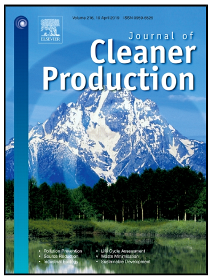
Accepted Manuscript



Performance evaluation of thermal power plants considering CO2 emission: A multistage PCA, Clustering, Game theory and Data Envelopment Analysis

Reza Mahmoudi, Ali Emrouznejad, Hossein Khosroshahi, Mehdi Khashei, Parisa Rajabi

PII:	S0959-6526(19)30736-X
DOI:	10.1016/j.jclepro.2019.03.047
Reference:	JCLP 16051
To appear in:	*Journal of Cleaner Production*
Received Date:	20 November 2018
Accepted Date:	05 March 2019

Please cite this article as: Reza Mahmoudi, Ali Emrouznejad, Hossein Khosroshahi, Mehdi Khashei, Parisa Rajabi, Performance evaluation of thermal power plants considering CO2 emission: A multistage PCA, Clustering, Game theory and Data Envelopment Analysis, *Journal of Cleaner Production* (2019), doi: 10.1016/j.jclepro.2019.03.047

Performance evaluation of thermal power plants considering CO2 emission: A multistage PCA, Clustering, Game theory and Data Envelopment Analysis

Reza Mahmoudi[1], Ali Emrouznejad[2*], Hossein Khosroshahi[1], Mehdi Khashei[1], Parisa Rajabi[1]

[1]Department of Industrial and Systems Engineering, Isfahan University of Technology, Isfahan, Iran

[2]Aston Business School, Aston University, Birmingham, B47ET, UK

Abstract

Data envelopment analysis is a relative performance assessment method to evaluate performance of a group of decision making units. Empirically, when the number of decision making units is insufficient, the classical data envelopment analysis models cannot discriminate the efficient units perfectly. To overcome this issue, in this paper, several mathematical approaches, including "multivariate data analysis techniques", "game theory", "Shannon entropy" and "the technique for order of preference by similarity to ideal solution", are combined with data envelopment analysis. The proposed framework is applied to evaluate performance of Iranian thermal power plants. Inefficient performance of thermal power plants may end up in serious economic and environmental problems for example CO_2 emission. Therefore, evaluating performance of thermal power plants and identifying their weaknesses in order to improve their performance is a necessity. The obtained results are analyzed, and some practical suggestions are provided to achieve sustainable performance and a cleaner production system.

Keywords: Data Envelopment Analysis; CO2 emission; Multivariate data analysis; Thermal power plants; TOPSIS; Shannon entropy

Abbreviations

PCA	Principal Component Analysis
TOPSIS	Technique for Order of Preference by Similarity to Ideal Solution
TPP	Thermal Power Plant
DEA	Data Envelopment Analysis
DMU	Decision Making Unit
MDA	Multivariate Data Analysis
MADM	Multi-Attribute Decision Making

[*] Corresponding author at: Ali Emrouznejad, Aston Business School, Aston University, Birmingham, UK.
E-mail address: a.emrouznejad@aston.ac.uk, Website: http://www.deazone.com

PC	Principal Component
Cat1	Category 1
Cat2	Category 2
AHP	Analytic hierarchy process
VIKOR	Viekriterijumsko kompromisno rangiranje
LINMAP	Linear Programming Technique for Multidimensional Analysis of Preference
ELECTRE	ELimination Et Choix Traduisant la REalité
CO_2	Carbon dioxide

Nomenclature

x_{ij}	**i^{th} input of** DMU_j
y_{rj}	r^{th} output of DMU_j
u_r	The relative weight assigned to output j
v_i	The relative weight assigned to input i
e_o	Efficiency of DMU_o.
$\boldsymbol{H_j}$	The Shannon Entropy of i^{th} criterion
$\boldsymbol{w_j}$	The weight of criterion j
$\boldsymbol{C_j}$	The relative closeness coefficient of each alternative
$\boldsymbol{D_n}$	The weighted normalized decision matrix in TOPSIS
ESS_j	Sum of square deviations of members of cluster j from the cluster's center

1. Introduction

Sustainability has three different dimensions including, environment, social and economic. A truly sustainable system must be sustainable in all dimensions (Mahmoudi and Rasti-Barzoki, 2017). Recently the environmental and sustainability issues have been categorized as one of the most important and crucial problems across the board. Therefore, both academic and practical communities have focused on these problems in many industries. Among all industries, energy sector always is known as one of the most polluter industries. Based on this fact, there is an increasing attention toward the performance of power plants, especially Thermal Power Plants (TPPs). This trend seems a completely rational trend, since for example just in China, TPPs release more than 40% of CO_2 emissions across the country. In addition, they consume about 45% of China's energy supply (Wang et al., 2017). Hence, managers and researchers are seeking out possible strategies to overcome challenging problems related to TPPs activities. To choose and apply an efficient and applicable strategy in order to improve performance of a TPP, first step is a comprehensive performance analysis to identify existing weaknesses. There are different methods suggested to performance analysis such as stochastic frontier analysis, data envelopment analysis (DEA) etc.

DEA is a nonparametric method to evaluate relative performance of a group of Decision Making Units (DMUs). By introduction the first DEA model (classic CCR) by Charnes et al. (1978), DEA has been widely used to evaluate the performance of DMUs in different areas such as environmental issues, urban sustainability, banking, hospitals and healthcare systems, supply chains, transportation etc. (Emrouznejad and Yang, 2018; Mahmoudi et al., 2018b).

Analyzing trend of recently published articles shows that DEA is a very useful and popular tool in efficient use of energy in a wide variety of applications (e.g. Munisamy and Arabi (2015) and Yang et al. (2018)). In particular, using DEA to study performance of the power plants has been one of the most noteworthy topics for researchers in recent years. Lam and Shiu (2001) applied DEA to a technical efficiency assessment of china's thermal power plants. Their analyses were based on cross-sectional data of 30 provinces, autonomous regions, and municipalities, each of them is considered as a DMU. The results show that fuel efficiency and capacity factor significantly affect technical efficiency. Based on the data set containing 65 thermals, hydro and wind power plants, owned by private and public sectors, Sarıca and Or (2007) used constant returns to scale, variable returns to scale and assurance region DEA models to survey the performance of electricity generation plants in Turkey. Liu et al. (2010) evaluated the performance efficiency of 9 thermal power plants in Taiwan during 2004–2006 using DEA approach. According to their results the combined cycle power plants were the most efficient ones. Rezaee et al. (2012) evaluated the performance of thermal power plants in Iran. They combined the Nash bargaining game model with DEA classic CCR model and proposed a new DEA model. By using DEA, Yadav et al. (2013) evaluated the performance of small scale thermal power plants in India from 2008 to 2009. They used slack analyzes to identify the slack values in inputs and outputs. Coşgun and Kaya (2014) used DEA and goal programming to rank natural gas based and coal based power plants in order to select the best thermal power plant construction scenario. Azad et al. (2015) applied DEA to survey how electricity competitive market establishment affects technical thermal power plants performance. The results illustrate that market restructuring positively affects technical efficiency of power plants. Arabi et al. (2016) proposed some new models based on DEA to assess the performance of power plants considering fuel consumption and environmental issues. Sahoo et al. (2017) used DEA models to evaluate the relation between energy saving potential of thermal power plants and their efficiencies and investigate the target set rationality for the Indian power set.

While DEA is widely used in the literature to study the relative performance of TPPs (Liu et al., 2018; Liu and Wen, 2012), DEA models, as many other mathematical methods, have number of limitations and conditions. Minimum number of needed DMUs is one of the most challenging initial conditions to apply a DEA model. There is a rule of thumb that says the relation between the number of DMUs (n) and the number of inputs (m) and outputs (s) must be $3(m+s)<n$ (Friedman and Sinuany-Stern, 1998). This rule is a critical initial condition to guaranty the discrimination power of DEA models.

In the real world, when governments, policy makers or managers try to design a cleaner production system and achieve sustainable situation, they should analyze performance of existing DMUs. In many industries, number of existing DMUs are not enough to fulfill the initial condition of DEA. Therefore, evaluating process cannot be conducted. In particular, in many small or developing countries, while the performance of TPPs are mostly inefficient and the process of evaluating and identifying weaknesses is a necessity, number of existing DMUs are significantly low. Although different DEA models have been suggested to overcome this limitation, it should be noted that mostly the rankings and efficiency scores obtained by different approaches are different from each other. To the best of our knowledge, there is not any meter to determine which method produce more accurate and reliable scores. Therefore, it is essential to develop an approach to evaluate the performance of TPPs in such a situation and provide a unique rank based on the results of different models.

In this paper, several integrated approaches based on Multivariate Data Analysis (MDA), DEA and game theory have been used to overcome this classic shortcoming of DEA models. Then proposed approaches are applied to evaluate the performance of Iranian TPPs. 8 inputs and 3 outputs have been considered for this study. According to the number of considered inputs and outputs and initial condition, at least 34 DMUs are needed. However, the available data set of Iranian TPPs does not satisfy required minimum number of DMUs (only data of 24 TPPs is available). After applying the proposed approaches for this real case, using Shannon entropy and the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), an approach is suggested to analyze results using different methods for providing the final ranks.

From sustainability perspective, there are a variety of factors which affect performance of TPPs, such as energy consumption, labor performance, environmental and managerial performance (Bi et al., 2018; Jin et al., 2018). In particular, the efficiency of TPPs strongly depend on their managerial strategies, and these strategies will be significantly affected by the ownership, especially in developing countries. Hence, in addition to improving operational factors, reforms in ownership and management board structures, and other non-operational factors, must be considered as a possible strategy to make performance of TPPs more sustainable (Arabi et al., 2016; Omrani et al., 2018). Considering these explanations, this study especially tries to answer the following research questions:

- In evaluating the performance of TPPs, how can the evaluators overcome the problem of insufficient number of DMUs?
- How can a unique rank be achieved according to the results of different methods?
- How sustainable is the performance of Iranian TPPs? What are their weaknesses?
- What are the candidate short/long term and micro/macro policies to improve the environmental, managerial and operational performance of Iranian TPPs?
- Should the government provide financial supports for TPPs to have cleaner production system.
- Could privatization help to improve the performance of Iranian TPPs?

The rest of the paper is organized as follows: In section 2 first a background about the used methodologies is presented. Then an algorithm is proposed to overcome the problem of insufficient number of DMUs. A Multi-Attribute Decision Making (MADM) method is suggested to obtain unique ranks based on the results of different approaches in section 3. In section 4, the real case study of the Iranian thermal power plant sector is presented. The results and analyses of the application are provided and analyzed in section 5. Conclusions are summarized in section 6. Finally, section 7 proposes some possible research directions for future studies.

2. Methodology – background

In this section, first some backgrounds have been presented about the mathematical methods used in this study. Then the proposed methodologies based on DEA, MDA and game theory are introduced.

2.1. Classic DEA

Based on Relative Performance Evaluation Theory, the performance of a specific DMU must be evaluated in the presence of its competitors. DEA is a group relative performance assessment and ranking method proposed by Charnes et al. (1978). The first DEA model, named classic CCR model, is as follows:

$$
\begin{aligned}
& Max \sum_{r=1}^{s} u_r y_{ro} \\
& s.t. \\
& \quad \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \qquad j=1,\dots,n \\
& \quad \sum_{i=1}^{m} v_i x_{io} = 1, \\
& \quad u_r > 0, \quad r=1,\dots,s, \\
& \quad v_i > 0, \quad i=1,\dots,m
\end{aligned} \tag{1}
$$

Model (1) obtain the efficiency score of DMU_O, where x_{ij} and y_{rj} are the i^{th} input and r^{th} output of DMU_j, respectively. Also u_r and v_i are the relative importance of each output, and input, respectively.

2.2. Multivariate DEA

There is a variety of phenomena in the world that scientific analysts tend to explain them. Most of them are affected by many different variables. In these cases that scientists need to elicit information from data including many observations on different variables, multivariate data analysis (MDA) could be a very useful methodology (Johnson and Wichern, 2007). By detection of interactions and relations between different variables of a phenomenon, MDA is a set of techniques used for exploring information from a data set in high dimensions, (Härdle and Simar, 2007; Murtagh and Heck, 2012). In this paper Principal Component Analysis (PCA) and clustering, which are two common MDA techniques, are combined with DEA to reduce the outputs and inputs dimensions. In the following, PCA and clustering are briefly explained.

2.2.1. PCA

PCA is a very common technique that can be used for reducing dimensions of a data set by producing uncorrelated linear combinations of original data set (Andrejić et al., 2016). Also PCA is an useful tool to interpretation (Johnson and Wichern, 2007). Based on the theory of PCA, a few uncorrelated linear combinations of original data that can explain the most percentage of variation in data set, can be considered instead of those lots of variables (Murtagh and Heck, 2012).

If X=[X_1,X_2,...,X_p] is a random variable with covariance matrix Σ and eigenvalue-eigenvector pairs $(\lambda_1,e_1),(\lambda_2,e_2),\ldots,(\lambda_p,e_p)$, $Y_1,Y_2,\ldots,Y_p$ as the principal components (PC) of $X_1,X_2,\ldots,X_p$, can be calculated as follows (Johnson and Wichern, 2007).

$$Y_j = e'_j X = e_{j1}X_1 + e_{j2}X_2 + \ldots + e_{jp}X_p,\ j = 1,2,\ldots p \tag{2}$$

where the following relations are established between each two different linear combination Y_k and Y_l:

$$Var(Y_k) = \lambda_k \tag{3}$$

$$Cov(Y_k, Y_l) = 0 \tag{4}$$

Therefore, the proportion of total population variance explained by k^{th} PC (pr_k) is:

$$pr_k = \frac{\lambda_k}{\lambda_1 + \lambda_2 + \ldots \lambda_p} \tag{5}$$

If m PC with the highest pr value are selected that could totally explain at least 0.75 proportion of variance ($pr^m = \sum_{[j]=1}^{m} pr_{[j]}$), those m PCs could be considered instead of p original variables. This is a rule of thumb and as it is mentioned by Johnson and Wichern (2007) and Jolliffe (2011), based on the analyst opinion and the context of the dataset the threshold value can be set higher than 0.75 or even less than 0.75.

In this research, PCA is used to reduce the number of inputs and outputs of DMUs. Then the classic CCR DEA model and game-DEA model have been applied for the new combinations of inputs and outputs.

2.2.2. Clustering

Clustering is an MDA technique used for grouping observations based on similarities. The main idea in clustering is that each member in a cluster have the most similarity with other members of that cluster, and have the most dissimilarity with the members in all other clusters (Faceli et al., 2007; Jain and Dubes, 1988). In cluster analysis, no assumption is considered for the number of groups or group structure and the clustering is performed only based on similarities. For example in k-mean clustering, the goal is to find K groups in the data, where the algorithm works iteratively to assign each data point to one of K groups based on the features that are provided. Data points are clustered based on feature similarity. The clustering algorithms can be classified in hierarchical and non-hierarchical methods. The Ward's clustering method is a criterion applied in hierarchical cluster analysis. Ward's minimum variance method is a special case of the objective function approach. In this method, ESS_j is defined as the sum of square deviations of members of cluster j from the cluster's center. Therefore, if there is k clusters, the following relation could be defined (Johnson and Wichern, 2007):

$$ESS = ESS_1 + ESS_2 + \ldots + ESS_K \tag{6}$$

It should be noted that if the whole N samples are placed in one cluster, the following relation is established (Johnson and Wichern, 2007):

$$ESS = \sum_{j=1}^{N} (x_j - \bar{x})'(x_j - \bar{x}) \tag{7}$$

where, x_j is the measurement associated with the j^{th} item and $\bar{x}$ is the mean of all samples.

In this study, we applied this technique for clustering input variables. Then based on the identified categories, game-DEA and classic CCR models are used to obtain the efficiency scores of each DMUs.

2.3. Game-DEA model

There is a variety of game-DEA models in the literature proposed for different purposes (Cook et al., 2010). Mahmoudi et al. (2018a) developed a game-DEA approach to evaluate the performance of DMUs in a case that inputs have been categorized into different groups. Their approach is useful to overcome the problem of insufficient number of DMUs, too. Indeed, they considered each category as a player in the bargaining game concept and the efficiency of each category as the utility of that category.

Suppose that there are n DMUs, which use two different types of inputs to produce outputs. Each DMU_j uses m_1 inputs from first type, denoted by x_{ij}^1 ($i = 1,\ldots, m_1$), and m_2

inputs from second type, denoted by x_{ij}^2 ($i = 1,\dots, m_2$), to produce s outputs, denoted by y_{rj} ($i = 1,\dots, m_1$). The proposed game-DEA model by Mahmoudi et al. (2018a) to evaluate the performance of a set of DMUs when inputs are categorized into two different types, is as follows:

$$
\begin{aligned}
\max\ & e_o = \left(\sum_{r=1}^{s} \mu_r^1 y_{ro} - \theta_1\right)\left(\gamma \sum_{r=1}^{s} \mu_r^1 y_{ro} - \theta_2\right) \\
s.t.\quad & \sum_{r=1}^{s} \mu_r^1 y_{ro} \geq \theta_1 \\
& \gamma \sum_{r=1}^{s} \mu_r^1 y_{ro} \geq \theta_2 \\
& \sum_{r=1}^{s} \mu_r^1 y_{rj} - \sum_{i=1}^{m_1} v_i^1 x_{ij}^1 \leq 0, \qquad j = 1,\dots,n \\
& \gamma \sum_{r=1}^{s} \mu_r^1 y_{rj} - \sum_{k=1}^{m_2} v_k^2 x_{kj}^2 \leq 0, \qquad j = 1,\dots,n \\
& \sum_{i=1}^{m_1} v_i^1 x_{io}^1 = 1 \\
& \sum_{k=1}^{m_2} v_k^2 x_{ko}^2 = 1 \\
& \gamma, \mu_r^1, v_i^1, v_k^2, > 0, \quad i = 1,\dots,m_1,\ k = 1,\dots,m_2,\ r = 1,\dots,s.
\end{aligned}
\tag{8}
$$

where e_o denotes the efficiency of DMU_o. θ_1 and θ_2 are the breakdown points of first type and second type of inputs, respectively, in Bargaining game model. Breakdown points are minimum achievable efficiency in each category. From the perspective of game theory, a DMU will withdraw from the game, if they obtain optimal efficiency scores lower than the breakdown points in each categories (Mahmoudi et al., 2014). Different methods are used in several studies to obtain breakdown points (Mahmoudi et al., 2018a; Tavana et al., 2018). For example Rezaee et al. (2012) proposed a method based on the cross-efficiency scores as follows:

a) Use the optimal weights of other DMUs to obtain the cross-efficiency of DMU_j by

$$E_{qj} = \frac{\sum_{r=1}^{s} u_{rq}^* z_{rj}}{\sum_{i=1}^{m} v_{iq}^* x_{ij}}, \quad q, j = 1,..,n \text{ and then } \bar{E}_j = \frac{1}{n}\sum_{q=1}^{n} E_{qj}.$$

b) Consider $\theta_{cross} = \inf_{q,j}(E_{qj})$ as the breakdown point.

We used this method to obtain the breakdown points. Also in this study, the game-DEA model (model (8)), as a useful model to overcome the problem of the insufficient number of DMUs, is applied to evaluate the performance of DMUs when inputs are classified into two groups.

2.4. An Algorithm for assessing the performance of TPPs when the number of DMUs is insufficient

According to the presented backgrounds, in this section a new algorithm is presented to evaluate the performance of TPPs when the number of DMUs is insufficient to meet the initial condition of DEA. For this purpose, DEA, MDA techniques and game theory have been integrated. The proposed algorithm is as follows (Note. S_i is Strategy i):

Step 1. Identifying inputs and outputs of the DMUs and collect related data.

Step 2. Based on the nature of the inputs, categorize the inputs in two initial categories, if it is possible. Name these categories as Cat1 and Cat2. For example, for TPPs, inputs can be classified into operational and non-operational inputs.

Step 3. Use PCA to identify PCs for inputs, Cat1 inputs, Cat2 inputs and outputs, separately.

Step 4. Calculate the minimum number of PCs for inputs ($\Omega_1 = Min_{In}^{PC}$), minimum number of PCs for outputs ($\Omega_2 = Min_{Out}^{PC}$), minimum number of PCs for inputs in Cat1 ($\Omega_3 = Min_{Cat1}^{PC}$) and minimum number of PCs for inputs in Cat2 ($\Omega_4 = Min_{Cat2}^{PC}$).

Step 5. Categorize inputs in proper number of groups using clustering approach (expert base).

Step 6. Determine efficiency of first and second cluster using Classic DEA, separately and calculate mean of two cluster efficiency scores for each DMU (S_1).

Step 7. Include all inputs and outputs and calculate the efficiency of each DMU using standard DEA model (S_2).

Step 8. Include all inputs and outputs and calculate the efficiency of each DMU using Game-DEA model (S_3).

Step 9. Determine efficiency of each DMU based on Ω_1 and normal outputs which satisfy the conditions $pr^{\Omega_1} \geq pr^*$ and $3(\Omega_1 + s) < n$ (if possible) using classic DEA model (S_4).

Step 10. Determine efficiency of each DMU based on normal inputs and Ω_2 which satisfy the conditions $pr^{\Omega_2} \geq pr^*$ and $3(m + \Omega_2) < n$ (if possible) using classic DEA model (S_5).

Step 11. Determine efficiency of each DMU based on Ω_1 and Ω_2 which satisfy the conditions $pr^{\Omega_1} \geq pr^*$, $pr^{\Omega_2} \geq pr^*$ and $3(\Omega_1 + \Omega_2) < n$ (if possible) using classic DEA model (S_6).

Step 12. Determine efficiency of each DMU based on two identified clusters in Step 4 using Game-DEA model (S_7).

Step 13. Determine efficiency of each DMU based on Ω_3, Ω_4, and normal outputs which satisfy the conditions $pr^{\Omega_3} \geq pr^*$, $pr^{\Omega_4} \geq pr^*$ and $3(\Omega_3 + \Omega_4 + s) < n$ (if possible) using Game-DEA model (S_8).

Step 14. Determine efficiency of each DMU based on Ω_3, Ω_4, and Ω_2 which satisfy the condition $pr^{\Omega_3} \geq pr^*$, $pr^{\Omega_4} \geq pr^*$, $pr^{\Omega_2} \geq pr^*$ and $3(\Omega_2 + \Omega_3 + \Omega_4) < n$ (if possible) using Game-DEA model (S_9).

It should be noted that different approaches could be used to calculate the mean values. It depends on the opinion of experts and researchers. Also the pr^* value is a managerial parameters which in most studies is assumed as 0.75 (Johnson and Wichern, 2007). In this paper the weighted geometric mean is used and it is assumed that $pr^* = 0.75$.

All of the proposed strategies (except S_2, classic DEA) are able to handle the problem of insufficient number of DMUs since these models are based on either decreasing the number of inputs (PCA) or categorizing the inputs (clustering) using game-DEA model or combinations of these. However, as it is mentioned before, different DEA models provide different efficiency scores and ranks. On the one hand, the results of a method cannot be ignored and on the other hand different efficiency scores and ranks for a specific DMU will make a problem for analyzing and identifying weaknesses of that DMU. Therefore, in the next section, an approach is proposed to obtain a unique rank according to the results of different strategies.

3. An MADM method to increase discrimination power of DMUs

If a decision maker tends to rank some alternatives based on some deferent criteria, multi attribute decision-making (MADM) methods will be helpful tools. In this paper, DMUs are assumed as alternatives and each method for finding their efficiency could be considered as a criterion. At first, the weight of each method is calculated based on Shannon Entropy method. Then, the DMUs are ranked based on TOPSIS, which is an MADM method.

One of the most important bases of MADM methods is the weights of criteria. There are several methods to determine the weight of criteria including Shannon Entropy, AHP, ANP, eigenvector, etc. The main advantage of Shannon Entropy is choosing higher weigh for the criterion with more variation in its scores. In DEA, a method with the most discrimination capability has the most variation in its scores. Therefore, in this paper Shannon Entropy method is applied. Below, Shannon Entropy and TOPSIS methods are briefly introduced.

3.1. Shannon Entropy method

Shannon Entropy method is a very common method for determining weights of criteria. This method doesn't need considering decision making's opinion. In other words, this method determines weights based on the amount of variation among the alternatives' values in a criterion. The more variations in values of criteria, the more criterion's weight (Qiu, 2002). This method could be performed based on the following explanations (Zou et al., 2006).

- If there are m alternatives and n criteria, the decision matrix is defined as the alternatives' values in different criteria and illustrated as follows:

$$D = [\mathrm{x}_{ij}]_{m\times n}, i = 1,2,\dots,m, j = 1,2,\dots,n \tag{9}$$

- The Shannon Entropy of i^{th} criterion is defined as follows:

$$H_j = -k\sum_{i=1}^{m} f_{ij}Ln(f_{ij}), j = 1,2,\dots,n \tag{10}$$

where f_{ij} and k is calculated as follows:

$$f_{ij} = \frac{x_{ij}}{\sum_{p=1}^{n} x_{ip}}, i = 1,2,\dots,m, j = 1,2,\dots,n \tag{11}$$

$$k = \frac{1}{Ln(m)} \tag{12}$$

- The weight of each criterion is calculated as follows.

$$w_j = \frac{1 - H_j}{m - \sum_{p=1}^{n} H_p}, j = 1,2,\dots,n \tag{13}$$

3.2. TOPSIS method

TOPSIS is one of the most used and popular MADM methods for ranking alternatives based on the values of criteria in each alternative. Ease of use and no need of decision maker's opinion are some of TOPSIS's advantages (Ishizaka and Nemery, 2013). However, there are a variety of methods such as AHP, VIKOR, LINMAP, ELECTRE that can be used. In this paper, TOPSIS is chosen and applied.

The algorithm of TOPSIS is given as follows (Ishizaka and Nemery, 2013):

A. The criteria's values for alternatives must be normalized by Euclidean normalization by the following relation.

$$D = [x_{ij}]_{m\times n} \rightarrow D_n = [r_{ij}]_{m\times n} \tag{14}$$

where

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{l=1}^{m} x_{lj}^2}}, i = 1,2,...,m, j = 1,2,...,n \tag{15}$$

B. The weighted normalized decision matrix is calculated as follows:

$$D_n = [x_{ij}]_{m\times n} \rightarrow V = [v_{ij}]_{m\times n} \tag{16}$$

where

$$v_{ij} = w_{ij} \times r_{ij}, i = 1,2,...,m, j = 1,2,...,n \tag{17}$$

C. By collecting the best and worst value in each criterion of weighted normalized decision matrix, the positive ideal vector A^+ and negative ideal vector A^- could be defined as follows:

$$A^+ = (v_1^+, v_2^+, ..., v_m^+) \tag{18}$$

$$A^- = (v_1^-, v_2^-, ..., v_m^-) \tag{19}$$

where $v_j^+ = \max_i \{v_{ij}\}$ and $v_j^- = \min_i \{v_{ij}\}$, if the criterion that must be maximized. In addition, $v_j^+ = \min_i \{v_{ij}\}$ and $v_j^- = \max_i \{v_{ij}\}$, if the criterion that must be minimized.

D. The distance of each alternative from the positive and negative ideal vector must be calculated as follows:

$$d_i^+ = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_j^+)^2}, i = 1,2,...,m \tag{20}$$

$$d_i^- = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_j^-)^2}, i = 1,2,...,m \tag{21}$$

E. The relative closeness coefficient of each alternative will be calculated as follows:

$$C_i = \frac{d_i^-}{d_i^- + d_i^+}, i = 1, 2, ..., m \quad (22)$$

F. The alternatives must be ranked based on the relative closeness coefficient. The more relative closeness coefficient of each alternative, the better rank for it.

3.3. An approach to obtain the final ranks

Based on the results of the different strategies and using Shannon Entropy and TOPSIS, in this section an approach is proposed to obtain a unique rank and score for each DMU and analyzing these values, as follows:

Step 1. Standardize efficiency score in each strategy.

Step 2. Consider each strategy as a criterion for ranking the DMUs and determine weight of each one by using Shannon entropy.

Step 3. Rank DMUs based on the relative closeness coefficient obtained by TOPSIS.

Step 4. Obtain the correlation between C (vector of the relative closeness coefficient) and inputs/outputs and analyze the results.

4. Performance assessment of Iranian TPPs

In this section, the proposed approaches introduced in the previous sections, are applied to the performance assessment of Iranian TPPs. Based on literature (Arabi et al., 2016; Azad et al., 2015; Coşgun and Kaya, 2014), 8 inputs and 3 outputs are considered which are introduced as follows:

Inputs

Input 1 (x_1): Generation capacity (MW)

Input 2 (x_2): The total hours of operation in a power plant per period (hour)

Input 3 (x_3): The internal consuming (MWh)

Input 4 (x_4): The fuel consumption amount (Terajoule)

Input 5 (x_5): The number of non-operational employees

Input 6 (x_6): The number of operational employees

Input 7 (x_7): The cost of generated power per kWh (monetary unit)

Input 8 (x_8): The total cost of training (billions of monetary unit)

Output

Output 1 (y_1): The total revenue (billions of monetary unit)

Output 2 (y_2): The total amount of electricity generated (MWh)

Output 3 (y_3): CO_2 emission (1000 ton)

where the CO_2 emission is an undesirable factor which has been converted to the desirable value using the proposed approach by Seiford and Zhu (2002). The value of CO_2 emission is obtained using eq. (23) (NRI, 2019):

$$E = A \times EF \times (1 - \frac{ER}{100}) \tag{23}$$

where E is the value of released emission and A is activity rate (value of generated electricity (KW/h)). EF is emission releasing factor which shows the value of released emission per each unit of generated electricity. ER is total captured emission parentages, ER is 0, if a TTP do not use an emission capture system. Also "the cost of generated power per kWh" is summation of fuel costs, operation and labor costs, maintenance costs and depreciation costs. Generation cost or operation cost have been widely used as an input in the previous studies including Barros (2008), Sözen et al. (2010) and Liu and Wen (2012).

The Inputs are categorized in two groups including operational inputs (x_1 , x_2 , x_3 and x_4) and non-operational inputs (x_5 , x_6 , x_7 and x_8). The data used for this application is reported in Table 1. It should be noted that all types of power plants in Iran have governmental structure. Detailed data related to power plants are categorized as national secure data by Iranian government. Some general information about the data can be found at https://amar.tavanir.org.ir/en/. Given that there are large number of variables as compared to number of DMUs, the initial condition of DEA is not established. Hence, the proposed approach is used to improve the discrimination power of DEA. In the next section, the results are presented as well as some managerial analysis.

Table 1. Data for Iranian thermal power plants (NRI, 2019).

Thermal power plant	Inputs								Output		
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	y_1	y_2	y_3
TPP_1	830	85,023.40	383,122	47	105	275	235.61	11.72	679.64	5,159,787	5297.16
TPP_2	225	72,910.20	123,128	15.69	88	192	330.63	3.18	179.09	1,354,178	7917.85
TPP_3	1280	82,106.20	521,124	61.81	192	470	302.26	18.08	965.43	7,329,520	3221.02
TPP_4	640	87,754.70	341,280	36.82	102	269	260.94	9.04	552.25	4,175,810	6149.3
TPP_5	757	60,704.60	344,217	39.97	110	272	230.61	10.69	572.76	4,330,898	5885.7
TPP_6	1830	74,592.10	595,521	86.27	214	478	158.06	2.29	1971.37	10,516,063	2011.51
TPP_7	1011	38,193.10	152,810	50.98	106	288	331.4	1.26	739.94	2,221,051	4554.1
TPP_8	1300	87,223.40	565,394	57.22	195	470	157.6	1.62	980.64	7,444,968	4441.88
TPP_9	444	53,002.95	140,031	19.64	96	260	256.5	0.55	288.63	2,191,235	7586.66

TPP_{10}	352	24,543.10	104,650	20.16	93	222	288.71	0.44	261.56	1,985,783	7543.28
TPP_{11}	1938	73,852.90	527,861	88.23	155	356	137.05	1.19	1325.15	10,060,484	2432.66
TPP_{12}	1958	67,793	618,690	97.67	158	357	147.01	7.52	1441.64	10,944,887	1057.5
TPP_{13}	400	40,037.80	89,828	14.31	95	252	274.31	1.54	194.76	1,478,648	8033.23
TPP_{14}	1000	91,828	413,842	49.44	99	313	250.38	3.84	762.24	5,468,079	5092.93
TPP_{15}	1600	88,986.70	823,033	78.33	131	455	190.54	6.14	1533.27	11,640,505	500
TPP_{16}	600	86,860.40	258,318	36.11	120	329	113.71	2.3	385.73	3,692,738	6208.68
TPP_{17}	894	86,464.40	404,306	42.02	101	256	246.54	3.43	927.14	5,790,639	5714.04
TPP_{18}	660	77,812.90	278,007	28.08	41	89	123	6.78	450.67	3,752,926	6880.48
TPP_{19}	740	75,002.40	215,593	35.74	98	220	303.33	7.6	434.44	3,298,268	7118.4
TPP_{20}	1323	80,235.10	423,452	69.71	133	450	185.45	13.6	888.07	6,742,225	6293.25
TPP_{21}	301	69,911.60	97,660	18.94	45	92	161.19	3.09	192.67	1,462,716	8111.46
TPP_{22}	1498	71,149.65	464,488	75.54	143	491	150	3.06	852.19	6,819,899	5830.48
TPP_{23}	960	82,500	406,448	42.15	100	340	144.4	1.96	720.85	5,472,634	6740
TPP_{24}	410	50,946.50	7593	15.56	47	113	316.01	0.84	169.76	1,288,818	8311.03

5. Results and discussion

Based on the number of inputs and outputs, at least 34 DMUs are needed to run the classic DEA model, but only data of 24 TPPs is available. To overcome this problem, the proposed approaches are applied.

Table 2 illustrates the total principal components (PCs) extracted from the PCA method applied to inputs. Each PC is a new attribute, which is a linear combination of inputs explaining a specific proportion of total variations. The proportions of PCs are shown in Table 3. If cumulative explained variance of a number of PCs is more than 0.75, these PCs are sufficient. From Table 3 it can be easily understood that two PCs must be selected for inputs since PC1 and PC2 cover about 0.76% of the total variations among inputs. It should be noted that by considering a large number of PCs, more cumulative explained variance will be covered, but in this study PCA is used to reduce the number of variables, as much as it is possible. The less number of variables, the more discrimination power of DEA models. Therefore, the minimum number of PCs (show this number by ϖ) which satisfy $pr^{\varpi} \geq pr^{*} = 0.75$, will be selected.

Table 2. The results of principal component analysis for inputs.

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
x_1	0.422	-0.19	-0.054	-0.307	-0.343	0.186	0.153	-0.716
x_2	0.254	0.525	0.458	0.529	-0.265	0.236	0.211	-0.017
x_3	0.429	0.077	0.074	0.017	-0.274	-0.509	-0.682	0.072
x_4	0.426	-0.131	-0.067	-0.338	-0.299	0.148	0.32	0.685

x_5	0.386	-0.159	-0.33	0.325	0.351	0.586	-0.375	0.059
x_6	0.399	-0.112	-0.257	0.361	0.343	-0.534	0.473	-0.072
x_7	-0.246	0.169	-0.72	0.291	-0.554	-0.012	0.025	0.008
x_8	0.152	0.774	-0.291	-0.437	0.315	0.007	0.002	-0.055

Table 3. Eigenvalues and cumulative values in principal component analysis for inputs.

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8
Eigenvalue	4.8882	1.1504	1.0518	0.4201	0.2933	0.1072	0.0795	0.0094
Proportion	0.611	0.144	0.131	0.053	0.037	0.013	0.01	0.001
Cumulative	0.611	0.755	0.886	0.939	0.975	0.989	0.999	1

Table 4 shows the results of PCA method on the outputs. As it is clear in the table, the first PC explains more than 0.86 proportions of original data's variance. Therefore, it could be acceptable to choose only one PC for outputs.

Table 4. The results of principal component analysis, eigenvalues and cumulative for outputs.

Variable	**PC1**	**PC2**	**PC3**		λ_1	λ_2	λ_3
y_1	0.545	-0.837	0.053	**Eigenvalue**	2.5866	0.3315	0.0819
y_2	0.596	0.342	-0.727	**Proportion**	0.8620	0.1100	0.027
y_3	0.590	0.428	0.685	**Cumulative**	0.8620	0.9730	1.0000

For clustering method, the inputs are divided into two different clusters using Ward clustering method. The Dendrogram of clustering is shown in Figure 1. Based on each cluster, the efficiencies are calculated. Finally, for each DMU, the total efficiency is obtained by calculating geometric mean of the efficiency scores of two clusters. In addition, these clusters have been considered as two different types of inputs in the Game-DEA model.

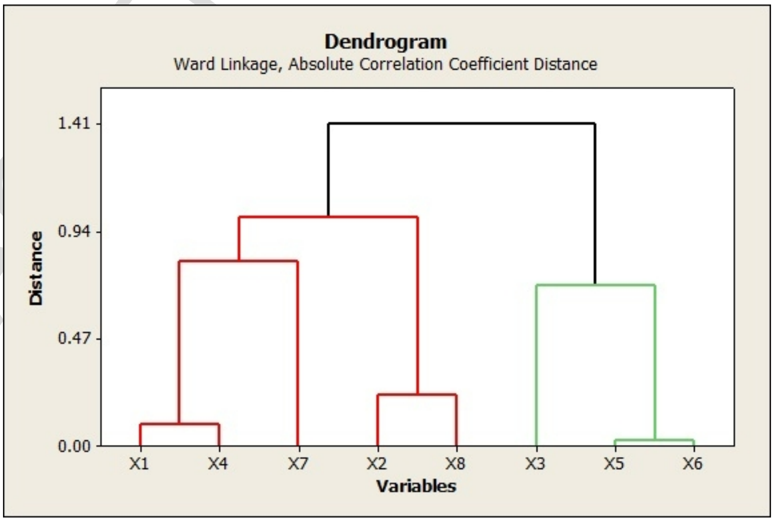


Figure 1. Dendrogram for inputs clustering.

To apply the strategies which include game-DEA model, first, the breakdown points must be calculated. Results for breakdown points in different strategies are $\theta_1 = 0.278$, $\theta_2 = 0.014$ for S_3 and S_8, $\theta_1 = 0.100$, $\theta_2 = 0.088$ for S_7, $\theta_1 = 0.118$, $\theta_2 = 0.014$ for S_9. Since used inputs and outputs in each strategy are different, obtained breakdown points in each strategy are different. By setting the breakdown points, identifying PCs and clusters, and running the proposed strategies, the efficiencies of DMUs are calculated and the results are given in Table 5.

Many interesting facts can be extracted from Table 5. S_2 is the classic DEA model applied for 11 original inputs and outputs. As it is clear from the results, this model has obtained efficiency score 1.00 for 18 DMUs from 24 DMUs (18 DMUs are identified as efficient ones), which means this model do not have discrimination power and the calculated results are not meaningful for analyzing. The results of other strategies show that the proposed approaches in this study can overcome this shortage of classic DEA model satisfactorily. For example, the classic DEA model is suggested 1.00 for the efficiency score of DMU_2, where the average efficiency score obtained by the proposed approaches for this DMU is about 0.45. Although the results of the other strategies show a good dispersion of efficiency score and discrimination power.

Table 5. Efficiency scores of TPPs in different scenarios.

DMU	S_1	S_2	S_3	S_4	S_6	S_7	S_8	S_9
1	0.792	0.933	0.932	0.660	0.644	0.600	0.376	0.356
2	0.566	1.000	1.000	0.374	0.375	0.266	0.315	0.247
3	0.792	0.896	0.880	0.759	0.748	0.238	0.388	0.391
4	0.758	0.984	0.913	0.567	0.563	0.486	0.472	0.331
5	0.710	0.978	0.932	0.634	0.633	0.410	0.385	0.338
6	1.000	1.000	0.998	1.000	0.960	1.000	0.686	0.560
7	1.000	1.000	1.000	1.000	0.673	1.000	0.971	0.375
8	0.804	1.000	0.934	0.680	0.704	0.613	0.477	0.344
9	0.681	1.000	0.859	0.638	0.639	0.418	0.498	0.381
10	0.765	1.000	0.955	0.894	0.896	0.501	0.503	0.506
11	1.000	1.000	0.957	1.000	1.000	1.000	0.486	0.577
12	1.000	1.000	0.989	1.000	1.000	1.000	0.518	0.545
13	0.545	1.000	1.000	0.631	0.633	0.247	0.472	0.403
14	0.786	0.875	0.801	0.648	0.631	0.522	0.391	0.368
15	1.000	1.000	0.971	0.808	0.803	1.000	0.522	0.415
16	0.744	1.000	0.883	0.636	0.607	0.624	0.367	0.360
17	0.959	1.000	0.989	0.712	0.691	0.910	0.465	0.404
18	0.948	1.000	1.000	0.607	0.607	0.888	0.474	0.356
19	0.646	0.862	0.832	0.642	0.643	0.372	0.516	0.385
20	0.754	1.000	0.901	0.789	0.789	0.530	0.572	0.445
21	0.676	1.000	0.975	0.465	0.466	0.413	0.361	0.319
22	0.744	1.000	0.866	0.758	0.786	0.516	0.530	0.393

23	0.858	1.000	0.919	0.657	0.657	0.712	0.478	0.375
24	0.770	1.000	0.999	1.000	1.000	0.549	1.000	1.000

Note. S_5 is an impossible strategy for this case study.

From Table 5 it is clear that the obtained ranks and scores by different strategies are different. However, all strategies have a specific result in common: most of Iranian TPPs are inefficient. To obtain a unique rank for each DMU, the proposed Shannon entropy and TOPSIS method is applied for the efficiencies obtained in different strategies. At first, based on Shannon entropy method, weights of each strategy is calculated for all strategies. In MADM for ranking DMUs, more variation among data in a criterion shows the importance of that criterion (Zou et al., 2006). The results of the Shannon entropy method are given in Table 6. From Table 6, among the all strategies the biggest weight is for S_7 because the amount of variation among the DMUs' values in this strategy is the highest one. Further, another evidence for inability of classic DEA model to discriminate between efficient and inefficient is the entropy score of this model in DMUs as reported in Table 6. The weight of S_2 is 0.004, less than all other scenarios, which shows very low variation among the results of this scenario. The relative closeness coefficient and DMUs' ranks calculated by TOPSIS methods are presented in

Table 7. Based on TOPSIS method, the best DMU is the one which has the highest C_i value and the DMUs are ranked based on the values of C_i. As it is clear in

Table 7, the best DMU is DMU_6 and the worst one is DMU_2. These ranks are in accordance with the efficiency scores presented in Table 5. As seen in this table, DMU6 has an efficiency score about 1 in 6 scenarios where DMU_2 is the weakest or one of the weakest DMUs in 6 scenarios. From

Table 7, it can be obviously seen how powerful the proposed approach in discrimination between DMUs is. Ranks and scores of the DMUs are completely different and there is not any group of DMUs with similar rank. Therefore, the proposed method is capable of overcoming the shortage of classic DEA model in evaluating the performance of a set of DMUs in the real world cases.

Table 6. Weights of the scenarios based on Shannon entropy.

	S_1	S_2	S_3	S_4
Weight	0.058	0.004	0.008	0.112
	S_6	S_7	S_8	S_9
Weight	0.103	0.347	0.181	0.187

Table 7. The relative closeness coefficient values obtained by TOPSIS.

	DMU_1	DMU_2	DMU_3	DMU_4	DMU_5	DMU_6	DMU_7	DMU_8
C_i	0.4141	0.1393	0.3087	0.3695	0.3300	0.5935	0.5703	0.4275
Rank	14	24	22	17	20	1	4	12
	DMU_9	DMU_{10}	DMU_{11}	DMU_{12}	DMU_{13}	DMU_{14}	DMU_{15}	DMU_{16}
C_i	0.3581	0.4315	0.5720	0.5700	0.3077	0.3863	0.5431	0.4212
Rank	18	10	3	5	23	16	6	13
	DMU_{17}	DMU_{18}	DMU_{19}	DMU_{20}	DMU_{21}	DMU_{22}	DMU_{23}	DMU_{24}
C_i	0.5176	0.5040	0.3452	0.4308	0.3118	0.4106	0.4630	0.5787
Rank	7	8	19	11	21	15	9	2

After obtaining the final scores and ranks of TPPs, the results must be analyzed to identify possible solutions to improve their performance and having efficient, cleaner and sustainable production activities. For this purpose, we used correlation analysis to establish if there are possible connections between variables (Cohen et al., 2014). Based on this, it can be understood how much each input/output affect the scores of DMUs. Then according to correlation values, by identifying which inputs and outputs have high positive or negative correlation with final scores, potential solutions to decrease/increase these inputs/outputs can be suggested.

Table 8 shows the correlation analysis between C and inputs/outputs. A primary analysis of the results illustrate that the relative closeness coefficient has the highest correlation with the total revenue (y_1) among the outputs. However, the correlation values for all outputs are approximately equal and between 0.5 and 0.6, which shows significantly strong correlation between all outputs and final scores of DMUs. This means any policy which improves the performance of TPPs in producing outputs, will improve their performance markedly. It should be noted that the desirable values of CO_2 emissions are considered in correlation calculation. Therefore, improving these outputs will happen by increasing desirable values or equally decreasing undesirable values (level of emission). Any applied strategies must increase the revenue and generated electricity values of TPPs. Also, among the inputs, the most negative correlation with the relative closeness coefficient is for "the cost of generated power per kWh (x_7)". "The total hours of operation in a power plant" and "the total cost of training" are two other inputs which have negative correlations. Other inputs have positive correlation values. Due to these correlation values, it can be concluded that which inputs must be decreased, and which ones must be increased to improve the performance of TPPs.

Table 8. The correlation between C and variables.

Model	x_1	x_2	x_3	x_4	x_5	x_6
C	0.5832	-0.0248	0.3912	0.5373	0.2131	0.2106
Model	x_7	x_8	y_1	y_2	y_3	
C	-0.3624	-0.2356	0.5967	0.5152	0.5370	

Note: C is the vector of the relative closeness coefficient of DMUs

In particular, the large positive correlation value of "total revenue" implies that the best way to increase the efficiency of TPPs is increasing their revenue. There are two possible solutions to achieve this target: decreasing the generation cost or increasing the supply price. In Iran, most of the TPPs have governmental structure. The government applies the subsidiary for energy consumption because Iran is one of the biggest energy producers in the world. In fact, Iran's energy tariffs are very low compared to international prices. According to Iranian Energy balance sheet (IEEO, 2006), in 2006, the total subsidiary imposed to energy was more than 41 billion dollars. This policy of Iranian government leads to lower price of energy and consequently lower revenue for the TPPs. According to correlation values of inputs and outputs illustrated in Table 8 and the structure of Iranian TPPs, three different short-term strategies can be suggested for improving the efficiency which are explained as follows:

a) Based on the results of Table 8, correlation between C and total revenue is the highest amount among outputs. Therefore, all factors related to revenue including cost and price are the most important parameters to investigate for receiving more efficiency ($Corr(C, y_1) = 0.5967$). Consequently, if power plants managements want to keep governmental structure, decreasing the energy subsidiary and closing the energy consumption price to the international prices, could be efficient policies to increase revenue. In addition to increasing the price, the management can apply strategies that reduce production cost.

b) According to the results of different scenarios (Table 5) most of Iranian TPPs are inefficient. All evaluated TPPs have governmental structure. Therefore, the managers of power plants are not allowed to make significant changes in non-operational inputs especially number of operational and non-operational employees and prices. As another suggestion for improving revenue factor, the power plants can move from governmental structure toward private. Privatizing the power plants will lead to significant benefits. For example, private TPPs will establish prices based on the international prices without government subsidiaries. Real prices will increase the revenues of TPPs. In addition, the correlation between C and non-operational/operational employees are positive values. Therefore, privatizing can increase the efficiency scores of power plants by improving the non-operational factors including employees' performance.

c) Another way to improve the performance of power plants is decreasing generation costs. There are noticeable evidences in Table 8 that confirm this suggestion including: the positive correlation between C and "the number of non-operational employees" and "the number of operational employees", and negative correlation between C and "the cost of generation", and "the total cost of training".

It is extracted from Table 8 that the correlation between C and "generation capacity" as well as "the fuel consumption amount" are high values. Hence, another way to improve the power plants efficiency is increasing the "generation capacity" or/and "the fuel consumption amount". It should be noted that both of these inputs are related to the power plants' structure and production systems. The other structural factors are "The total amount of electricity generated" and "CO_2 emission". To decrease the CO_2 emission, the power plants management board can decide to install the Carbon capture systems. Reforming the generation structure is a long-time process. Therefore, as long time strategies, following recommendations could be suggested:

a) Increasing the generation capacity and decreasing production cost by updating the generation system or turning the old system to new sustainable and cleaner system. The average age of Iranian thermal power plants is more than 30 years old, and all production systems at least 30 years old technologies.

b) Although the cost of fossil fuels in Iran is still low, by considering financial and other supports such as subsides, the government can promote green electricity production technologies such as wind and solar systems. This strategy can lead to significant effects in the future of energy production in Iran. Fewer fossil fuels will be used and fewer emission will be released. Even more fuel can be saved for expert or internal uses.

c) Decreasing the fuel consumption amount by improving the generation system.

d) None of power plants has installed any emission reduction system. It seems that installing the carbon capture systems to decrease the pollution emissions is an essential activity for power plants. As a result, this study aims to draw the attention of managers and government to the environmental issues in the thermal power plants. As it is clear from Table 8, the correlation between C and "CO_2 emission" is significantly high. Hence, to achieve sustainability targets and have an efficient TPPs, using a cleaner production system is an essential for Iranian TPPs.

6. Conclusion

Number of DMUs is a crucial initial condition for conducting an evaluating process by DEA models. In the most real world cases, especially in the national cases or large scale industries, the existing number of DMUs is insufficient. This study suggested different integrated game-DEA, MDA-DEA and game-MDA-DEA methods to overcome this problem. Different DEA models always provide different efficiency scores and ranks for a set of DMUs. An MADM method is proposed to obtain a unique rank and score for a set of DMUs when different scores are obtained by different methods.

The proposed methods are used to evaluate the performance of Iranian TPPs and the results are analyzed deeply. Based on the results increasing the revenue of the TPPs is the most effective strategy to improve their performance. For this purpose, privatization and stopping governmental subsidiaries to have a real power price is an efficient policy. Another suggestion is installing new and modern production systems. On the one hand, cleaner and modern generation systems will increase generation capacity and decrease generation costs, on the other hand, they will release less CO_2 emission. Emission level is one of the most important criteria in evaluating sustainability of a DMU. Therefore, decreasing CO_2 by modern generations system or installing Carbon capture systems (or other similar systems) can make the performance of TPPs more efficient and sustainable.

The results showed that most of Iranian TPPs have insufficient performance. Where energy sector is a critical sector for the local and national governments, insufficient performance of the TPPs can make serious environmental, economic and even social problems. Iranian government can enjoy the results and proposed methods in this study to improve the performance of TPPs and achieve a sustainable situation. In addition to the results of this study, the authors encourage the government to apply proposed methods on the updated data by considering other possible inputs or outputs, those ones the authors did not have access to them because of security reasons.

7. Directions for future researches

The proposed integrated methods are applied to evaluate the performance of TPPs. However, for future researches, these methods could be considered for other real case studies that the number of DMUs is not sufficient. It should be noted that in many real cases this problem exists, and the results showed that proposed approach can handle this problem very well. An integrated Shannon entropy and TOPSIS method is proposed to obtain unique ranks and scores based on the results of different methods. Future studies can develop different MADM methods to present unique rank based on different results and compare their results to ours. Cross-sectional data are considered in this study. Using panel data, customizing proposed strategies for panel data and analyzing results can be other interested research directions. In this study 11 inputs/outputs are considered. The authors did not have access to a more data. However, considering more inputs/outputs will lead to more accurate results. Hence future studies can apply the proposed method to include more inputs and outputs.

Acknowledgements

The authors would like to thank Professor *Giorgio Besagni*, the associate editor of *Journal of Cleaner Production*, and three anonymous reviewers for their insightful comments and suggestions. As results this paper has been improved substantially.

References

Andrejić, M., Bojović, N., Kilibarda, M., 2016. A framework for measuring transport efficiency in distribution centers. Transport Policy 45, 99-106.

Arabi, B., Munisamy, S., Emrouznejad, A., Toloo, M., Ghazizadeh, M.S., 2016. Eco-efficiency considering the issue of heterogeneity among power plants. Energy 111, 722-735.

Azad, M.A.M., Covich, M.P., Sojoodi, S., 2015. The impact of electricity competitive market establishment on technical efficiency of thermal power plants in Iran. International Journal of Energy Economics and Policy 5(4), 1010-1015.

Barros, C.P., 2008. Efficiency analysis of hydroelectric generating plants: A case study for Portugal. Energy Economics 30(1), 59-75.

Bi, G.-b., Shao, Y., Song, W., Yang, F., Luo, Y., 2018. A performance evaluation of China's coal-fired power generation with pollutant mitigation options. Journal of Cleaner Production 171, 867-876.

Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. European journal of operational research 2(6), 429-444.

Cohen, P., West, S.G., Aiken, L.S., 2014. Applied multiple regression/correlation analysis for the behavioral sciences. Psychology Press.

Cook, W.D., Liang, L., Zhu, J., 2010. Measuring performance of two-stage network structures by DEA: a review and future perspective. Omega 38(6), 423-430.

Coşgun, Ö., Kaya, G.O., 2014. Data Envelopment Analysis Application in Turkish Energy Market. International Journal of Computational Intelligence Systems 7(4), 636-649.

Emrouznejad, A., Yang, G.-l., 2018. A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. Socio-Economic Planning Sciences 61, 4-8.

Faceli, K., de Carvalho, A.C., de Souto, M.C., 2007. Multi-objective clustering ensemble. International Journal of Hybrid Intelligent Systems 4(3), 145-156.

Friedman, L., Sinuany-Stern, Z., 1998. Combining ranking scales and selecting variables in the DEA context: The case of industrial branches. Computers & Operations Research 25(9), 781-791.

Härdle, W., Simar, L., 2007. Applied multivariate statistical analysis. Springer.

IEEO, 2006. Iran energy efficiency organization, Iranian energy balance sheet. http://en.saba.org.ir/en/home. 2017).

Ishizaka, A., Nemery, P., 2013. Multi-criteria decision analysis: methods and software. John Wiley & Sons.

Jain, A.K., Dubes, R.C., 1988. Algorithms for clustering data. Prentice-Hall, Inc.

Jin, M., Shi, X., Emrouznejad, A., Yang, F., 2018. Determining the optimal carbon tax rate based on data envelopment analysis. Journal of Cleaner Production 172, 900-908.

Johnson, R.A., Wichern, D.W., 2007. Applied multivariate statistical analysis. Prentice hall Upper Saddle River, NJ.

Jolliffe, I., 2011. Principal Component Analysis, in: Lovric, M. (Ed.) International Encyclopedia of Statistical Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1094-1096.

Lam, P.-L., Shiu, A., 2001. A data envelopment analysis of the efficiency of China's thermal power generation. Utilities Policy 10(2), 75-83.

Liu, C., Lin, S.J., Lewis, C., 2010. Evaluation of thermal power plant operational performance in Taiwan by data envelopment analysis. Energy policy 38(2), 1049-1058.

Liu, X., Wang, B., Du, M., Zhang, N., 2018. Potential economic gains and emissions reduction on carbon emissions trading for China's large-scale thermal power plants. Journal of Cleaner Production 204, 247-257.

Liu, X., Wen, Z., 2012. Best available techniques and pollution control: a case study on China's thermal power industry. Journal of Cleaner Production 23(1), 113-121.

Mahmoudi, R., Emrouznejad, A., Rasti-Barzoki, M., 2018a. A bargaining game model for performance assessment in network DEA considering sub-networks: a real case study in banking. Neural Computing and Applications.

Mahmoudi, R., Emrouznejad, A., Shetab-Boushehri, S.-N., Hejazi, S.R., 2018b. The origins, development and future directions of Data Envelopment Analysis approach in transportation systems. Socio-Economic Planning Sciences.

Mahmoudi, R., Hafezalkotob, A., Makui, A., 2014. Source selection problem of competitive power plants under government intervention: a game theory approach. Journal of Industrial Engineering International 10(3), 1-15.

Mahmoudi, R., Rasti-Barzoki, M., 2017. Sustainable supply chains under government intervention with a real-world case study: An evolutionary game theoretic approach. Computers & Industrial Engineering.

Munisamy, S., Arabi, B., 2015. Eco-efficiency change in power plants: using a slacks-based measure for the meta-frontier Malmquist–Luenberger productivity index. Journal of Cleaner Production 105, 218-232.

Murtagh, F., Heck, A., 2012. Multivariate data analysis. Springer Science & Business Media.

NRI, 2019. Niroo Research Institue. https://www.nri.ac.ir/en. 2019).

Omrani, H., Alizadeh, A., Emrouznejad, A., 2018. Finding the optimal combination of power plants alternatives: A multi response Taguchi-neural network using TOPSIS and fuzzy best-worst method. Journal of Cleaner Production 203, 210-223.

Qiu, W., 2002. Management decision and applied entropy. Beijing: China Machine Press.

Rezaee, M.J., Moini, A., Makui, A., 2012. Operational and non-operational performance evaluation of thermal power plants in Iran: A game theory approach. Energy 38(1), 96-103.

Sahoo, N.R., Mohapatra, P.K.J., Sahoo, B.K., Mahanty, B., 2017. Rationality of energy efficiency improvement targets under the PAT scheme in India – A case of thermal power plants. Energy Economics 66, 279-289.

Sarıca, K., Or, I., 2007. Efficiency assessment of Turkish power plants using data envelopment analysis. Energy 32(8), 1484-1499.

Seiford, L.M., Zhu, J., 2002. Modeling undesirable factors in efficiency evaluation. European journal of operational research 142(1), 16-20.

Sözen, A., Alp, I., Özdemir, A., 2010. Assessment of operational and environmental performance of the thermal power plants in Turkey by using data envelopment analysis. Energy Policy 38(10), 6194-6203.

Tavana, M., Khalili-Damghani, K., Santos Arteaga, F.J., Mahmoudi, R., Hafezalkotob, A., 2018. Efficiency decomposition and measurement in two-stage fuzzy DEA models using a bargaining game approach. Computers & Industrial Engineering 118, 394-408.

Wang, K., Wei, Y.-M., Huang, Z., 2017. Environmental efficiency and abatement efficiency measurements of China's thermal power industry: A data envelopment analysis based materials balance approach. European Journal of Operational Research.

Yadav, V.K., Jha, D.K., Kumar, N., Rajagopal Peesapatib, V.V., 2013. Benchmarking of Indian thermal power plants, IEEE Power and Energy Society General Meeting.

Yang, L., Wang, K.-L., Geng, J.-C., 2018. China's regional ecological energy efficiency and energy saving and pollution abatement potentials: An empirical analysis using epsilon-based measure model. Journal of Cleaner Production 194, 300-308.

Zou, Z.-h., Yun, Y., Sun, J.-n., 2006. Entropy method for determination of weight of evaluating indicators in fuzzy synthetic evaluation for water quality assessment. Journal of Environmental Sciences 18(5), 1020-1023.

Highlights

- Multistage data analysis and game theory are combined with DEA to overcome the problem insufficient number of DMUs
- An approach is proposed to integrate the results of different DEA models
- The proposed approaches are applied to evaluate the performance of Iranian thermal power plants
- The power plants should move from governmental structure toward becoming private
- New cleaner production systems should be installed