, \quad i = 1, 2, \ldots, m, \\
& \sum_{j=1}^{n} \eta_j y_{rj}^{\mathrm{g}} - s_r^+ = h_0 y_{r0}^{\mathrm{g}}, \quad r \in G, \\
& \sum_{j=1}^{n} \eta_j y_{tj}^{\mathrm{b}'} - s_t^+ = h_0 y_{t0}^{\mathrm{b}'}, \quad t \in B, \\
& \sum_{j=1}^{n} \eta_j = 1, \\
& \eta_j, s_i^-, s_r^+, s_t^+ \geqslant 0 \quad \text{for all } j, i, r, t,
\end{aligned} \tag{1}$$

where $G = (1, 2, \ldots, k)$ is the index set of desirable outputs, and $B = (k+1, k+2, \ldots, s)$ is the index set of undesirable outputs. Model (1) implicates that a DMU has the ability to expand desirable outputs $Y^{\mathrm{g}}$ and undesirable outputs $Y^{\mathrm{b}'}$ simultaneously. A $\mathrm{DMU}_0$ is said to be efficient if $h_0^* = 1$ and all $s_i^{-*} = s_r^{+*} = s_t^{+*} = 0$. If $h_0^* > 1$ and (or) some $s_i^{-*}$, $s_r^{+*}$, or $s_t^{+*}$ is non-zero, then the $\mathrm{DMU}_0$ is inefficient.

Evidently, efficiency in model (1) is measured radially; that is, model (1) requires equiproportional augmentation in each output (desirable outputs $y_j^{\mathrm{g}}$ and transformed undesirable $y_j^{\mathrm{b}'}$). Considering non-proportional augmentation in each output, following the main idea of non-radial Russell measure, we give the following DEA-based non-radial linear programming:

$$\begin{aligned}
\max \quad & \phi_0 = \frac{1}{s}\sum_{r=1}^{k} \alpha_r + \frac{1}{s}\sum_{t=k+1}^{s} \beta_t \\
\text{s.t.} \quad & \sum_{j=1}^{n} \eta_j x_{ij} + s_i^- = x_{i0}, \quad i = 1, 2, \ldots, m, \\
& \sum_{j=1}^{n} \eta_j y_{rj}^{\mathrm{g}} = \alpha_r y_{r0}^{\mathrm{g}}, \quad r \in G, \\
& \sum_{j=1}^{n} \eta_j y_{tj}^{\mathrm{b}'} = \beta_t y_{t0}^{\mathrm{b}'}, \quad t \in B, \\
& \sum_{j=1}^{n} \eta_j = 1, \\
& \alpha_r \geqslant 1, \beta_t \geqslant 1 \quad \text{for all } r, t, \\
& \eta_j, s_i^- \geqslant 0 \quad \text{for all } j, i.
\end{aligned} \tag{2}$$

Values of $\alpha_r$, $\beta_t$ obtained from solving model (2) are employed to estimate the amount of augmentation possible in each desirable output and the amount of reduction possible in each undesirable output, respectively. If some transformed undesirable outputs $y_{t0}^{\mathrm{b}'}$ are zero, then we set $\beta_t = 1$ for those $t$ with $y_{t0}^{\mathrm{b}'} = 0$.

Since non-discretionary inputs cannot be improved, to incorporate the non-discretionary inputs into output-oriented DEA model, a non-discretionary input should be converted into a reverse non-discretionary input (please refer to [18] for the detail of reverse inputs) by applying a monotone decreasing transformation (e.g., $1/z_j$). For the purpose of preserving convexity relations [18], we adopt a linear monotone decreasing transformation, $z'_j = w - z_j \geqslant 0$ $(j = 1, 2, \ldots, n)$, where $w$ is a proper translation vector that makes $z'_j > 0$.

Under the assumption that all outputs are desirable, Banker and Morey [10] evaluated efficiency of a DMU by comparing it with those DMUs that are in the similar or more difficult environments, i.e.,

$$\sum_{j=1}^{n} \eta_j z'_{lj} \geqslant z'_{l0} \quad (l = 1, 2, \ldots, p). \tag{3}$$

In the presence of undesirable outputs, Eq. (3) is inappropriate because it considers only positive impacts of non-discretionary inputs on desirable outputs while omitting any impact of non-discretionary inputs on undesirable outputs.

As stated in [19], the basic pollution problems satisfy the null-jointness property, i.e., when we produce desirable outputs, the bad outputs are byproducts of the production process. Reducing undesirable outputs in production may not be possible without assuming certain costs [9]. If a DMU has a relatively low non-discretionary input (e.g., emission quota), it has two possible choices to satisfy the non-discretionary input constraint: reduce its discretionary input scale, or divert some of its discretionary inputs to the "cleanup" of those undesirable outputs. Both choices indicate a distortion of technical efficiency of an evaluated DMU. Therefore, efficiency of a DMU cannot be evaluated by comparing it with those DMUs that are in more favorable environments, because this kind of comparisons cannot reflect the true efficiency of the evaluated DMU.

Similarly, a DMU cannot be evaluated by comparing it with those DMUs that are in more difficult environments. We thus suggest that, to consider the impacts of non-discretionary inputs on desirable outputs and on undesirable outputs simultaneously, efficiency of a DMU be evaluated by comparing it with those DMUs that are in the similar environments, i.e.,

$$\sum_{j=1}^{n} \eta_j z'_{lj} = z'_{l0} \quad (l = 1, 2, \ldots, p). \tag{4}$$

Eq. (4) requires that reference units utilize the same levels of the transformed non-discretionary inputs as that of the assessed DMU on average.

We can find similar suggestions to Eq. (4) in the literature on defining reference set. In [20], Anderson et al. use Wilcoxin ran sum test, the median test, and the Van der Waerden non-parametric tests to investigate the impacts of operating environments on REITs' (real-estate investment trust) efficiencies. Their results show that it is not appropriate that DMUs with different environments are measured together.

Based on the reference set defined in Eq. (4), we extend model (2) to the following output-oriented DEA model:

$$\begin{aligned}
\max \quad & \rho_0 = \frac{1}{s}\sum_{r=1}^{k} \alpha_r + \frac{1}{s}\sum_{t=k+1} \beta_t \\
\text{s.t.} \quad & \sum_{j=1}^{n} \eta_j x_{ij} + s_i^- = x_{i0}, \quad i \in D, \\
& \sum_{j=1}^{n} \eta_j z'_{lj} = z'_{l0}, \quad l \in ND, \\
& \sum_{j=1}^{n} \eta_j y^{\mathrm{g}}_{rj} = \alpha_r y^{\mathrm{g}}_{r0}, \quad r \in G, \\
& \sum_{j=1}^{n} \eta_j y^{\mathrm{b}'}_{tj} = \beta_t y^{\mathrm{b}'}_{t0}, \quad t \in B, \\
& \sum_{j=1}^{n} \eta_j = 1, \\
& \alpha_r \geqslant 1, \beta_t \geqslant 1 \quad \text{for all } r, t, \\
& \eta_j, s_i^- \geqslant 0 \quad \text{for all } j, i.
\end{aligned} \tag{5}$$

For any given non-discretionary inputs, model (5) can augment each desirable output and each transformed undesirable output simultaneously with different proportions.

Let $(\alpha_r^*, \beta_t^*, s_i^{-*}, \eta^*)$ be the optimal solution to model (5). A $\mathrm{DMU}_0$ is said to be efficient if $\varphi_0^* = 1$ (e.g., $\alpha_r^* = 1$ and $\beta_t^* = 1$) and all $s_i^{-*} = 0$. If $\varphi_0^* > 1$ (e.g., any of $\alpha_r^*$ or $\beta_t^*$ is greater than one) and (or) some of $s_i^{-*}$ are non-zero, then the $\mathrm{DMU}_0$ is inefficient.

In the later case, model (5) determines the efficient target in the output-oriented sense:

$$\begin{aligned}
&\hat{x}_{i0} = x_{i0} - s_i^{-*}, \quad i = 1, 2, \ldots, m, \\
&\hat{y}^{\mathrm{g}}_{r0} = \alpha_r^* y^{\mathrm{g}}_{r0}, \quad r = 1, 2, \ldots, k, \\
&\hat{y}^{\mathrm{b}'}_{t0} = \beta_t^* y^{\mathrm{b}'}_{t0}, \quad t = k+1, k+2, \ldots, s,
\end{aligned} \tag{6}$$

in which the 'hat' represents the point on the efficient frontier. Attribute to the nature of non-discretionary input, any target for decreasing non-discretionary input may not be meaningful.

Because the optimal solution to model (5) is a feasible solution for model (2), the efficiency score of any DMU based on model (5) is less than or equal to that based on model (2).

## 4. Returns analysis

Returns-to-scale (RTS) analysis is usually defined with respect to all the inputs that the DMU uses to produce the output. In the presence of non-discretionary inputs, we substitute returns analysis for RTS. To do DMUs' returns analysis, consider the dual to model (5):

$$
\begin{aligned}
\min \quad & \zeta_0 = \sum_{i=1}^{m} u_i x_{i0} + \sum_{l=1}^{p} \lambda_l z'_{l0} + u_0 \\
& \qquad - \sum_{r=1}^{k} \omega_r - \sum_{t=k+1}^{s} \omega_t \\
\text{s.t.} \quad & \sum_{i=1}^{m} u_i x_{ij} - \sum_{r=1}^{k} \mu_r y_{rj}^{\mathrm{g}} - \sum_{t=k+1}^{s} \mu_t y_{tj}^{\mathrm{b}'} + \sum_{l=1}^{p} \lambda_l z'_{lj} \\
& \quad + u_0 \geqslant 0, \quad j = 1, 2, \ldots, n, \\
& \mu_r y_{r0}^{\mathrm{g}} - \omega_r \geqslant 1/s, \quad r = 1, 2, \ldots, k, \\
& \mu_t y_{t0}^{\mathrm{b}'} - \omega_t \geqslant 1/s, \quad t = k+1, k+2, \ldots, s, \\
& u_i, \mu_r, \mu_t, \omega_r, \omega_t \geqslant 0, \lambda_l, u_0 \\
& \text{free, for all } i, r, t, l, \qquad (7)
\end{aligned}
$$

where $u_i$, $\lambda_l$, $\mu_r$, $\mu_t$ are multipliers for discretionary inputs, non-discretionary inputs, desirable outputs and undesirable outputs, respectively.

For given non-discretionary inputs, we shall examine a DMU's returns at a point on the efficient production surface, and relate it to the sign of the intercept term $u_0$ in model (7). By comparing model (7) and the model in [21], it can be easily observed that, although non-discretionary inputs are incorporated into model (7), the approach of estimating DMUs' RTS presented in [21] is still applicable to our return analysis. If there is a unique solution to model (7), then we can follow the approach presented in [21] to estimate DMUs' returns:

(1) Increasing returns (IR) prevail at $(X_0, Z_0, Y_0)$ if and only if $u_0^* < 0$.
(2) Decreasing returns (DR) prevail at $(X_0, Z_0, Y_0)$ if and only if $u_0^* > 0$.
(3) Constant returns (CR) prevail at$(X_0, Z_0, Y_0)$ if and only if $u_0^* = 0$.

The above approach is only suitable for the BCC frontier DMUs (i.e., $\rho_0^* = 1$) and it is always possible that there are more than one optimal solution to model (7). For these general cases (i.e., a DMU is not efficient or there exist alternate optima for $u_0$) we follow the approach presented in [22] and give the following modification to characterize returns on $\mathrm{DMU}_0$ with $u_0^* > 0$ in model (7), i.e.,

$$
\begin{aligned}
\min \quad & \hat{u}_0 \\
\text{s.t.} \quad & \sum_{i=1}^{m} u_i x_{i0} + \sum_{l=1}^{p} \lambda_l z'_{l0} + \hat{u}_0 - \sum_{r=1}^{k} \omega_r \\
& - \sum_{t=k+1}^{s} \omega_t = \rho_0^*, \\
& \sum_{i=1}^{m} u_i x_{ij} - \sum_{r=1}^{k} \mu_r y_{rj}^{\mathrm{g}} - \sum_{t=k+1}^{s} \mu_t y_{tj}^{\mathrm{b}'} + \sum_{l=1}^{p} \lambda_l z'_{lj} \\
& + \hat{u}_0 \geqslant 0, \quad j = 1, 2, \ldots, n, \\
& \mu_r y_{r0}^{\mathrm{g}} - \omega_r \geqslant 1/s, \quad r = 1, 2, \ldots, k, \\
& \mu_t y_{t0}^{\mathrm{b}'} - \omega_t \geqslant 1/s, \quad t = k+1, k+2, \ldots, s, \\
& u_i, \mu_r, \mu_t, \hat{u}_0, \omega_r, \omega_t \geqslant 0, \lambda_l \\
& \text{free, for all } i, r, t, l. \qquad (8)
\end{aligned}
$$

In the case of solution of the unrestricted variable in model (7) yielding $u_0^* < 0$, we replace $\hat{u}_0 \geqslant 0$ in model (8) by $\hat{u}_0 \leqslant 0$ and change the objective of model (8) from minimization to maximization. Then we have the following proposition:

**Proposition 1.** (1) *Given the existence of an optimal solution with* $u_0^* > 0$ *in model* (7), *the returns at* $\mathrm{DMU}_0$ *are CR if and only if the optimal value which model* (8) *achieves is zero, i.e.,* $\hat{u}_0^* = 0$, *and DR if and only if* $\hat{u}_0^* > 0$; (2) *given the existence of an optimal solution with* $u_0^* < 0$ *in model* (7), *the returns at* $\mathrm{DMU}_0$ *are CR if and only if the optimal value which model* (8) *achieves is zero, i.e.,* $\hat{u}_0^* = 0$, *and IR if and only if* $\hat{u}_0^* < 0$.

The proof of Proposition 1 is similar to that of Theorem 7 in [22] and thus omitted.

## 5. Eco-efficiency analyses of paper mills

In this section, we illustrate the proposed model using real data set of 32 paper mills along the Huai River in Anhui Province, China. Each paper mill uses discretionary inputs (labor and capital) to produce paper products while emitting BOD (Pollutants generated by paper mills include BOD, sulfur oxides, particulates and suspended solids. We in this paper take BOD as the main pollutant considering the practical situation of the Huai River). We collect total amount of waste water that

Table 1
Characteristics of the data set of 32 paper mills (2001)[a,b]

| Characteristics | Inputs (D) | | Input (ND) | Outputs (G) | Output (B) |
|---|---|---|---|---|---|
| | Labor | Capital | BOD-Q (kg) | Paper (ton) | BOD (kg) |
| Max | 1090 | 5902 | 33,204 | 29,881 | 28,487.7 |
| Min | 122 | 1368.5 | 1504.2 | 5186.6 | 1453.3 |
| Mean | 630 | 3534.1 | 13,898.6 | 17,054 | 10,698.5 |
| Std. dev | 326 | 1222.4 | 8782 | 7703.3 | 7969.2 |

[a] *Data source*: Anhui Environmental Protection Bureau, the Fuyang Environmental Protection Bureau, the Huainan Environmental Protection Bureau.
[b] *Note*: Capital are stated in units of 10,000 RMB Yuan. Labor is expressed in units of 1 person.

flowed out from each mill into the River in 2001, and multiply the amount by its average BOD concentration of that year to obtain the real emission amount of BOD. Since the main pollutant of paper mills is BOD, we only collect data of emission quotas of BOD (BOD-Q) allocated for the 32 paper mills.

In this study, emission quota of BOD is viewed as a non-discretionary input. This is because, firstly, emission quota of BOD is an administrative constraint, which is an external factor and cannot be controlled by managers of paper mills; secondly, it has substantial impacts on both desirable output (paper) and undesirable output (BOD) of each paper mill.

We collect these data from the Anhui Environmental Protection Bureau, the Fuyang Environmental Protection Bureau, and the Huainan Environmental Protection Bureau. The characteristics of the data set are summarized in Table 1. The detailed data set for the 32 paper mills is available from the authors.

By applying our model, we first estimate efficiencies and desirable/undesirable outputs targets for the 32 paper mills. To elaborate the rationality of the proposed model, we then compare our approach with the BM model [10] and the Ruggiero model [11]. Results of returns analyses are illustrated at last.

Set the translation vectors of $v$=50,000 and $w$=60,000 for undesirable output (BOD) and non-discretionary input (BOD-Q). Table 2 gives the efficiency results obtained from model (1), model (2), and model (5) for all paper mills.

Note in Table 2 that, when we ignore the non-discretionary input, for BCC models, 14 paper mills are deemed as efficient under both model (1) and model (2); for CCR models, there are only 9 paper mills are evaluated as efficient. However, we have only 11BCC-inefficient paper mills under model (5), namely, paper mills 4, 5, 6, 14, 15, 18, 19, 22, 24, 26, 27. For CCR model (5), we also have 17 efficient paper mills, which is larger than that obtained from model (1) or model (2). These results show that, when undesirable outputs are considered in performance evaluation, if the impacts of non-discretionary inputs on DMUs' efficiencies are not dealt with properly, the ranking of DMUs' performance may be severely distorted. This conclusion can also be verified by comparing mean efficiencies under different models, i.e., mean efficiency under model (5) is always lower than that under model (2).

To further analyze the impacts of non-discretionary inputs on DMUs' efficiencies, we take mills 14 and 29 as examples to compare their efficient targets under different models (model (2) and model (5)). Relative results are reported in Table 3.

It can be observed from Table 3 that, output-oriented efficient targets for mill 14 are 11,077.6 ton of paper and 4292 kg of BOD under CCR model (5); while these targets are 12,488.5 ton of paper and 4533.4 kg of BOD under CCR model (2). Because the current BOD output of mill 14 is 4533.4 kg, this implies that, although CCR model (2) suggests that mill 14 need not reduce BOD, CCR model (5) suggests that mill 14 should reduce its undesirable output of BOD. Results of mill 29 in Table 3 also confirm this observation.

To illustrate the rationality of the developed model, we compare our model with the BM and the Ruggiero models (We transform the original input-oriented models into output-oriented models in our comparison). The characteristics of BCC measures under different models are reported in Table 4.

It can be found in Table 4 that, there are 14 efficient paper mills under the BM model; while under the Ruggiero model, only 11 paper mills are deemed as BCC-efficient and there exist 4 infeasible solutions (mills 12, 20, 23 and 25). These results indicate that the BM model and the Ruggiero should be extended to consider impacts of non-discretionary inputs on DMUs' undesirable outputs.

Table 2
Efficiency results[a]

| Paper mills | Model (1) | | Model (2) | | Model (5) | |
|---|---|---|---|---|---|---|
| | BCC | CCR | BCC | CCR | BCC | CCR |
| 1 | 1.0000 | 1.0170 | 1.0000 | 1.0937 | 1.0000 | 1.0000 |
| 2 | 1.0000 | 1.0077 | 1.0000 | 1.5756 | 1.0000 | 1.0000 |
| 3 | 1.2294 | 1.9995 | 1.4028 | 2.0164 | 1.0000 | 1.1350 |
| 4 | 1.0755 | 1.9111 | 1.0814 | 2.1498 | 1.0569 | 1.4232 |
| 5 | 1.1586 | 1.7184 | 1.1644 | 1.9030 | 1.1437 | 1.4051 |
| 6 | 1.0719 | 2.3279 | 1.9412 | 3.7820 | 1.5951 | 2.0566 |
| 7 | 1.0092 | 1.5918 | 1.1063 | 1.6010 | 1.0000 | 1.0000 |
| 8 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 9 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 10 | 1.1099 | 1.1997 | 1.2790 | 1.3106 | 1.0000 | 1.0000 |
| 11 | 1.0536 | 1.9517 | 1.5171 | 2.3394 | 1.0000 | 1.0000 |
| 12 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 13 | 1.0210 | 1.6884 | 1.0431 | 1.7240 | 1.0000 | 1.2631 |
| 14 | 1.0033 | 1.0107 | 1.1041 | 1.2684 | 1.0860 | 1.1842 |
| 15 | 1.0592 | 1.1666 | 1.2754 | 1.3866 | 1.2672 | 1.3539 |
| 16 | 1.0785 | 1.2284 | 1.4318 | 1.4541 | 1.0000 | 1.0000 |
| 17 | 1.0000 | 1.4750 | 1.0000 | 2.9729 | 1.0000 | 1.0000 |
| 18 | 1.2300 | 1.5689 | 1.2318 | 1.7596 | 1.2071 | 1.3715 |
| 19 | 1.0768 | 1.4850 | 1.4529 | 1.7789 | 1.3086 | 1.6205 |
| 20 | 1.0000 | 1.5908 | 1.0000 | 3.3802 | 1.0000 | 3.3801 |
| 21 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 22 | 1.2778 | 1.5259 | 1.4731 | 2.4185 | 1.1647 | 1.2027 |
| 23 | 1.0000 | 1.0468 | 1.0000 | 1.6596 | 1.0000 | 1.3468 |
| 24 | 1.0560 | 1.1009 | 1.1157 | 1.7374 | 1.0472 | 1.0598 |
| 25 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 26 | 1.1363 | 2.2353 | 1.2321 | 2.2857 | 1.2294 | 1.7068 |
| 27 | 1.0397 | 1.2132 | 1.3739 | 1.5599 | 1.3606 | 1.5293 |
| 28 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 29 | 1.1334 | 1.1380 | 1.1365 | 1.1481 | 1.0000 | 1.0000 |
| 30 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 31 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 32 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| Mean | 1.0569 | 1.3500 | 1.1676 | 1.6658 | 1.0771 | 1.2512 |
| # of efficient DMUs | 14 | 9 | 14 | 9 | 21 | 17 |

[a]All calculations are by the authors.

Table 3
Targets for mill 14 and mill 29 to eliminate inefficiencies

| Mill | | Current level | Model (2) | | Model (5) | |
|---|---|---|---|---|---|---|
| | | | BCC | CCR | BCC | CCR |
| 14 | Paper | 8126.8 | 9818.8 | 12,488.5 | 9523.8 | 11,077.6 |
| | BOD | 4533.4 | 4533.4 | 4533.4 | 4533.4 | 4292 |
| | Labor | 671 | 671 | 671 | 671 | 671 |
| | Capital | 1789.7 | 1789.7 | 1789.7 | 1789.7 | 1789.7 |
| 29 | Paper | 15,151 | 16,869.1 | 19,637.2 | 15,151 | 15,151 |
| | BOD | 10,613 | 4326.8 | 10,613 | 10,613 | 10,613 |
| | Labor | 166 | 166 | 166 | 166 | 166 |
| | Capital | 3968.1 | 3968.1 | 3968.1 | 3968.1 | 3968.1 |

Table 4
Characteristics of BCC measures under different models

| | BM | Ruggiero | Model (5) |
|---|---|---|---|
| Infeasible | 0 | 4 | 0 |
| Mean | 1.1669 | 1.1780 | 1.0771 |
| Efficient | 14 | 11 | 21 |

Table 5
Results of DMUs' returns under different models

| Mill | Model (1) | Model (2) | Model (5) |
|---|---|---|---|
| 1 | DR | DR | CR |
| 3 | DR | DR | IR |
| 10 | DR | DR | CR |
| 14 | DR | DR | DR |
| 29 | IR | IR | CR |

It is interesting to analyze the impacts of the non-discretionary input on DMUs' returns under different models. By taking mills 1, 3, 10, 14 and 29 as examples, we list results of returns estimation under different models in Table 5.

It can be found from Table 5 that returns estimations under model (1) are the same as those under model (2), but returns estimations under model (5) are different from those under model (2) or model (1). For example, DR prevails for mill 10 under model (1) and model (2); while under model (5), the mill is operating at the most productive scale size for the discretionary inputs. These results suggest that environment protection policies, e.g., BOD emission quota, have significant effects on DMUs' productions.

Detailed estimation of returns of all paper mills under model (5) show that, 15 of the 32 mills are operating at scales smaller than the most productive scale size (DR) for discretionary inputs. Therefore, for nearly half of the paper mills along the Huai River, a prevailing policy for them is to reduce the discretionary inputs to eliminate technical inefficiencies.

## 6. Conclusions

This paper addresses the ecological efficiency evaluation problem for paper mills along the Huai River in China. The characteristic of the ecological efficiency evaluation problem is that undesirable outputs and non-discretionary inputs should be considered simultaneously. Since current efficiency evaluation approaches are not directly applicable to the problem, a DEA-based non-radial model is developed. In the proposed model, we describes a new approach of defining reference set that requires reference units operate in a similar environment on average. Using real data of 32 paper mills, an empirical study is employed to illustrate the impacts of non-discretionary inputs on DMUs' eco-efficiencies.

Possible extension of this research is to investigate non-discretionary input allocation mechanism. This will gain deeper insights into the causes of eco-inefficiency, and give further supports on environmental protection policy for the Huai River basin.

## Acknowledgements

The authors would like to thank the anonymous referees for their helpful comments and suggestions, which significantly improved the paper. This research was supported by the Program for New Century Excellent Talents in the University of China (Grant no. NCET-04-0570), and the national Natural Science Foundation of China (Grant No. 70525001).

## References

[1] Total amount control. http://www.zhb.gov.cn/english/
[2] China papermaking statistics yearbook (1999). Beijing: China Light Industry Press; 2000.
[3] Färe R, Grosskopf S, Lovell CAK, Pasurka C. Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. The Review of Economics and Statistics 1989;71:90–8.
[4] Scheel H. Undesirable outputs in efficiency valuations. European Journal of Operational Research 2001;132:400–10.
[5] Seiford LM, Zhu J. Modeling undesirable factors in efficiency evaluation. European Journal of Operational Research 2002;142:16–20.
[6] Vencheh AH, Matin RK, Kajani MT. Undesirable factors in efficiency measurement. Applied Mathematics and Computation 2005;163:547–52.
[7] Sarkis J, Talluri S. Ecoefficiency measurement using data envelopment analysis: research and practitioner issues. Journal of Environmental Assessment Policy and Management 2004;6(1):91–123.
[8] Korhonen PJ, Luptacik M. Eco-efficiency analysis of power plants: an extension of data envelopment analysis. European Journal of Operational Research 2004;154:437–46.
[9] Färe R, Grosskopf S. Modeling undesirable factors in efficiency evaluation: comment. European Journal of Operational Research 2004;157:242–5.
[10] Banker RD, Morey R. Efficiency analysis for exogenously fixed inputs and outputs. Operations Research 1986;34:513–21.
[11] Ruggiero J. On the measurement of technical efficiency in the public sector. European Journal of Operational Research 1996;90:553–65.
[12] Ruggiero J. Non-discretionary inputs in data envelopment analysis. European Journal of Operational Research 1998; 111:461–9.

[13] Ruggiero J. Performance evaluation when non-discretionary factors correlate with technical efficiency. European Journal of Operational Research 2004;159:250–7.
[14] Ray SC. Resource use efficiency in public schools: a study of Connecticut data. Management Science 1991;37:1620–8.
[15] Charnes A. Cooper WW, Lewin A, Morey RC, Rousseau J. Efficiency analysis with non-discretionary resources. Research report CCS 379, University of Texas, Austin, 1980.
[16] Muñiz MA. Separating managerial inefficiency and external conditions in data envelopment analysis. European Journal of Operational Research 2002;143:625–43.
[17] Färe R, Lovell CAK. Measuring the technical efficiency of production. Journal of Economic Theory 1978;19: 150–62.
[18] Lewis HF, Sexton TR, Data envelopment analysis with reverse inputs, Paper presented at North America productivity workshop, Union College, Schenectady, NY, July 1999.
[19] Färe R, Primont D. Multi-output production and duality: theory and applications. Boston: Kluwer Academic Publishers; 1995.
[20] Anderson RI, Fok R, Springer T, Webb J. Technical efficiency and economies of scale: a non-parametric analysis of REIT operating efficiency. European Journal of Operational Research 2002;139:598–612.
[21] Banker RD, Charnes A, Cooper WW. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management Science 1984;30:1078–92.
[22] Seiford LM, Zhu J. An investigation of returns to scale in data envelopment analysis. Omega 1999;27:1–11.