RAIRO-Oper. Res. 58 (2024) 3189–3202
https://doi.org/10.1051/ro/2024119

RAIRO Operations Research
www.rairo-ro.org

A BOOTSTRAP DATA ENVELOPMENT ANALYSIS MODEL WITH STOCHASTIC REDUCIBLE OUTPUTS AND EXPANDABLE INPUTS: AN APPLICATION TO POWER PLANTS

Alireza Amirteimoori[1,*], Tofigh Allahviranloo[1] and Asunur Cezar[2]

Abstract. Clean production of electricity is not only cost-effective but also effective in reducing pollutants. Toward this end, the use of clean fuels is strongly recommended by environmentalists. Benchmarking techniques, especially data envelopment analysis, are an appropriate tool for measuring the relative efficiency of firms with environmental pollutants. In classic data envelopment analysis models, decision-makers are faced with production processes in which reducible inputs are used to produce expandable outputs. In this contribution, we consider production processes when the input and output data are given in stochastic form and some throughputs are reducible and some others are expandable. A stochastic directional distance function model is proposed to calculate the relative technical efficiency of firms. In order to evaluate firm-specific technical efficiency, we apply bootstrap DEA. We first calculate the technical efficiency scores of firms using the classic DEA model. Then, the double bootstrap DEA model is applied to determine the impact of explanatory variables on firm efficiency. To demonstrate the applicability of the procedure, we present an empirical application wherein we employ power plants.

Mathematics Subject Classification. 90B30.

Received November 1, 2023. Accepted May 26, 2024.

1. Introduction

Despite the fact that electricity is classified as a clean fuel, its production causes environmental pollutants. Over the last three decades, the rapid growth of pollutant production by refineries and other energy-producing plants has caused concern among environmentalists. A major concern of people around the world, especially environmentalists, is reducing the level of production of undesirable outputs by industrial plants. Although a lot of research has been conducted on reducing undesirable outputs by industrial countries, pollutant production is still one of the major concerns of countries. Recently, important steps have been taken by researchers to increase productivity and reduce pollutants; however, these studies have not yet been implemented. Benchmarking is one of the most frequently used approaches in performance evaluation and has attracted considerable attention among researchers. In the last two decades, data envelopment analysis (DEA), as a powerful benchmarking tool,

Keywords. Stochastic DEA, double bootstrap, undesirable outputs, technical efficiency, pollutants, power plant.

[1] Faculty of Engineering & Natural Sciences, Istinye University, Istanbul, Turkey.

[2] Department of Management, Boğaziçi University, Bebek, Istanbul 34342, Turkey.

[*]Corresponding author: alireza.amirteimoori@istinye.edu.tr

has frequently been used in the performance evaluation and target setting of a set of homogeneous production units. After the original works of Charnes *et al.* [9] and Banker *et al.* [5], there have been a lot of developments and applications in the literature (see, for instance [2, 3, 6, 7, 14, 29, 33, 36, 43], and many more). An important feature of DEA is its extension to stochastic environments. To consider random variations of inputs and outputs, several authors have considered various approaches in the DEA framework to calculate technical and allocative efficiencies (see, for instance [12, 13, 19, 20, 23, 26–28, 35, 38, 41]).

One of the most frequently studied subjects in the framework of DEA is performance evaluation in the presence of undesirable products. Two important assumptions in handling undesirable outputs in the production process are weak disposability and managerial disposability assumptions. When we use the weak disposability assumption, we actually reduce the level of undesirable outputs by reducing the level of activity, while in the latter, undesirable outputs are reduced by increasing some or all inputs. Kuosmanen [22] made an extension of the weak disposable technology set of Färe and Grosskopf [15] to develop a DEA model to evaluate the relative efficiency of firms in the presence of undesirable outputs. Sueyoshi and Goto [39] have used a non-radial DEA model and strong complementary slackness conditions to unify technical efficiency under managerial and natural disposability assumptions. Sueyoshi *et al.* [40] have used the DEA technique for environmental assessment, employing both, natural and managerial disposability assumptions. Zhou *et al.* [44] have used the transformation technique to handle undesirable outputs for calculating environmental efficiency in the DEA framework. Al-Mezeini *et al.* [1] have used double bootstrap two-stage DEA to investigate the efficiency of greenhouse production in Oman. Chen and Liu [10] have used slack-based super-efficiency with undesirable outputs to improve eco-efficiency in coal mining for sustainable development. Qu *et al.* [31] have used the DEA technique with undesirable outputs to calculate the green factor productivity of industrial enterprises in Zhejiang province of China from 2012 to 2017 (see also [8, 17, 25, 32, 34, 42], and many more).

All the above-mentioned studies assume that the outputs are partitioned into two groups: desirable and undesirable. In such cases, desirable outputs are expanded and inputs and undesirable outputs are reduced to improve the productivity of the process. However, in some real applications, there are also input variables that need to be maximized. In this case, the input set is divided into two groups: the first group of inputs needs to be reduced, and the second group needs to be increased. Some researchers have referred to these inputs as undesirable inputs. Hua and Bian [18] presented a review of DEA models with undesirable factors (inputs and outputs). DEA models with undesirable inputs and outputs are studied by Liu *et al.* [24]. Kordrostami *et al.* [21] studied the problem of the weak disposability assumption of inputs and outputs in a nonparametric production analysis.

While production processes with undesirable outputs have been extensively studied, there have been few studies about DEA with undesirable inputs. In this sense, there is a gap in evaluating the performance of firms with expandable inputs and reducible outputs in the presence of explanatory variables.

To fill this gap, in this paper, we consider a production process where the inputs and outputs are divided into two groups: inputs and outputs that need to be decreased and those inputs and outputs that need to be increased. To the best of our knowledge, there is no DEA-based research work dealing with the existence of such inputs and outputs in a stochastic environment. Therefore, in this contribution, we focus on the evaluation of technical efficiency when inputs and outputs are random variables, and some inputs and outputs need to be reduced while others must be increased. We first extend the traditional weak disposability assumption [37] to incorporate such inputs and outputs in a deterministic environment. A directional distance function model is proposed to calculate the technical efficiency of firms. Then, the obtained deterministic models are extended to a stochastic environment. In order to estimate the firm-specific technical efficiency, we apply bootstrap DEA to estimate the bias and bias-corrected efficiency scores. We first calculate the technical efficiency of the DMUs using our proposed model. Then, the double bootstrap DEA technique is used to determine the impact of explanatory variables on firm efficiency. To demonstrate the applicability of our proposed approach, we present a real application wherein we employ Iranian power plants.

The next sections of this paper proceed as follows: Section 2 discusses the extended weak disposability assumption in a deterministic environment. In Section 3, we set up the stochastic and deterministic versions of our proposed efficiency model. Section 4 presents a real application using data from 20 power plants in Iran. Conclusions appear in Section 5.

2. Methodology

Suppose we have n firms, and each firm uses t inputs to generate S desirable outputs and K undesirable outputs. We assume that the inputs are divided into two groups: the first group of inputs must be reduced, and the second group must be increased. The first group is referred to as reducible inputs, and the second group of inputs is referred to as expandable inputs. We aim to reduce reducible inputs and undesirable outputs while simultaneously increasing expandable inputs and desirable outputs. Let DMU_j be a firm that uses reducible inputs $x_j^R = (x_{1j}^R, \dots, x_{Mj}^R)^T$ and expandable inputs $x_j^E = (x_{1j}^E, \dots, x_{Lj}^E)^T$ to generate the desirable outputs $v_j = (v_{1j}, \dots, v_{Sj})^T$ and the undesirable outputs $w_j = (w_{1j}, \dots, w_{Kj})^T$. Formally and generally, the technology set Y is defined as follows:

$$Y = \left\{ \left(x^R, x^E, v, w\right) : \left(x^R, x^E\right) \text{ can produce } (v, w) \right\}.$$

The weak disposability assumption is used to reduce the level of undesirable outputs by reducing the levels of activities of firms. When the input set consists of reducible and expandable inputs, when we reduce the level of activity, in addition to desirable and undesirable outputs, we must also reduce the level of expandable inputs. In this case, the traditional weak disposability assumption of Shephard [37] is re-written as follows:

Definition 1 (Extended weak disposability)**.** Desirable and undesirable outputs satisfy the weak disposability assumption if

$$\left(x^R, x^E, v, w\right) \in Y \Rightarrow 0 \le \left(x^R, \theta x^E, \theta v, \theta w\right) \in Y, \qquad \forall \theta \ge 0.$$

The extended weak disposability assumption states that if the outputs (v, w) can be produced by inputs $\left(x^R, x^E\right)$, then undesirable outputs can be reduced by a scaling factor θ by reducing the desirable outputs and expandable inputs by a single factor θ.

The technology set Y in a variable return to scale environment is defined as follows:

$$\begin{aligned} Y = \Bigg\{ &\left(x^R, x^E, v, w\right) : \sum_{j=1}^n \lambda_j x_j^R \le x^R, \sum_{j=1}^n \lambda_j \theta_j x_j^E \ge x^E, \sum_{j=1}^n \lambda_j \theta_j v_j \ge v, \\ &\sum_{j=1}^n \lambda_j \theta_j w_j = w, \sum_{j=1}^n \lambda_j = 1,\ \lambda_j \ge 0,\ 0 \le \theta_j \le 1,\ j = 1, \dots, n \Bigg\}. \end{aligned} \tag{1}$$

Clearly, the technology set Y in equation (1) is in a non-linear format. Let $\lambda_j \theta_j = \mu_j$ and $\lambda_j (1 - \theta_j) = \rho_j$, then we can rewrite λ_j as $\lambda_j = \mu_j + \rho_j$, where $\mu_j \ge 0$ and $\rho_j \ge 0$. By making these changes of variables, the technology set (1) can be transformed into the following linear format:

$$\begin{aligned} P_x = \Bigg\{ &\left(x^R, x^E, v, w\right) : \sum_{j=1}^n (\mu_j + \rho_j) x_j^R \le x^R, \sum_{j=1}^n \mu_j x_j^E \ge x^E, \sum_{j=1}^n \mu_j v_j \ge v, \\ &\sum_{j=1}^n \mu_j w_j = w, \sum_{j=1}^n (\mu_j + \rho_j) = 1,\ \mu_j,\ \rho_j \ge 0,\ j = 1, \dots, n \Bigg\}. \end{aligned} \tag{2}$$

Now, we can easily use Farrell's [16] output- or input-oriented efficiency models to analyze the technical efficiency of firms. In this study, we focus on measuring the firm's efficiency with a focus on revenue generation while reducing undesirable outputs. To achieve this, we use the directional distance function approach. In this sense, we solve the following LP model:

$$
\begin{aligned}
&\beta_o^* = \text{Max } \beta \\
&\text{s.t.} \\
&\qquad \sum_{j=1}^{n} (\mu_j + \rho_j) x_{mj}^R \le x_{mo}^R, && m = 1, \ldots, M, \\
&\qquad \sum_{j=1}^{n} \mu_j x_{lj}^E \ge x_{lo}^E, && l = 1, \ldots, L, \\
&\qquad \sum_{j=1}^{n} \mu_j v_{sj} \ge v_{so} + \beta d_{so}^v, && s = 1, \ldots, S, \\
&\qquad \sum_{j=1}^{n} \mu_j w_{kj} = w_{ko} - \beta d_{ko}^w, && k = 1, \ldots, K, \\
&\qquad \sum_{j=1}^{n} (\mu_j + \rho_j) = 1, \\
&\qquad \mu_j, \rho_j \ge 0, && j = 1, \ldots, J
\end{aligned} \tag{3}
$$

where (d_o^v, d_o^v) is a user-defined direction in which the desirable outputs are increased and simultaneously, the undesirable outputs are reduced. In Model (3), $\mu_j = 0, j = 1, \ldots, n, j \neq o, \mu_o = 1, \rho_j = 0, j = 1, \ldots, n, \beta = 0$ is a feasible solution and clearly β is bounded. These guarantee the feasibility and boundedness of Model (3). The relative efficiency of DMU_o is defined as follows:

$$
e_o^* = \frac{1 - \frac{\beta_o^*}{\sum_{j=1}^{n} \beta_j^*}}{1 + \frac{\beta_o^*}{\sum_{j=1}^{n} \beta_j^*}}\cdot
$$

Clearly, $0 \le e_o^* \le 1$, and DMU_o is said to be efficient if $e_o^* = 1$.

If we are given the prices of the desirable outputs of firm j as $u_j = (u_{1j}, \ldots, u_{Sj})$, we can calculate the revenue-based efficiency of firm j. In order to generate the output vector v_o by consuming (x_o^R, x_o^E) by firm o, its revenue-based efficiency measure is calculated as follows:

$$
\begin{aligned}
&\beta_o^* = \text{Max } \beta \\
&\text{s.t.} \\
&\qquad \sum_{j=1}^{n} (\mu_j + \rho_j) x_{mj}^R \le x_{mo}^R, && m = 1, \ldots, M, \\
&\qquad \sum_{j=1}^{n} \mu_j x_{lj}^E \ge x_{lo}^E, && l = 1, \ldots, L, \\
&\qquad \sum_{j=1}^{n} \mu_j \left(\sum_{s=1}^{S} u_{sj} v_{sj} \right) \ge \left(\sum_{s=1}^{S} u_{so} v_{so} \right) + \beta d_o^v, && s = 1, \ldots, S, \\
&\qquad \sum_{j=1}^{n} \mu_j w_{kj} = w_{ko} - \beta d_{ko}^w, && k = 1, \ldots, K, \\
&\qquad \sum_{j=1}^{n} (\mu_j + \rho_j) = 1,
\end{aligned}
$$

$$\mu_j,\ \rho_j \geq 0, \qquad j = 1,\dots,J. \tag{4}$$

The LP model (4) is used to calculate efficiency for every unit under consideration. It is easy to show that the production set in Model (4) is a convex set, and its frontier is a piecewise linear and concave production function. Regarding the validity of our proposed model, Banker [4] extended a statistical foundation for DEA models and claimed that all DEA estimators are not only consistent, but they also maximize likelihood. Banker [4] also showed that DEA estimators of the best practice are monotone increasing, and concave production functions are maximum likelihood estimators.

3. Extension to stochastic environment

Suppose $\tilde{X}_j^R = (\tilde{x}_{1j}^R,\dots,\tilde{x}_{Mj}^R)^T$ and $\tilde{X}_j^E = (\tilde{x}_{1j}^E,\dots,\tilde{x}_{Lj}^E)^T$ represent the random reducible and expandable input vectors that are used to generate the random desirable outputs $\tilde{V}_j = (\tilde{v}_{1j},\dots,\tilde{v}_{Sj})^T$and random undesirable outputs $\tilde{W}_j = (\tilde{w}_{1j},\dots,\tilde{w}_{Kj})^T$. An important assumption is that the input/output data follows normal distribution with known mean and variance. Let $X_j^R = (\tilde{x}_{1j}^R,\dots,\tilde{x}_{Mj}^R)^T$, $X_j^E = (\tilde{x}_{1j}^E,\dots,\tilde{x}_{Lj}^E)^T$, $V_j = (\tilde{v}_{1j},\dots,\tilde{v}_{Sj})^T$ and $W_j = (\tilde{w}_{1j},\dots,\tilde{w}_{Kj})^T$ represent their associated means. We propose the following program to evaluate the revenue-based efficiency of a specific DMU_o:

$$\begin{aligned}
& e_o^* = \text{Max } \beta \\
& \text{s.t.} \\
& \quad \mathbb{P}\left\{\sum_{j=1}^{n}(\mu_j+\rho_j)\tilde{x}_{mj}^R \leq \tilde{x}_{mo}^R\right\} \geq 1-\alpha, && m=1,\dots,M,\\
& \quad \mathbb{P}\left\{\sum_{j=1}^{n}\mu_j\tilde{x}_{lj}^E \geq \tilde{x}_{lo}^E\right\} \geq 1-\alpha, && l=1,\dots,L,\\
& \quad \mathbb{P}\left\{\sum_{j=1}^{n}\mu_j\left(\sum_{s=1}^{S}u_{sj}\tilde{v}_{sj}\right) \geq \left(\sum_{s=1}^{S}u_{so}\tilde{v}_{so}\right)+\beta d_o^v\right\} \geq 1-\alpha, && s=1,\dots,S,\\
& \quad \mathbb{P}\left\{-\epsilon \leq \sum_{j=1}^{n}\mu_j\tilde{w}_{kj} - (\tilde{w}_{ko}-\beta d_{ko}^w) \leq \epsilon\right\} \geq 1-\alpha, && k=1,\dots,K,\\
& \quad \sum_{j=1}^{n}(\mu_j+\rho_j) = 1,\\
& \quad \mu_j,\ \rho_j \geq 0.
\end{aligned} \tag{5}$$

Applying the single error structure proposed by Cooper *et al.* [11], the linear format of Model (6) is given by:

$$\begin{aligned}
& \beta_o^* = \text{Max } \beta \\
& \text{s.t.} \\
& \quad -\Phi^{-1}\{\gamma\}\left(\sum_{j=1}^{n}(\mu_j+\rho_j)a_{mj}^R - a_{mo}^R\right) \leq \left(x_{mo}^R - \sum_{j=1}^{n}(\mu_j+\rho_j)x_{mj}^R\right), && \forall m,\\
& \quad -\Phi^{-1}\{\gamma\}\left(\sum_{j=1}^{n}(\mu_j+\rho_j)a_{mj}^R - a_{mo}^R\right) \geq -\left(x_{mo}^R - \sum_{j=1}^{n}(\mu_j+\rho_j)x_{mj}^R\right), && \forall m,
\end{aligned}$$

$$
\begin{aligned}
& -\Phi^{-1}\{\gamma\}\left(\sum_{j=1}^{n}\mu_j a_{lj}^E - a_{lo}^E\right) \leq \left(\sum_{j=1}^{n}\mu_j x_{lj}^E - x_{lo}^E\right), && \forall l,\\
& -\Phi^{-1}\{\gamma\}\left(\sum_{j=1}^{n}\mu_j a_{lj}^E - a_{lo}^E\right) \geq -\left(\sum_{j=1}^{n}\mu_j x_{lj}^E - x_{lo}^E\right), && \forall l,\\
& -\Phi^{-1}\{\gamma\}\left(-\sum_{j=1}^{n}\mu_j d_{sj}+d_{so}\right) \leq \left(\sum_{j=1}^{n}\mu_j\left(\sum_{s=1}^{S}u_{sj}v_{sj}\right) - \left(\sum_{s=1}^{S}u_{so}v_{so}\right) + \beta d_{so}^v\right), && \forall s,\\
& -\Phi^{-1}\{\gamma\}\left(-\sum_{j=1}^{n}\mu_j d_{sj}+d_{so}\right) \geq -\left(\sum_{j=1}^{n}\mu_j\left(\sum_{s=1}^{S}u_{sj}v_{sj}\right) + \left(\sum_{s=1}^{S}u_{so}v_{so}\right) + \beta d_{so}^v\right), && \forall s,\\
& -\Phi^{-1}\left(\frac{\gamma}{2}\right)\left(\sum_{j=1}^{n}\mu_j c_{kj} - c_{ko}\right) \leq \left(\sum_{j=1}^{n}\mu_j w_{kj} - w_{ko} + \beta d_{ko}^w + \epsilon\right), && \forall k,\\
& -\Phi^{-1}\left(\frac{\gamma}{2}\right)\left(\sum_{j=1}^{n}\mu_j c_{kj} - c_{ko}\right) \geq -\left(\sum_{j=1}^{n}\mu_j w_{kj} - w_{ko} + \beta d_{ko}^w + \epsilon\right), && \forall k,\\
& \sum_{j=1}^{n}(\mu_j+\rho_j) = 1, &&\\
& \mu_j,\ \rho_j \geq 0 && \text{for all } j. \quad (6)
\end{aligned}
$$

Again, the technical efficiency of DMU_o is defined as follows:

$$
e_o^* = \frac{1 - \frac{\beta_o^*}{\sum_{j=1}^{n}\beta_j^*}}{1 + \frac{\beta_o^*}{\sum_{j=1}^{n}\beta_j^*}}.
$$

Clearly, e_o^* is well-defined and $0 \leq e_o^* \leq 1$, and DMU_o is said to be efficient if $e_o^* = 1$.

Now, suppose that in addition to firm-specific inputs and outputs, there are N explanatory factors $\tilde{Z}_j = (\tilde{z}_{1j}, \ldots, \tilde{z}_{Nj})^T$ that affect the efficiency of the firms. Let $Z_j = (z_{1j}, \ldots, z_{Nj})^T$ represent its associated mean. In the second stage, we examine the impact of explanatory factors that affect the technical efficiency. Toward this end, we use the following regression model to account for the impact of such variables on the calculated technical efficiency scores:

$$
\mathrm{Log}(e_o^*) = a_0 + a_1 z_1 + a_2 z_2 + \ldots + a_N z_N + U \quad (7)
$$

in which $\mathrm{Log}(e_o^*)$ is the logarithm of e_o^*, and $z_1, z_2, \ldots, z_N$ are the above-mentioned contextual variables. U is the error term and it is a truncated normal random variable that is distributed normally with zero mean and σ^2 variance.

Simar and Wilson [38] suggested a two-stage double bootstrap DEA procedure to obtain a more accurate efficiency score. In this paper, in the first stage, a parametric bootstrap method on Model (6) is used to calculate the bias-corrected efficiency scores. Then, in the second stage, the regression Model (7) is used to calculate the effect of explanatory variables on bias-corrected efficiencies obtained from stage 1. We will use the second approach of Simar and Wilson [38] with 2000 bootstrap estimates. Toward this end, we use the following modified procedure proposed by Simar and Wilson [38]:

Step 1. Solve the proposed Model (6) to calculate β_j^*, the efficiency score of $\mathrm{DMU}_j : (x_j, y_j)$ for $j = 1, 2, \ldots, J$.
Step 2. For $k = 1$ to 2000, repeat the following steps.
Step 3. Select at random with replacement $\bar{\beta}_{jk}^* : j = 1, \ldots, n$ from $\{\beta_1^*, \beta_2^*, \ldots, \beta_n^*\}$.

Step 4. Set $x_j^{*k} = \frac{\beta_j^*}{\beta_{jk}^*} x_j$, and calculate the $\tilde{\beta}_{jk}^*$, the efficiency score of $\tilde{\mathrm{DMU}}_{jk} : (x_j^{*k}, y_j)$ for $j = 1, 2, \ldots, n$ by solving Model (6).

$\tilde{\beta}_{jk}^*$ is considered as an estimate for β_j^*.

Now, we proceed with our two-step double-bootstrap DEA procedure. In the first step of our proposed double-bootstrap DEA procedure, we first use the above-mentioned procedure to calculate the bias-corrected efficiency of all DMUs. In the second step, we evaluate the effect of contextual variables on bias-corrected efficiency by using the regression Model (7).

4. An illustrative application

In this section, we focus on the performance evaluation of 20 power plants in Iran, given the importance of the power industry and its role in producing pollutants. Experts in the electricity industry emphasize the significance of using natural gas, as it is significantly less polluting than fuel oil and gasoil. Efforts are being made to replace fuel oil and gasoil with natural gas as a clean fuel in power plants. In this context, all three fuels – natural gas, fuel oil and gasoil – are inputs to the plants. However, to reduce the production of pollutants, priority is given to consuming natural gas. Therefore, natural gas is classified as an expandable input, while fuel oil and gasoil are considered reducible inputs.

This research includes 20 power plants in Iran during the ten-year period between 2011 and 2020. To conduct the efficiency analysis of these power plants, we need to specify their inputs, outputs, and explanatory variables. In this study, we consider three scenarios: First, we calculate plant-wise efficiency with a focus on electricity generation while keeping the level of pollutants constant. Second, we focus on minimizing the level of pollutants and maximizing electricity generation. Third, we use the weak disposability assumption of Shephard [37], where undesirable output is reduced and desirable output is increased.

We considered net electricity generation (NEG) (million kilowatt hours (kWh)) as our single desirable output and environmental pollution (EP) ($1000\,\mathrm{m}^3$) as our undesirable output. The inputs include total costs (staff costs and operational costs) (TC) (1000 rials), fuel oil and gasoil (FOG) (1000 Liters), and natural gas (NG) ($1000\,\mathrm{m}^3$). According to Qin *et al.* [30] and based on data availability, we considered three explanatory variables: economic development (ED) (1000 million rials), power structure (PS) (100 000 million rials) and technological innovation (TI) (10 000 million rials). The means and standard deviations of the data are listed in Table 1. The first five columns of Table 1 list the means of the data, and the second five columns give the standard deviations. The last three columns of the table give the mean of the explanatory factors. To evaluate the stochastic efficiencies of the observed power plants using our extended model, with a specific focus on their ability to generate net electricity and reduce the level of pollutants, we used two different directions $(d^v, d^w) = (1, 1)$ and $(d^v, d^w) = (v_o, w_o)$. We also considered two different confidence levels: $\gamma = 0.1$ and $\gamma = 0.5$. When using direction $(d^v, d^w) = (v_o, w_o)$, our approach is transformed into the traditional assumption of weak disposability. The results of technical efficiencies are presented in Table 2. According to the results, with confidence level $\gamma = 0.1$, twelve plants (Isfahan, Markazi, Tehran, Khozestan, Lorestan, S. Khorasan, Semnan, Boushehr, Hormozgan, Yazd, Qazvin, and N. Khorasan) are found to be efficient in both directions. However, when the confidence level is changed to $\gamma = 0.5$, nine plants (Markazi, Tehran, Khozestan, Semnan, Boushehr, Hormozgan, Yazd, Qazvin, and N. Khorasan) are found to be efficient.

After calculating the stochastic technical efficiencies of power plants, we have computed the Pearson correlation coefficients of the variables by pairing technical efficiency with all of the variables used in this application. We saw that there is a positive correlation between technical efficiency and the three explanatory variables. This test also revealed that the highest correlation of technical efficiency is associated with power structure, with a coefficient of 0.6010.

We have also determined the optimal values for variables in inefficient plants in these three directions. The results for both directions in the deterministic case ($\gamma = 0.5$) are presented in Table 3.

In the weak disposable technology with $(d^v, d^w) = (v_o, w_o)$, our results indicate that if we reduce the level of staff costs by 20.4% and the level of fuel oil and gasoil by 3.1%, and if we increase the level of natural gas by

TABLE 1. Mean and STD of the inputs and outputs along with the means of explanatory variables.

	Means					Std					Explanatory variables		
	TC	FOG	NG	NEG	EP	TC	FOG	NG	NEG	EP	ED	PS	TI
Isfahan	4.42E+09	2.67E+05	8.09E+05	3.63E+06	9.70E+04	4.08E+08	3.75E+05	3.04E+05	9.09E+05	2.20E+04	7.37E+00	8.35E+02	1.45E+02
Markazi	1.79E+09	3.24E+05	1.62E+06	6.95E+06	1.75E+05	1.47E+08	4.45E+05	4.33E+05	5.22E+05	9.99E+03	2.41E+00	1.30E+03	6.93E+02
Hamedan	2.00E+09	6.51E+05	5.26E+05	4.22E+06	1.06E+05	2.02E+08	1.50E+05	3.73E+05	1.40E+06	3.45E+04	1.52E+00	1.00E+03	8.45E+02
Tehran	1.30E+10	9.52E+04	1.04E+06	2.86E+06	1.03E+05	1.26E+09	1.12E+05	1.88E+05	2.85E+05	1.11E+04	2.81E+01	1.17E+03	6.11E+01
Khozestan	5.27E+09	5.76E+05	2.70E+06	1.13E+07	2.95E+05	6.22E+08	4.38E+05	7.24E+05	8.34E+05	3.13E+04	1.88E+01	6.64E+02	5.63E+01
Sistan & Balouchestan	3.48E+09	6.40E+05	5.39E+05	3.27E+06	1.06E+05	4.22E+08	2.70E+05	5.09E+05	8.35E+05	2.41E+04	1.78E+00	2.68E+02	1.96E+02
Kermanshah	1.72E+09	6.95E+05	3.75E+05	3.98E+06	9.64E+04	1.18E+08	8.38E+04	1.73E+05	4.55E+05	1.09E+04	1.91E+00	6.40E+02	4.29E+02
Guilan	3.44E+09	8.86E+04	4.40E+05	1.60E+06	4.77E+04	1.75E+08	6.12E+04	6.01E+04	1.91E+05	4.50E+03	2.80E+00	3.60E+02	1.77E+02
Mazandaran	4.74E+09	1.33E+06	1.56E+06	1.11E+07	2.59E+05	1.64E+08	3.96E+05	4.32E+05	9.49E+05	1.83E+04	4.19E+00	1.78E+03	5.43E+02
East Azerbayjan	3.65E+09	4.78E+05	6.89E+05	3.97E+06	1.05E+05	3.56E+08	3.71E+05	3.57E+05	3.23E+05	7.24E+03	4.45E+00	7.50E+02	2.29E+02
West Azerbayjan	4.18E+09	5.87E+03	4.30E+04	1.05E+05	4.52E+03	4.43E+08	5.28E+03	1.25E+04	1.69E+04	7.95E+02	2.54E+00	6.00E+01	4.27E+01
Lorestan	2.52E+09	5.68E+02	3.99E+04	7.95E+04	3.77E+03	2.46E+08	9.09E+02	1.91E+04	3.73E+04	1.69E+03	1.40E+00	9.00E+01	9.49E+01
S. Khorasan	1.24E+09	1.26E+05	7.45E+05	2.65E+06	7.85E+04	1.69E+08	5.80E+04	8.00E+04	2.40E+05	7.46E+03	6.35E-01	6.36E+02	1.28E+03
Semnan	9.60E+08	6.84E+04	2.36E+05	8.67E+05	2.51E+04	1.45E+08	4.20E+04	1.24E+05	3.99E+05	1.16E+04	1.14E+00	3.24E+02	3.65E+02
Boushehr	1.12E+09	5.34E+01	3.25E+05	6.83E+05	2.94E+04	1.40E+08	9.10E+01	3.66E+04	8.48E+04	3.29E+03	7.62E+00	2.14E+02	4.99E+01
Fars	4.82E+09	9.45E+03	1.38E+05	3.36E+05	1.34E+04	2.95E+08	1.02E+04	3.49E+04	8.87E+04	3.55E+03	5.97E+00	1.96E+02	4.40E+01
Hormozgan	2.41E+09	1.09E+06	2.69E+06	1.23E+07	3.40E+05	4.86E+08	2.66E+05	6.49E+05	1.25E+06	3.91E+04	2.54E+00	9.90E+02	5.10E+02
Yazd	2.14E+09	1.31E+05	9.26E+05	4.63E+06	9.53E+04	2.20E+08	7.87E+04	1.15E+05	4.12E+05	7.74E+03	2.29E+00	8.85E+02	4.98E+02
Qazvin	1.38E+09	1.06E+06	1.88E+06	1.14E+07	2.65E+05	9.27E+07	5.22E+05	6.09E+05	1.44E+05	9.87E+03	1.91E+00	1.04E+03	7.06E+02
N. Khorasan	1.07E+09	1.32E+05	8.78E+05	3.37E+06	9.10E+04	3.73E+07	5.68E+04	1.84E+05	7.76E+05	1.81E+04	6.35E-01	1.27E+03	2.56E+03

Table 2. The technical efficiency of the firms at different confidence levels.

	$\gamma = 0.1$			$\gamma = 0.5$		
$(d^v, d^w) =$	(v_o, w_o)	$(1, 1)$	$(1, 0)$	(v_o, w_o)	$(1, 1)$	$(1, 0)$
Isfahan	1.0000	1.0000	1.0000	0.9357	0.8250	0.8263
Markazi	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Hamedan	0.8863	0.7260	0.7260	0.9096	0.7261	0.7463
Tehran	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Khozestan	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Sistan & Balouchestan	0.7418	0.5353	0.5353	0.8421	0.6355	0.7300
Kermanshah	0.8887	0.7448	0.7448	0.8988	0.7117	0.7673
Guilan	0.8690	0.8690	0.8690	0.9171	0.8957	0.9158
Mazandaran	0.9865	0.9104	0.9104	0.9911	0.9240	0.6827
East Azerbayjan	0.8365	0.6387	0.6387	0.8967	0.7070	0.7627
West Azerbayjan	0.3850	0.9433	0.9433	0.5582	0.9537	0.9638
Lorestan	1.0000	1.0000	1.0000	0.5285	0.9618	0.9674
S. Khorasan	1.0000	1.0000	1.0000	0.9766	0.9514	0.9069
Semnan	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Boushehr	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Fars	0.7815	0.9499	0.9499	0.8615	0.9611	0.9589
Hormozgan	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Yazd	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Qazvin	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
N. Khorasan	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Average	0.9188	0.9159	0.9159	0.9158	0.9127	0.9114

49.3%, we should expect a 35.5% increase in net electricity generation, while the level of pollutants is reduced by 7.6%.

However, in the direction $(d^v, d^w) = (1, 1)$, if we reduce the level of staff costs by 20% and the level of fuel oil and gasoil by 3%, and if we increase the level of natural gas by 49%, we should expect a 35% increase in net electricity generation while the level of pollutants is reduced by 2%. These results are reasonable because when we replace natural gas with fuel oil and gasoil, we have prevented both the production of pollutants caused by these two fossil fuels and saved on human resources because natural gas requires fewer personnel.

Now, we are focusing on increasing electricity generation while keeping the level of pollutants constant. To achieve this, we set $(d^v, d^w) = (1, 0)$. The technical efficiencies at two different confidence levels are given in columns 4 and 7 of Table 2. The optimal values of inputs and desirable output for $\gamma = 0.5$ are given in Table 5. Our results indicate that in order to keep the level of pollutants constant, we should reduce the first and second inputs by 13% and 3%, respectively, and increase the level of the third input by 52%. In this case, net electricity generation will increase by 34%.

An interesting finding is that the direction we choose significantly influences the prioritization of plants. In the direction $(d^v, d^w) = (v_o, w_o)$, which represents the classic weak disposability assumption, the most inefficient plant is West Azerbaijan. However, when we change the direction to $(d^v, d^w) = (1, 1)$, the most inefficient plant becomes Sistan & Balouchestan.

As the efficiency results in Table 2 show, at confidence level $\gamma = 0.5$, out of 31 power plants, nine (30% of the plants) are efficient. This is not an acceptable discrimination, and we are interested in increasing discrimination power. In this sense, we use bootstrap DEA to calculate the bias and bias-corrected efficiency scores. The bias and bias-corrected efficiency scores are listed in Table 4. The average original stochastic BCC efficiency score in direction (v_o, w_o) is 0.9158, while the average bias-corrected efficiency score in this direction is reduced to 0.9142. Before applying the bootstrap DEA, we had nine efficient plants, while, after correcting bias, the discrimination

TABLE 3. Optimal values of inputs and outputs in different directions for $\gamma = 0.5$.

	TC	FOG	NG	NEG	EP
Original data	6.53E+10	7.76E+06	1.82E+07	8.94E+07	2.34E+06
$(d^v, d^w) = (1,1)$	5.20E+10	7.52E+06	2.72E+07	1.21E+08	2.30E+06
$(d^v, d^w) = (1,0)$	5.68E+10	7.52E+06	2.77E+07	1.20E+08	2.34E+06
$(d^v, d^w) = (v_o, w_o)$	5.20E+10	7.52E+06	2.72E+07	1.21E+08	2.16E+06

TABLE 4. Bias and bias-corrected efficiencies for $\gamma = 0.5$.

	Bias			Bias-corrected efficiency		
Company	(v_o, w_o)	$(1,1)$	$(1,0)$	(v_o, w_o)	$(1,1)$	$(1,0)$
Isfahan	0.0019	0.0016	0.0015	0.9338(12)	0.8234(16)	0.8248(15)
Markazi	0.0001	0.0011	0.0024	0.9999(1)	0.9989(2)	0.9976(9)
Hamedan	0.0024	0.0019	0.0017	0.9072(14)	0.7242(17)	0.7446(17)
Tehran	0.0017	0.0015	0.0013	0.9983(6)	0.9983(6)	0.9987(3)
Khozestan	0.0005	0.0002	0.0006	0.9995(2)	0.9998(1)	0.9994(1)
Sistan & Balouchestan	0.0024	0.0020	0.0019	0.8397(18)	0.6335(20)	0.7281(16)
Kermanshah	0.0007	0.0005	0.0004	0.8981(16)	0.7112(18)	0.7669(18)
Guilan	0.0015	0.0016	0.0013	0.9156(13)	0.8941(15)	0.9145(13)
Mazandaran	0.0020	0.0018	0.0014	0.9891(10)	0.9222(14)	0.6813(20)
East Azerbayjan	0.0018	0.0012	0.0025	0.8949(17)	0.7058(19)	0.7602(19)
West Azerbayjan	0.0009	0.0008	0.0007	0.5573(19)	0.9529(12)	0.9631(11)
Lorestan	0.0019	0.0018	0.0022	0.5266(20)	0.9600(10)	0.9652(10)
S. Khorasan	0.0016	0.0015	0.0009	0.9750(11)	0.9499(13)	0.9060(14)
Semnan	0.0015	0.0019	0.0015	0.9985(5)	0.9981(7)	0.9985(5)
Boushehr	0.0013	0.0019	0.0019	0.9987(4)	0.9980(8)	0.9981(7)
Fars	0.0024	0.0013	0.0018	0.8591(17)	0.9598(11)	0.9571(12)
Hormozgan	0.0023	0.0015	0.0015	0.9977(8)	0.9985(5)	0.9984(6)
Yazd	0.0023	0.0020	0.0021	0.9976(9)	0.9979(9)	0.9979(8)
Qazvin	0.0011	0.0013	0.0011	0.9989(3)	0.9987(3)	0.9989(2)
N. Khorasan	0.0019	0.0013	0.0014	0.9981(7)	0.9986(4)	0.9986(4)
Average	0.0016	0.0014	0.0015	0.9142	0.9112	0.9099

power has been substantially increased. In this sense, the bias-corrected efficiency scores are used to rank all plants. The ranking results are given in parentheses in Table 4.

Now, we use the bias-corrected efficiency scores to remove the impact of contextual variables on firm efficiency. We focus on the deterministic case ($\alpha = 0.5$). The bias-corrected efficiency score is our dependent variable, and the three explanatory variables (economic development, power structure, and technological innovation) are independent variables. The results of the regression model are given in Table 5. The results showed that all three variables have a direct relation with technical efficiency. However, economic development and power structure are statistically significant. It is interesting to see that although technological innovation has a direct relationship with technical efficiency, this relationship is not statistically significant.

We need to point out that the R-square value is the percentage of the dependent variable variation (logarithm of technical efficiency) that a linear model explains. Our regression has an adjusted R-squared of 0.28. This reveals that 28% of the variability observed in the target variable is explained by the regression model.

TABLE 5. Regression statistics.

Multiple R	0.6303				
R Square	0.3973				
Adjusted R Square	0.2843				
Standard Error	0.11656				
Observations	20				
		ANOVA			
	df	SS	MS	F	Significance F
Regression	3	0.14328	0.0478	3.5155	0.0395
Residual	16	0.21737	0.0136		
Total	19	0.36065			
	Coefficient s	Standard Error	t Stat	P-value	
Intercept	0.7688	0.0520	14.7876	9.46E−11	
Economic development	0.0044	0.00461	0.9609	0.0351	
Power Structure	0.0001	7.1E−05	2.0341	0.0059	
Technological Innovation	3.93E−05	5.97E−05	0.6577	0.5201	

Policy implications

Our findings in this application have important policy implications for performance analysis and discriminatory power.

First, as we should expect, the confidence level plays a crucial role in determining which power plants are identified as efficient. We found that inefficient plants tend to experience a decline in their relative efficiency performance with an increase in confidence level.

At confidence level $\gamma = 0.5$ (the deterministic case), 45% of the plants were efficient. Most of the inefficiency of power plants is related to gas fuel consumption. Therefore, it is strongly recommended to increase the consumption of gas fuel and reduce the consumption of fuel oil and gasoil.

We have found that the direction we use to reduce the level of undesirable output plays a crucial role in determining which companies are identified as efficient. Different directions prioritize different aspects of the production process and the balance between desirable and undesirable outputs. As a result, the efficiency scores may vary significantly depending on the selected direction.

We also found that the first two explanatory variables (economic development and power structure) have a direct relationship with technical efficiency. The explanatory variables, economic development and power structure are statistically significant. It is interesting to see that although technological innovation has a direct relationship with technical efficiency, this relationship is not statistically significant.

All simulations can easily be conducted by PYTHON, R and GAMS (General Algebraic Modeling System) software. In this application, we used GAMS software on a personal computer with Core i7 processor. However, the experiment can be conducted using the same compiler or a personal computer with Core i5 processor.

5. Concluding remarks

The indiscriminate consumption of fossil fuels and the resulting production of pollutants have posed serious threats to our environment. It is highly recommended to shift towards the use of renewable energies or fossil fuels with lower pollution levels, such as natural gas. In this study, we focused on estimating the relative technical efficiency of firms when the data are given in stochastic form and some inputs and outputs need to be reduced and some others must be increased. To achieve this, we developed a stochastic directional distance function program to calculate the relative performance of firms. Our approach was illustrated through an application to power plants over a ten-year period between 2011 and 2020. The most important finding of these applications is that the most important source of inefficiency in power plants is related to not using natural gas fuel and optimal net electricity generation. Another finding from this analysis is that while all three explanatory variables (economic development, power structure, and technological innovation) have a direct relationship with technical efficiency, technological innovation is not statistically significant.

Our results also revealed that by reducing the levels of staff costs and fuel oil and gasoil, and simultaneously increasing the level of natural gas by a certain amount, there can be a substantial increase in net electricity generation while keeping the level of pollutants constant.

One of the research limitations is that the evaluation process of the power plants depends mainly on the input and output data; thus, data availability was an important issue.

As potential future research, one can explore other types of imprecise data, such as fuzzy-type inputs and outputs. Developing an efficiency analysis model in the presence of reducible outputs and expandable inputs is recommended. Another extension could involve the use of both weak and managerial disposability assumptions in handling reducible and expandable inputs and outputs in a production process. These further investigations could provide valuable insights into improving the efficiency and environmental performance of firms.

Acknowledgements

The authors would like to thank the Editor-in-Chief, the Area Editor, and the anonymous reviewers for their valuable feedback on the previous version of this manuscript.

Data availability statement

The dataset is available per request.

References

[1] N.K. Al-Mezeini, A. Oukil and A.M. Al-Ismaili, Investigating the efficiency of greenhouse production in Oman: a two-stage approach based on Data Envelopment Analysis and double bootstrapping, *J. Clean. Prod.* **247** (2020) 119160.

[2] A. Amirteimoori, T. Allahviranloo, M. Zadmirzaei and F. Hasanzadeh, On the environmental performance analysis: A combined fuzzy data envelopment analysis and artificial intelligence algorithms. *Expert Syst. App.* **224** (2023) 119953.

[3] A. Amirteimoori, B.K. Sahoo and S. Mehdizadeh, Data envelopment analysis for scale elasticity measurement in the stochastic case: with an application to Indian banking. *Finan. Innov.* **9** (2023) 31.

[4] R.D. Banker, Maximum likelihood, consistency and data envelopment analysis: a statistical foundation. *Manage. Sci.* **39** (1993) 1265–1273.

[5] R.D. Banker, A. Charnes and W.W. Cooper, Models for the estimation of technical and scale inefficiencies in data envelopment analysis. *Manage. Sci.* **30** (1984) 1078–1092.

[6] R.D. Banker, A. Amirteimoori and R.P. Sinha, An integrated Data Envelopment Analysis and generalized additive model for assessing managerial ability with application to the insurance industry, *Decis. Anal. J.* **4** (2022) 100115.

[7] A.S. Camanho and R.G. Dyson, A generalization of the Farrell cost-efficiency measure applicable to non-fully competitive settings. *Omega* **36** (2008) 147–162.

[8] V. Charles, M. Kumar and S. Irene Kavitha, Measuring the efficiency of assembled printed circuit boards with undesirable outputs using data envelopment analysis. *Int. J. Prod. Econ.* **136** (2012) 194–206.

[9] A. Charnes, W.W. Cooper and E. Rhodes, Measuring the efficiency of decision-making units. *Eur. J. Oper. Res.* **2** (1978) 429–444.

[10] Y. Chen and L. Liu, Improving eco-efficiency in coal mining area for sustainability development: an emergy and super-efficiency SBM-DEA with undesirable output. *J. Clean. Prod.* **339** (2022) 130701.

[11] W.W. Cooper, Z. Huang, V. Lelas, S.X. Li and O.B. Olesen, Chance constrained programming formulations for stochastic characterizations of efficiency and dominance in DEA. *J. Prod. Anal.* **9** (1998) 53–79.

[12] W.W. Cooper, H. Deng, Z. Huang and S.X. Li, Chance constrained programming approaches to congestion in stochastic data envelopment analysis. *Eur. J. Oper. Res.* **155** (2004) 487–501.

[13] W.W. Cooper, Z. Huang and S.X. Li, Chance-constrained DEA, in Handbook on Data Envelopment Analysis. Springer, Boston, MA (2011) 211–240.

[14] A. Emrouznejad, B.R. Parker and G. Tavares, Evaluation of research in efficiency and productivity: a survey and analysis of the first 30 years of scholarly literature in DEA. *Soc.-Econ. Planning Sci.* **42** (2008) 151-157.

[15] R. Färe and S. Grosskopf, Non-parametric productivity analysis with undesirable outputs. *Am. J. Agric. Econ.* **85** (2003) 1070–1074.

[16] M.J. Farrell, The measurement of productive efficiency. *J. R. Stat. Soc. (Ser. A)* **120** (1957) 253–90.

[17] A. Hailu and T.S. Veeman, Non-parametric productivity analysis with undesirable outputs: an application to the Canadian Pulp and paper industry. *Am. J. Agric. Econ.* **83** (2001) 605–616.

[18] Z. Hua and Y. Bian, DEA with undesirable factors, in Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis, edited by J. Zhu and W.D. Cook. Springer, Boston, MA (2007).

[19] Z. Huang and S.X. Li, Stochastic DEA models with different types of input-output disturbances. *J. Prod. Anal.* **15** (2001) 95–113.

[20] C. Kao and S.T. Liu, Stochastic efficiency measures for production units with correlated data. *Eur. J. Oper. Res.* **273** (2019) 278-287.

[21] S. Kordrostami, A. Amirteimoori, F. Seihani Parashkouh, M. Mahboubi and M. Jahani Sayyad Noveiri, Weak Disposability of input and output in a nonparametric production analysis. *Iran. Econ. Rev.* **25** (2021) 551–565.

[22] T. Kuosmanen, Weak disposability in nonparametric production analysis with undesirable outputs. *Am. J. Agric. Econ.* **87** (2005) 1077–1082.

[23] K.C. Land, C.A.K. Lovell and S. Thore, Productive efficiency under capitalism and state socialism: an empirical inquiry using chance-constrained data envelopment analysis. *Technol. Forecasting Soc. Change* **46** (1994) 139–152.

[24] W.B. Liu, W. Meng, X.X. Li and D.Q. Zhang, DEA models with undesirable inputs and outputs. *Ann. Oper. Res.* **173** (2010) 177–194.

[25] M. Mahdiloo, A.H. Jafarzadeh, R. Farzipoor Saen, Y. Wu and J. Rice, Modelling undesirable outputs in multiple objective data envelopment analysis. *J. Oper. Res. Soc.* **69** (2018) 1903–1919.

[26] O.B. Olesen, Comparing and combining two approaches for chance-constrained DEA. *J. Prod. Anal.* **26** (2006) 103–119.

[27] O.B. Olesen and N.C. Petersen, Stochastic data envelopment analysis – a review. *Eur. J. Oper. Res.* **251** (2016) 2–21.

[28] O.B. Olesen and N.C. Petersen, Foundation of chance constrained data envelopment analysis for Pareto-Koopmann efficient production possibility sets, in International DEA Symposium 2000, Measurement and Improvement in the 21st Century. The University of Queensland (2000) 313–349.

[29] A. Oukil, N. Channouf and A. Al-Zaidi, Performance evaluation of the hotel industry in an emerging tourism destination: the case of Oman. *J. Hospitality Tourism Manage.* **29** (2016) 60–68.

[30] Q. Qin, Y. Jiao, X. Gan and Y. Liu, Environmental efficiency and market segmentation: an empirical analysis of China's thermal power industry. *J. Clean. Prod.* **242** (2020) 118560.

[31] Y. Qu, J. Li and S. Wang, Green total factor productivity measurement of industrial enterprises in Zhejiang Province, China: a DEA model with undesirable output approach. *Energy Rep.* **8** (2022) 307–317.

[32] K. Rashidi and R. Farzipoor Saen, Measuring eco-efficiency based on green indicators and potentials in energy saving and undesirable output abatement. *Energy Econ.* **50** (2015) 18–26.

[33] B.K. Sahoo and K. Tone, Non-parametric measurement of economies of scale and scope in non-competitive environment with price uncertainty. *Omega* **41** (2013) 97–111.

[34] L. Seiford and J. Zhu, Modeling undesirable factors in efficiency evaluation. *Eur. J. Oper. Res.* **42** (2002) 16–20.

[35] J.K. Sengupta, Transformations in stochastic DEA models. *J. Econ.* **46** (1990) 109–123.

[36] J.K. Sengupta, Dynamic and Stochastic Efficiency Analysis: Economics of Data Envelopment Analysis. World Scientific, Singapore (2000).

[37] R.W. Shephard, Theory of Cost and Production Functions. Princeton University Press, Princeton (1970).

[38] L. Simar and P.W. Wilson, Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Manage. Sci.* **44** (1998) 49–61.

[39] T. Sueyoshi and M. Goto, DEA radial and non-radial models for unified efficiency under natural and managerial disposability: theoretical extension by strong complementary slackness conditions. *Energy Econ.* **34** (2012) 700–713.

[40] T. Sueyoshi, M. Goto and M.A. Snell, DEA environmental assessment: measurement of damages to scale with unified efficiency under managerial disposability or environmental efficiency. *Appl. Math. Modell.* **37** (2013) 7300–7314.

[41] G. Wei, J. Chen and J. Wang, Stochastic efficiency analysis with a reliability consideration. *Omega* **48** (2014) 1–9.

[42] A. Zanella, A. Camanho and T.G. Dias, Undesirable outputs and weighting schemes in composite indicators based on data envelopment analysis. *Eur. J. Oper. Res.* **245** (2015) 517–530.

[43] P. Zhou, B.W. Ang and K.L. Poh, A survey of data envelopment analysis in energy and environmental studies. *Eur. J. Oper. Res.* **189** (2008) 1–18.

[44] Z. Zhou, G. Xu, C. Wang and J. Wu, Modeling undesirable output with a DEA approach based on an exponential transformation: an application to measure the energy efficiency of Chinese industry. *J. Clean. Prod.* **236** (2019) 117717.