Expert Systems with Applications 42 (2015) 760–773

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A data envelopment analysis model with interval data and undesirable output for combined cycle power plant performance assessment

Kaveh Khalili-Damghani [a], Madjid Tavana [b,c,*], Elham Haji-Saami [a]

[a] Department of Industrial Engineering, South-Tehran Branch, Islamic Azad University, Tehran, Iran
[b] Business Systems and Analytics Department, Lindback Distinguished Chair of Information Systems and Decision Sciences, La Salle University, Philadelphia, PA 19141, USA
[c] Business Information Systems Department, Faculty of Business Administration and Economics, University of Paderborn, D-33098 Paderborn, Germany

ARTICLE INFO

Article history:
Available online 30 August 2014

Keywords:
Economic scale size
Return to scale
Data envelopment analysis
Imprecise data: undesirable outputs
Combined cycle power plant

ABSTRACT

Determining the optimal scale size of a combined cycle power plant is inherently a complex problem often with multiple and conflicting criteria as well as uncertain factors. The complexity of the problem is compounded by the production of undesirable outputs and the presence of natural and managerial disposability. We propose a customized data envelopment analysis (DEA) method for solving the return to scale (RS) problem in the presence of uncertain data and undesirable outputs. A combined cycle power plant is considered a decision making unit (DMU) which consumes fuels to produce electricity and emissions. The uncertainty of the inputs and outputs are modeled with interval data and the emissions are assumed to be undesirable outputs. The proposed DEA method determines the interval efficiency scores of the DMUs and offers a practical benchmark for enhancing the efficiency scores. We demonstrate the applicability of the proposed method and exhibit the efficacy of the procedure with a six-year study of 17 combined cycle power plants in Iran. The main contributions of this paper are six fold: we (1) model the uncertainties in the input and output data using interval data; (2) consider undesirable outputs; (3) determine the efficiency scores of the DMUs as interval values; (4) develop a group of indices to distinguish between the efficient and inefficient DMUs; (5) determine the most economic scale size for the efficient DMUs; and (6) determine practical benchmarks for the inefficient DMUs.

1. Introduction

Energy generation, distribution and consumption play a vital role in the economy. Generation and distribution of clean energy and sustainable development considerations are essential for future generations especially in developing countries. Therefore, energy systems should systematically be assessed and updated in order to improve the performance, to determine the economic scale size and to efficiently utilize resources.

A great deal of research has been devoted to the field of energy system assessment and evaluation during recent years. Azade, Ghaderi, and Maghsoudi (2008) proposed an integrated hierarchical approach based on data envelopment analysis (DEA), principal component analysis and numerical taxonomy to locate solar plants in different regions and cities in Iran. Aguirre, Villalobos, Phelan, and Pacheco (2011) presented a methodology to measure relative industrial energy efficiency across plants within a manufacturing sector through the use of energy-production signatures. They used linear programming, regression, benchmarking and simulation models to study the behavior of a representative manufacturing plant. The methodology was validated using data from the Department of Energy database. Bampatsou, Papadopoulos, and Zervas (2013) used a DEA model to determine the technical efficiency index of EU-15 countries. Salazar-Ordóñez, Pérez-Hernández, and Martín-Lozano (2013) estimated the potential efficiency of the sugar beet crop in Spain using a DEA model. Kagawa, Takezono, Suh, and Kudoh (2013) evaluated the efficiency of an advanced bio-diesel plant in Japan using DEA. Fang, Hu, and Lou (2013) used an input-oriented variable return to scale (VRS) DEA model to compute the pure technical efficiency and energy-saving targets. Fang et al. (2013) utilized a four-stage DEA and studied the effects of industry characteristics on the energy-saving targets.

Chang, Zhang, Danao, and Zhang (2013) proposed a non-radial DEA model with slack-based measures to analyze the environmental

* Corresponding author at: Business Systems and Analytics Department, Lindback Distinguished Chair of Information Systems and Decision Sciences, La Salle University, Philadelphia, PA 19141, USA. Tel.: +1 215 951 1129; fax: +1 267 295 2854.

E-mail addresses: kaveh.khalili@gmail.com (K. Khalili-Damghani), tavana@lasalle.edu (M. Tavana), elham.hajisami@gmail.com (E. Haji-Saami).

URLs: http://kaveh-khalili.webs.com (K. Khalili-Damghani), http://tavana.us/ (M. Tavana).

http://dx.doi.org/10.1016/j.eswa.2014.08.028
0957-4174/

efficiency of the Chinese transportation sector. Zhou, Xing, Fang, Liang, and Xu (2013) proposed a new non-radial DEA approach by integrating the entropy weight and the slack-based model to evaluate efficiency of the Chinese power industry at the provincial level. Çelen (2013) used stochastic frontier analysis to analyze the efficiency performances of Turkish electricity distribution companies. Çelen (2013) measured "how the efficiency performances of the electricity distribution regions were affected by the merges between distribution regions". Kuosmanen, Saastamoinen, and Sipiläinen (2013) compared the impact of three methods on cost efficiency estimation and analysis. Kuosmanen et al. (2013) validated the performance of their approach using Monte Carlo simulations in a study of the electricity distribution industry in Finland. Vazhayil and Balasubramanian (2013) proposed deterministic and stochastic DEA models to optimize energy planning in the Indian power sector.

Wu, An, Xiong, and Chen (2013) proposed a method that considered undesirable outputs and analyzed congestion of industrial regions in China. Bian, He, and Xu (2013) evaluated regional energy efficiency in China based on a non-radial DEA model. Bian et al. (2013) treated non-fossil energy as a fixed input in order to measure energy savings as well as reduce Carbon dioxide (CO_2) emission for improving efficiency. Zhang, Zhou, and Choi (2013) modeled energy and CO_2 emission performance in electricity generation and proposed a meta-frontier non-radial directional distance function to consider the heterogeneous group of electricity generation, non-radial slacks, and undesirable outputs, simultaneously. Zhang et al. (2013) studied electricity generation in Korea and estimated the CO_2 emissions and the potential reductions in energy usage under different technological assumptions. Riccardi, Oggioni, and Toninelli (2012) estimated the efficiency of high energetic and CO_2 emissions in the cement production process in 21 countries. Riccardi et al. (2012) compared standard DEA models with a directional distance function approach to measure the efficiency in the presence and in the absence of CO_2 emission. Sueyoshi and Goto (2010) reformulated the original non-radial DEA model to handle undesirable (bad) outputs. They applied their method on the operational, environmental and combined efficiency measures of US coal-fired power plants. Wu, Fan, Zhou, and Zhou (2012) measured industrial energy efficiency by constructing both static and dynamic energy efficiency performance indices. Wu et al. (2012) considered undesirable outputs such as CO_2 emissions in their modeling framework. Mandal (2010) evaluated the Indian cement industry in the presence of emissions and undesirable outputs using a DEA model.

Although electricity is considered a form of clean energy, production of electricity using gas, steam, and combined cycle power plants often causes some emissions and pollutions. Therefore, the output of electricity production is often mixed with some unwanted and undesirable outputs. Consequently, it is critical to improve the process of efficiency measurement and the determination of economic scale size of electricity power plants. Moreover, better estimate of emissions and pollutions produced by power plants can enhance the strategic planning of sustainable development.

The problem of performance assessment of electrical power plants is challenging and complex. This problem usually involves multiple and often conflicting criteria and undesirable outputs which are often difficult to assess because of environmental uncertainties. The production of a variety of emissions, pollutions, and other undesirable outputs such as CO_2 causes further complications.

To the best of our knowledge, there is no single method in the literature that can measure the efficiency scores and to determine the most economic scale size of decision making units (DMUs) in the presence of undesirable outputs and uncertain data. In this paper, we propose a DEA method for measuring the performance of combined cycle power plants in the presence of pollution production and data uncertainty. A combined cycle power plant is assumed to be a DMU which consumes fossil fuels to produce electricity as desirable outputs and polluting gases (i.e., CO_2, SO_x and NO_x) as undesirable outputs. Moreover, the uncertainty of inputs and outputs during the planning horizon is modeled using interval data. The proposed approach is used to determine the most economic scale size of power plants and to present practical suggestions for efficiency improvement of inefficient DMUs. The proposed method is customized and applied to a case study of Iranian electrical power plants to illustrate its applicability and efficacy. The theoretical contributions of this paper are six fold: we (1) model the uncertainties in the input and output data using interval data; (2) consider undesirable outputs; (3) determine the efficiency scores of the DMUs as interval values; (4) develop a group of indices to distinguish between the efficient and inefficient DMUs; (5) determine the most economic scale size for the efficient DMUs; and (6) determine practical benchmarks for the inefficient DMUs.

The remainder of this paper is organized as follows. In Section 2, we briefly review the relevant literature on the conventional DEA models, uncertainty in DEA models, undesirable outputs in DEA models, RS in DEA models, and the applications of DEA models in power plants assessment and energy generation. In Section 3, we briefly outline the mathematical basis of the DEA models. In Section 4, we develop the proposed DEA model in the presence of interval data and undesirable outputs. In Section 5, we demonstrate the applicability of the proposed method and exhibit the efficacy of the procedure using a six-year study of 17 combined cycle power plants in Iran. In Section 6, we discuss the managerial implications and in Section 7, we present our conclusion and future research directions.

2. Literature review

DEA is a non-parametric method for evaluating the relative efficiency of DMUs with multiple inputs and multiple outputs (Charnes, Cooper, & Rhodes, 1978). The first DEA model (i.e., the CCR model) was proposed by Charnes et al. (1978) by considering the constant RS assumption. Banker, Charnes, and Cooper (1984) extended the Charnes et al.'s (1978) model by proposing the BCC model and considering the VRS assumption (Cooper, Seiford, & Tone, 2007).

2.1. Uncertainty in DEA models

In general, classical DEA problems are solved under the assumption that the values of parameters are specified precisely in a crisp environment. However, the observed values of the input and output data in real-world problems are often imprecise or vague. Imprecise evaluations is primarily the result of unquantifiable, incomplete and non-obtainable information. The most common form of uncertainty in DEA problems occurs when some or all of the relevant data are not known precisely. This type of uncertainty is called *ambiguity* (Inuiguchi & Ramík, 2000). The ambiguity in data can be modeled using the possibility approach parameterized through fuzzy sets, or the probability approach parameterized through random variables, or interval data. Several optimization methods such as stochastic programming, fuzzy mathematical programming, and interval mathematical programming are proposed to take into account various uncertainties in the optimization process (Liu, Huang, Liu, Fuller, & Zeng, 2003).

Stochastic programming methods model uncertainties with probability distributions derived from historical data (Peidro,