J Ind Eng Int (2015) 11:199–208
DOI 10.1007/s40092-014-0096-0

ORIGINAL RESEARCH

Robust DEA under discrete uncertain data: a case study of Iranian electricity distribution companies

Ashkan Hafezalkotob · Elham Haji-Sami · Hashem Omrani

Received: 30 January 2014 / Accepted: 30 June 2014 / Published online: 18 December 2014

Abstract Crisp input and output data are fundamentally indispensable in traditional data envelopment analysis (DEA). However, the real-world problems often deal with imprecise or ambiguous data. In this paper, we propose a novel robust data envelopment model (RDEA) to investigate the efficiencies of decision-making units (DMU) when there are discrete uncertain input and output data. The method is based upon the discrete robust optimization approaches proposed by Mulvey et al. (1995) that utilizes probable scenarios to capture the effect of ambiguous data in the case study. Our primary concern in this research is evaluating electricity distribution companies under uncertainty about input/output data. To illustrate the ability of proposed model, a numerical example of 38 Iranian electricity distribution companies is investigated. There are a large amount ambiguous data about these companies. Some electricity distribution companies may not report clear and real statistics to the government. Thus, it is needed to utilize a prominent approach to deal with this uncertainty. The results reveal that the RDEA model is suitable and reliable for target setting based on decision makers (DM's) preferences when there are uncertain input/output data.

Keywords Data envelopment analysis · Discrete uncertain data · RDEA · Robust optimization

Introduction

In the highly competitive and dynamic markets derived from globalization, the domestic firms should find a competitive edge that enables them to survive in the market. Moreover, limited natural resources and growing environmental concerns and regulations about production processes are new considerations influence the firms' operations. Therefore, the operational efficiencies would play an important role in survival and growth of firms. Especially in electricity distribution companies, operational efficiency is the most crucial issue among regulators (Sadjadi and Omrani 2008).

Data envelopment analysis (DEA) is a well-known nonparametric technique that measures the relative operational efficiency of similar decision-making units (DMUs). The most important capability of DEA is its ability to compare several parameters (inputs/outputs) concurrently and sum up them into a scalar measure of relative efficiency. The efficiencies of DMUs are obtained from weights corresponding to each input and output that computed through the optimal solution of linear programming (LP) problems. In fact, DEA is a data-oriented method for measuring and benchmarking the relative efficiency of peer DMUs. Target setting and improvement of DMU's performance are important features of DEA technique. There are several successful real-world applications of DEA method in different public and private sector industries such as banks, software development, health care, pharmacies, auto manufacturing, fisheries and search engines (Saranga and Phani 2009). Sadjadi and Omrani (2008), for instance, used

A. Hafezalkotob (✉) · E. Haji-Sami
Industrial Engineering College, Islamic Azad University, South Tehran Branch, Entezari Alley, Oskoui Alley, Choobi Bridge, 1151863411 Tehran, Iran
e-mail: a_hafez@azad.ac.ir; Hafezalkotob@iust.ac.ir

E. Haji-Sami
e-mail: paieez_sami@yahoo.com

H. Omrani
Department of Industrial Engineering, Urmia University of Technology, Band road, 57155-419 Urmia, Iran
e-mail: omrani57@iust.ac.ir; h.omrani@uut.ac.ir

DEA method for measuring the relative efficiency of energy companies in Iran. Roghanian and Foroughi (2010) implemented DEA to compare efficiencies of all regional and international airports in Iran using different input/output data. Goto and Tsutsui (1998) employed DEA approach to measure overall cost and technical efficiencies between Japanese and US electricity power plants. Saranga and Phani (2009) employed non-parametric DEA models and parametric methods such as regression analysis to specify the factors that have contributed to the internal operational efficiencies of firms in Indian pharmaceutical industry.

One of the most important issues associated with DEA is the uncertainty associated with the data. Since the resulted formulation of DEA technique is in form of LP, one can use traditional sensitivity analysis when there are one or a few uncertain parameters. However, when all input data are subject to uncertain, it is practically impossible to use sensitivity analysis method to handle all uncertainties. There are several methods for estimating the efficiencies of DMUs under data uncertainty.

In the real-world problems, data are often contaminated by perturbations (Izadi and Kimiagari 2014; Khalaj et al. 2013; Hosseini and Tarokh 2011; Shad et al. 2014). Because of perturbations in data, the efficient frontier in DEA is changed and the determined targets may become incorrect. Thus, the correction in the proposed target setting models would be necessary such that perturbation in inputs and outputs data would be considered (Monfared and Safi 2013; Bashiri et al. 2013). In a survey study, Ben-Tal and Nemirovski (2000) showed that a small perturbation on data could lead to infeasible solutions for some benchmark optimization problems. On the other hand, Bertsimas and Sim 2003, 2004; Bertsimas and Thiele 2006 and Bertsimas et al. (2004) developed new LP to adjust the robustness of the model against conservatism of the solution. In our LP reformation of DEA model, the results of the efficiency estimation and target setting could be unreliable in many cases especially when the efficiency of a particular firm is close to another. Mulvey et al. (1995) suggested an alternative approach, which is called scenario-based robust optimization (RO). This approach integrates goal programming formulations with a scenario-based description of problem data. It is a series of solutions of the model data from a scenario set. This motivates us to use robust DEA model to achieve more reliable results.

Sadjadi and Omrani (2008) developed a DEA model based on robust optimization approach and proposed a new formulation of DEA which is more reliable for efficiency estimating and ranking strategies. Also, they showed that the robust DEA founded upon Bertsimas and Sim 2003, 2004; Bertsimas and Thiele 2006) and Bertsimas et al. (2004) is easier and more applicable than robust DEA based on Ben-Tal and Nemirovski (2000) approach. Robust optimization generally refers to the modeling of optimization problems with uncertain data to obtain a solution that is guaranteed to be good and feasible for all or most possible uncertain parameters Bashiri and Moslemi (2013). Uncertainty in the parameters is containing all (or most) possible values that may be realized for the corresponding parameters. Shokouhi et al. (2010) proposed DEA under uncertainty which was based on a robust optimization model that input and output parameters were constrained to be within an uncertainty set. They applied Monte Carlo simulation to compute the conformity of the ranking in the RDEA model.

Morita (2003) developed a method using DEA which dealt with the use of non-parametric production frontiers and did not require cost information on inputs and outputs for identifying the economies of scope. The most robust multipliers have been defined for evaluation of the dominance relation of efficient frontiers. Foroughi and Aouni (2012) determined efficiency based on DEA with interval data and setting up a full ranking of DMUs in two phases. At first, interval efficiencies have been computed; afterwards, they combined the lower and upper bounds of the interval efficiencies. Hatami-Marbini et al. (2012) developed a fuzzy DEA framework with a Banker, Charnes and Cooper (BCC) model for measuring crisp and interval efficiencies using alfa-level approach to convert BCC model into an interval programming model.

Our paper is closely related to Sadjadi and Omrani (2008). They studied robust DEA model under continuous uncertain data. To the best of authors' knowledge, no research was found that considers the discrete uncertainty regarding input and output data of DEA. Therefore, there are two main contributions in this study. For the first and foremost, we extend DEA model to scenario-based description of the uncertain data. Using RO approach of Mulvey et al. (1995), we develop DEA formulation to consider discrete uncertainty in input and output parameters as a set of possible scenarios. In the second place, we explore the effect of the discrete uncertain data on the degree of operational efficiency achieved by the Iranian electricity distribution companies (Satapathy and Mishra 2013).

The rest of the article is organized as follows. In "Data envelopment analysis (DEA)", the background of DEA approach has been described. In "Robust optimization", scenario-based robust approach based on Mulvey robust optimization has been expressed briefly. In "Robust DEA based on Mulvey approach", we formulate robust DEA model based on Mulvey approach. A real numerical example demonstrates the efficacy of the model in Iranian

Electricity Company in "Case study". At the end, concluding remarks and some directions for future research are given in "Conclusion".

Data envelopment analysis (DEA)

DEA is a non-parametric approach which determines a piecewise linear efficiency frontier along the most efficient companies (DMUs) and derives the relative efficiency measures for all other companies (DMUs). The method was first introduced by Charnes et al. (1978) and has been widely implemented by many researchers in various sectors. DEA identifies an efficient frontier made up of the best practice DMUs to measure the relative efficiency scores of the less efficient DMUs. We choose an input-oriented approach of DEA to adjust the output by changing the input parameters such that the efficiency is maximized.

Assume that n DMUs should be evaluated by DEA method where each DMU has m input and t output data. Let x_{ij} denotes ith input and y_{tj} represents tth output of DMU j. Moreover, let u_r and v_i be the dual variables associated with x_{ij} and y_{tj}, respectively. The fractional DEA model is formulated as follows:

$$\max z = \frac{\sum_{r=1}^{t} u_r y_{r0}}{\sum_{i=1}^{m} v_i x_{i0}}, \tag{1}$$

s.t.

$$\frac{\sum_{r=1}^{t} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \le 1, \qquad j = 1, \ldots, n, \tag{2}$$

$$u_r, v_i \ge 0.$$

Model (1–3) is a non-linear programming problem (NLP). The DEA model is solved n times to determine the relative efficiencies of different DMUs. Since model (1–4) is an NLP problem, Charnes et al. (1983) recommended a simple modification of the objective function to linearize the problem as follows:

$$\max z = \sum_{r=1}^{t} u_{r0} y_{r0}, \tag{3}$$

s.t.

$$\sum_{r=1}^{t} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \tag{4}$$

$$\sum_{i=1}^{m} v_i x_i = 1, \tag{5}$$

$$u_r, v_i \ge 0, \quad \forall j = 1, 2, \ldots, n.$$

LP problem (3–5) has been widely used for the past three decades and the results have been commonly accepted as measure of relative efficiencies of different DMUs. However, when there is uncertainty with regard to the inputs and the outputs data, specific techniques should be used to make sure that small changes in input/output data do not alter the resulted rankings.

Robust optimization

Classical modeling approaches in operation research under uncertainty assume full probabilistic characterizations. The learning which is needed to implement the policies derived from these models is accomplished either through classical statistical estimation procedures or subjective Bayesian priors. However, in many models, the uncertainty is ignored altogether, and a representative nominal value of the data is used simply (e.g., expected values). The classical approach to deal with uncertainty is stochastic programming (SP). Recently, RO is introduced as a complementary alternative to sensitivity analysis and SP. Indeed, RO, while not without limitations, has some pros over stochastic LP and it is more generally applicable. Soyster (1973) proposed the highest protection model of the nominal linear optimization problem which is the most conservative in practice in the sense of the robust solution. Ben-Tal and Nemirovski (2000) assumed that the true values of uncertain data entries in ith inequality constraint are obtained from the nominal values of the entries by random perturbations.

The need for robustness has been recognized in a number of application areas. Mulvey et al. (1995) dealt with optimization problems that have two distinct components: a structural component that is fixed and free of any noise in its input data, and a control component that is subjected to noise in its input data. Then, they introduced two sets of variables to formulate such problems:

$x \in R^{n_1}$, represents the vector of decision variables that their optimal values are not dependent upon the realization of the uncertain parameters. They are also called design variables that cannot be adjusted once a specific realization of the data is observed.

$y \in R^{n_2}$, represents the vector of control decision variables that their optimal value are contingent upon the realization of uncertain parameters as well as the optimal value of the design variables.

Assume an LP model with the following structure:

$$\min c^T x + d^T y, \; x \in R^{n_1}, y \in R^{n_2}, \tag{6}$$

s.t.

$$Ax = b, \tag{7}$$

$$Bx + Cy = e, \tag{8}$$

$$x, y \ge 0.$$

The objective function (11) consists of design and control decision variables. Equation (12) represents the structural constraints that their coefficients are assumed fixed and free of noise. Equation (13) represents the control constraints that their coefficients and parameters are subject to noise.

To formulate the RO problem, Mulvey et al. (1995) defined a set of probable scenarios $\Omega = \{1, 2, \ldots, S\}$ for LP model (14–15). To this end, for each scenario $s \in \Omega$, the set $\{d_s, B_s, C_s, e_s\}$ of realizations for the coefficients and parameters are associated. Moreover, the probability of scenarios are indicated by p_s,where $\left(\sum_{s=1}^{S} p_s = 1\right)$ The optimal solution of problem (6–8) will be robust with respect to optimality if it stays "close" to optimal for any probable scenario $s \in \Omega$. It is then called solution robust. The solution is also robust with respect to feasibility if it stays "almost" feasible for any probable scenario $s \in \Omega$. It is then called model robust.

It is improbable that any solution to program (6–8) will stay both feasible and optimal for all scenarios indicated by $s \in \Omega$. If the system that is being modeled inherently has substantial redundancies built in, then it might be possible to achieve solutions that stay both feasible and optimal. Otherwise, the RO model proposed by Mulvey et al. (1995) enables us to measure the tradeoff between solution and model robustness. Let us define a set $\{y_1, y_2, \ldots, y_s\}$ of control variables for each scenario $s \in \Omega$. Additionally, let a set $\{z_1, z_2, \ldots, z_s\}$ be the error vectors that measure the infeasibility allowed in the control constraints under scenario $s \in \Omega$. Now, consider the formulation of the RO model as follows:

$$\min \sigma\,(x, y_1, \ldots, y_s) + \omega\rho\,(z_1, \ldots, z_s), \tag{9}$$

s.t.

$$Ax = b, \tag{10}$$

$$B_s x + C_s + Z_s = e_s, \quad \forall s \in \Omega, \tag{11}$$

$$x_s \geq 0, y_s \geq 0, \quad \forall s \in \Omega.$$

With multiple probable scenarios, the objective function $\xi = c^T x + d^T y$ turns into a random variable that takes the value $\xi_s = c^T x + d_s^T y_s$, with probability p_s. Therefore, there is no longer a unique choice for an aggregate objective. Term $\rho(z_1, z_2, \ldots, z_s)$ penalizes violations of the control under some of the scenarios. Different alternatives can be employed for penalty function and it is also problem dependent (Mulvey et al. 1995).For instance, $\rho(z_1, z_2, \ldots, z_s) = \sum_{s\in\Omega} (p_s z_s^T z_s)$ is a quadratic penalty function for equality constrained problems where both positive and negative violations should be penalized. Penalty function $\rho(z_1, z_2, \ldots, z_s) = \sum_{s\in\Omega} p_s \max\{0, z_s\}$ can be applied for inequality control constraints when only positive violations should be penalized (i.e., negative values of show slack in the inequality constraints which are acceptable).By adjustment the goal programming weight ω, the RO model is able to generate a spectrum of solutions that measure tradeoff between solution and model robustness.

Von Neumann and Morgenstern (2007) interpreted the risk as the variance of output. High variance of $\xi_s = c^T x + d_s^T y_s$ shows that there is high fluctuation in outcome. Bar-Shira and Finkelshtain (1999) stated that using the function, which simultaneously raises the mean and reduces variance, is more robust than approaches based on expected value. The following equation demonstrates the mean–variance function for each scenario.

$$\sigma(x, y_1, \ldots, y_s) = \sum_{s\in S} p_s \xi_s + \lambda \sum_{s\in S} p_s (\xi_s - \sum_{s'\in S} p'_s\, \xi_{s'})^2. \tag{12}$$

An efficient frontier can be generated simply by parameterizing the tradeoff between risk and expected outcome (i.e., by changing λ, systematically).This formulation needs that the distribution of the random variable ξ_s be symmetric around its means as well as the third and higher moments of ξ_s are not considered (Mulvey et al. 1995).

Robust DEA based on Mulvey approach

As mentioned, it is almost impossible for DEA applications in many real cases to determine and capture the completely accurate data of the inputs and outputs. In other words, the real data are uncertain and the applications of the exact models could lead to incorrect results.

Since the DEA model (3–20) is an LP, uncertainty in output and input data (i.e., x_{ij} and y_{rj}, respectively) can be formulated by RO model based on approach of Mulvey et al. (1995) [i.e., RO model (21–22)]. A set of scenario of probable input and output data is indicated by $\Omega = \{1, 2, \ldots, S\}$ with incidence probability p_s, for each scenario $s \in \Omega$.Therefore, the robust DEA model based on Mulvey approach is as follows:

$$\max \sum_{s\in\Omega} \sum_{r=1}^{t} p_s u_r y_{ros} - \gamma \sum_{s\in\Omega} p_s \delta_{so} - \lambda \sum_{s\in\Omega} p_s \left(\xi_s - \sum_{s'\in\Omega} p_{s'} \xi_{s'}\right)^2, \tag{13}$$

s.t.

$$\sum_{i=1}^{m} v_i x_{io} + \delta_{so} = 1, \quad \forall s \in \Omega, \tag{14}$$

$$\sum_{r=1}^{t} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + \delta_{sj} = 0, \quad \forall s \in \Omega, \; j = 1, \ldots, n, \tag{15}$$

$v_i, u_r, \delta_{sj} \geq 0, \quad \forall i, r, s.$

The objective function of the robust formulation of DEA has three terms. The first term is the expected efficiency of the DMUs. The second term is the variance of the efficiency weighted by the goal programming parameter γ. δ_{sj} is error variable under scenario s for DMU$_j$ which adjusts how much DMU$_j$ can go out of feasibility space under s scenario. The infeasibility penalty value is measured by term $\left(\sum_{s\in\Omega} p_s \delta_{so}\right)$. Therefore, the third term of objective function (26) penalizes a norm of the infeasibilities, weighted by parameter λ. The coefficients γ and λ are user-defined parameters which identify the importance of variance and infeasibility terms, respectively. The robust formulation (13–15) is an NLP program and the problem can be more readily solved if it is transformed into an LP problem. Therefore, we use transformation variables Q_s^+ and Q_s^- for the quadratic term of the variance in the objective function. Therefore, NLP model (13–15) converts into the following LP model:

$$\max \sum_{s\in\Omega} \sum_{r=1}^{t} p_s u_r y_{ros} - \gamma \sum_{s\in\Omega} p_s \delta_{so} - \lambda \sum_{s\in\Omega} p_s \left(Q_s^+ + Q_s^-\right), \tag{16}$$

s.t.

$$\sum_{i=1}^{m} v_i x_{io} + \delta_{so} = 1, \qquad \forall s \in \Omega, \tag{17}$$

$$\sum_{r=1}^{t} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + \delta_{sj} = 0, \quad \forall s \in \Omega, \; j = 1, \ldots, n, \tag{18}$$

$$\sum_{r=1}^{t} p_s u_r y_{ros} - p_s \left(\sum_{s'\in\Omega} \sum_{r=1}^{t} p_{s'} u_r y_{ros} \right) = Q_s^+ - Q_s^-, \quad \forall s \in \Omega, \tag{19}$$

$v_i, u_r, \delta_{sj}, Q_s^+, Q_s^- \geq 0, \quad \forall i, r, s.$

The variance term in objective function (13) is quadratic and it has been transformed into linear form using variables Q_s^+ and Q_s^-. Constraint (19) computes the variance term of efficiency under each scenario. Term $\left(\sum_{s'\in\Omega} \sum_{r=1}^{t} p_{s'} u_r y_{ros}\right)$ in constraint (19) is the expected value of efficiencies which indicates the amount of DMU's efficiency under probable scenarios. Since one of the variables Q_s^+ and Q_s^- takes the positive value, this constraint measures the expected deviation of efficiency from its expected value such as variance. Therefore, the variance value in model (16–19) is demonstrated via term $\left(\sum_{s\in\Omega} p_s \left(Q_s^+ + Q_s^-\right)\right)$ in objective function.

Case study

To discuss the performance of the proposed robust DEA based on Mulvey approach, we employ the proposed method in the real problem of electricity distribution units. Since, Mulvey robust approach yields an NLP program, we implement the transformation variation to convert it into an LP problem [see model (16–19)].

The actual data of year 2008 for Iranian electricity distribution units (as DMUs) have been considered. The Iranian electricity distribution units, established in 1992, are public and operate under the supervision of TAVANIR Company[1] (Iran Power Generation, Transmission, and Distribution Management Company). TAVANIR has conceived that electricity distribution companies have high incentive to not report real and clear data. It may be beneficial for them to conceal real information and reveal deceptive input and output data. Moreover, real and accurate data about key performance criteria of all companies do not always exist. Therefore, it is important to TAVANIR to analyze efficiency of these companies under a large amount of uncertainty. The analyzers of TAVANIR are able to determine pessimistic, medium, and optimistic scenarios for output and input data of the companies. We propose RDEA to deal with these uncertain situations.

Jamasb and Pollitt (2000) extensively reviewed the electricity international case studies and identified the most often used inputs as operating cost, the number of employees, transformer capacity and network length. Moreover, the most frequently used outputs are recognized as units of energy delivered, the number of customers and the size of service area, as well. Hence, similar to Sadjadi and Omrani (2008), we take account of five parameters as inputs and output data for estimating the operational efficiency of the electricity distribution units. The inputs are the number of labors, transformer capacity, and network length. The outputs are also units of total electricity sales and the number of customers. It is noteworthy that the measurement units for transformer capacity, the network length and total electricity sales are MVA, Kilometer (Km) and MWh, respectively.

[1] TAVANIR is responsible for electricity generation, transmission, and distribution in Iran. This company operates under the supervision of Ministry of Energy.

The case study contains annual data on 38 companies observed in 2008 obtained from Power Generation, Transmission, and Distribution Management Company publications. Note that in the real world, these data are not precise and they are estimated with a specific error level. For instance, the network length and transformers capacity data for electricity distribution units are capital parameters and their actual values are not often available. Moreover, total electricity sales and number of customers are not often reported precisely. We take three scenarios into account for data that are called pessimistic, medium, and optimistic scenarios. The occurrence probabilities of these scenarios are estimated as 0.25, 0.5, and 0.25, respectively. Table 1 demonstrates the inputs and outputs of units under these scenarios.

(s_1 = Pessimistic, s_2 = Medium, s_3 = Optimistic).

The robust DEA model based on Mulvey approach has been applied to evaluate 38 companies for Iranian electricity distribution units under imprecise data in year 2008. In this model, the expected value and variance impact of probable scenarios have been considered. As mentioned earlier, tradeoff between solution robustness and model robustness can be derived by penalty parameter γ. Figure 1 illustrates the effects of the penalty function on the expected value of efficiency of DMUs. We know from the figure that the expected value of efficiency of DMUs is decreasing with the penalty parameter γ. On the other hand, by parameterizing the value λ, the tradeoff between expected value and variance can be constructed. The larger the parameter λ, the more importance of variance of scenarios will be. Figure 2 demonstrates the effects of parameter λ on efficiency of DMUs, as well. From the figure, we know that when λ increases, the expected value of efficiency decreases. From these figures, it is also found that in these cases, the parameters γ and λ significantly alter DMUs' efficiencies, but they rarely change DMUs ranking. Note that in other cases, when there are considerable variations among probable scenarios, changing the parameter λ may remarkably alters the ranking as well.

To get deep insight from robust DEA model (16–19), we compare all terms of objective function systematically. To this end, the goal programming parameters γ and λ are assumed fixed ($\gamma = 3$ & $\lambda = 0.8$). Solving the robust DEA model, all terms of objective function are calculated and reported in columns of Table 2. Next to each column, the ranking is indicated as well.

Now, let us consider the probable scenarios in more detail. The pessimistic efficiency value of each DMU is the worst and the optimistic efficiency of each DMU is the best value of efficiency amount, which are found by efficiency ranking under probable scenarios. For calculating the efficiencies under different scenarios, we should set appropriate values for probabilities. For instance, under pessimistic scenario, we should set $p_1 = 1$, $p_2 = 0$, and $p_3 = 0$. The obtained efficiencies for DMUs in pessimistic, medium, and optimistic scenarios are indicated in Table 2. The efficiencies in these cases also give us the ranking of DMUs in standard DEA formulation under each probable scenario. Comparing the ranking of DMUs, we found that the position of the electricity distribution units has slightly changed under different probable scenarios.

As we saw earlier, different efficiencies and rankings have obtained under various scenarios. Robust DEA model (16–19) gives us an aggregate measure. Expected value and variance terms of the model are indicated in Table 2. In the last column of the table, expected value and variance measures are integrated with coefficient λ. The penalty function $\sum_{s\in\Omega} p_s\delta_{so}$ estimates the infeasibility allowed in different scenarios. The infeasibility penalty measure associated with each DMU is also reported in the table. The highest rank of infeasibility penalty measure demonstrates the electricity distribution unit which has the lowest infeasibility measure.

From the table we found that, the 20th electricity distribution unit is the highest efficient unit. For this DMU, the efficiencies under pessimistic, medium, and optimistic scenarios are 0.81981982, 0.9009009, and 1, respectively. Hence, the expected value of the efficiency of this unit would be 0.90540541. According to the infeasibility penalty measures, we derive that 20th DMU has a low influence on the quality of its efficiency. Total objective value of robust DEA for this DMU is 0.6658089; therefore, it attains the first rank. Moreover, we obtain that 22th DMU has a high influence on the quality of its efficiency where the total objective value of robust DEA for this DMU is 0.1558; hence, it means that 22th DMU attains the last position in all DMUs ranking. Moreover, we derive from the Table 2, that the changing in parameters λ and γ does not trigger a dramatic change in ranking of DMUs.

Conclusion

Uncertainty is an inherent part of the real performance evaluation problems. On the one hand, some precise real data about companies may not always available, on the other hand, some companies may have an incentive to conceal their real output and input data. Although, two approaches of robust optimization for DEA was proposed by Sadjadi and Omrani (2008), discrete data uncertainty based on probable scenarios has not been considered yet. For scenario-based uncertainty about input and output data, we presented a new robust DEA model founded upon the approach of Mulvey et al. (1995). One of the main advantages of this approach is that it enables decision makers to draw a tradeoff between

Table 1 The inputs and outputs of electricity distribution units under three pessimistic, medium and optimistic scenarios

DMU	X_1 (Labor)			X_2 (Network length)			X_3 (Capacity transformers)			Y_1 (Total electricity sales)			Y_2 (Number of customers)		
	S_1	S_2	S_3	S_1	S_2	S_3	S_1	S_2	S_3	S_1	S_2	S_3	S_1	S_2	S_3
1	500	540	567	8,055	8,790	9,014	1,005	1,106	1,204	1,775.75216	1951.376	2,166.02736	480.2687	50,5.546	540.93422
2	801	849	893	16,967	17,680	18,650	106	1,030	1,152	1,319.3999	1,449.89	1,609.3779	455.55065	479.527	513.09389
3	982	1,001	1,057	19,075	20,118	219,870	1,178	1,345	1,426	1,994.06662	2,191.282	2,432.32302	617.96265	650.487	696.02109
4	399	405	463	9,458	10,363	11,358	511	530	590	694.02242	762.662	846.55482	277.6907	292.306	312.76742
5	502	575	589	7,593	8,771	9,412	1,207	1,356	1,407	2,868.78228	3,152.508	3,499.28388	599.2714	630.812	674.96884
6	995	1,028	1,095	24,794	26,399	27,632	2,405	2,866	3,006	4,748.562	5218.2	5,792.202	730.19755	768.629	822.43303
7	274	290	316	6,078	7,802	8014	564	583	600	756.15176	830.936	922.33896	177.5132	186.856	199.93592
8	562	585	604	14,183	15,032	16,824	1,498	1,578	1,681	2,331.88865	2,562.515	2,844.39165	388.80935	409.273	437.92211
9	504	573	597	12,079	13,347	1,489	1,158	1,217	1,326	1,741.70724	1,913.964	2,124.50004	406.36535	427.753	457.69571
10	515	566	586	11,382	12,498	13,587	1,000	1,060	1,156	1,139.86145	1,252.595	1,390.38045	318.63	335.4	358.878
11	625	664	684	11,075	12,183	13,752	2,093	2,368	2,456	4,160.79027	4,572.297	5,075.24967	801.75915	84,3.957	903.03399
12	489	512	537	4,026	5,191	6,420	1,796	1,842	2,004	4,243.85416	4663.576	5,176.56936	684.9519	721.002	771.47214
13	503	530	567	7,904	8,710	9,412	1,824	1,968	2,163	2,836.72207	3,117.277	3,460.17747	584.00775	614.745	657.77715
14	520	555	581	3,591	4,829	5,210	1,503	1,553	1,628	3,053.51683	3,355.513	3,724.61943	716.05395	753.741	806.50287
15	611	650	673	6,591	7,875	8,8622	1,679	1,706	1,876	3,500.89285	3847.135	4,270.31985	724.8671	763.018	816.42926
16	785	806	841	9,748	10,571	11,691	2,078	2,295	2,369	3,767.51921	4,140.131	4,595.54541	793.725	835.5	893.985
17	296	324	364	4,296	5,628	6,589	789	805	864	1,505.92624	1,654.864	1,836.89904	264.6054	278.532	298.02924
18	729	752	776	10,358	11,608	12,530	1,605	1,696	1,789	3,294.06077	3619.847	4,018.03017	741.20615	780.217	834.83219
19	1,295	1,439	1,604	45,952	47,237	50,741	2,507	2,605	2,796	5,833.78614	6,410.754	7,115.93694	1,005.00595	1,057.901	1,131.95407
20	500	524	576	4,962	5,385	6,230	2,178	2,297	2,378	5,675.943	6237.3	6,923.403	247.7733	260.814	279.07098
21	1,012	1,348	1,410	18,063	19,916	21,116	3,901	4,081	4,368	6,370.091	7,000.1	7,770.111	489.88175	515.665	551.76155
22	323	363	389	4,029	5,429	6,620	552	571	598	629.356	691.6	767.676	104.85625	110.375	118.10125
23	304	344	386	9,079	10,567	11,874	783	809	948	1,124.75272	1,235.992	1,371.95112	222.42255	234.129	250.51803
24	316	292	318	8,054	9,031	10,963	984	1,003	1,179	2,117.60549	2,327.039	2,583.01329	292.26465	307.647	329.18229
25	326	360	388	7,295	8,363	9,584	749	768	804	1,234.69073	1,356.803	1,506.05133	187.0968	196.944	210.73008
26	487	507	556	12,078	13,230	14,963	1,096	1,210	1,356	1,440.68106	1,583.166	1,757.31426	399.9272	420.976	450.44432
27	368	398	439	11,007	12,082	13,745	742	767	789	998.13896	1,096.856	1,217.51016	308.2978	324.524	347.24068
28	218	241	270	4,178	5,125	6,078	403	455	468	512.3664	563.04	624.9744	108.44345	114.151	122.14157
29	678	717	738	12,071	13,480	14,963	1,962	2,008	2,256	2,567.90989	2,821.879	3,132.28569	508.12555	534.869	572.30983
30	736	759	785	24,069	25,735	27,620	1,859	1,970	2,168	3,032.41939	3,332.329	3,698.88519	489.56635	515.333	551.40631
31	307	354	374	6,287	7,522	8,054	1,059	1,163	1,278	1,662.11318	1,826.498	2,027.41278	190.77615	200.817	214.87419
32	946	997	1,150	29,078	31,554	35,812	1,996	2,145	2,369	4,490.73534	4,934.874	5,477.71014	568.7156	598.648	640.55336
33	871	900	1,046	19,963	21,665	24,589	1,586	1,634	1,786	2,288.50076	2,514.836	2,791.46796	738.46255	777.329	831.74203
34	739	773	796	17,057	18,897	20,746	1,762	1,877	1,956	2,232.25457	2453.027	2,722.85997	667.60775	702.745	751.93715
35	215	269	288	5,210	6,224	7,786	601	645	675	790.22034	868.374	963.89514	256.2967	269.786	288.67102
36	378	407	462	9,350	1,0498	11,998	965	1,005	1,157	1,207.22966	1,326.626	1,472.55486	341.5611	359.538	384.70566
37	894	924	987	15,074	16,510	17,930	1,779	1,865	1,973	2,812.72992	3090.912	3,430.91232	289.6512	304.896	326.23872
38	542	596	637	11,096	12,595	13,875	997	1,022	1,157	2,124.73989	2334.879	2,591.71569	337.5768	355.344	380.21808

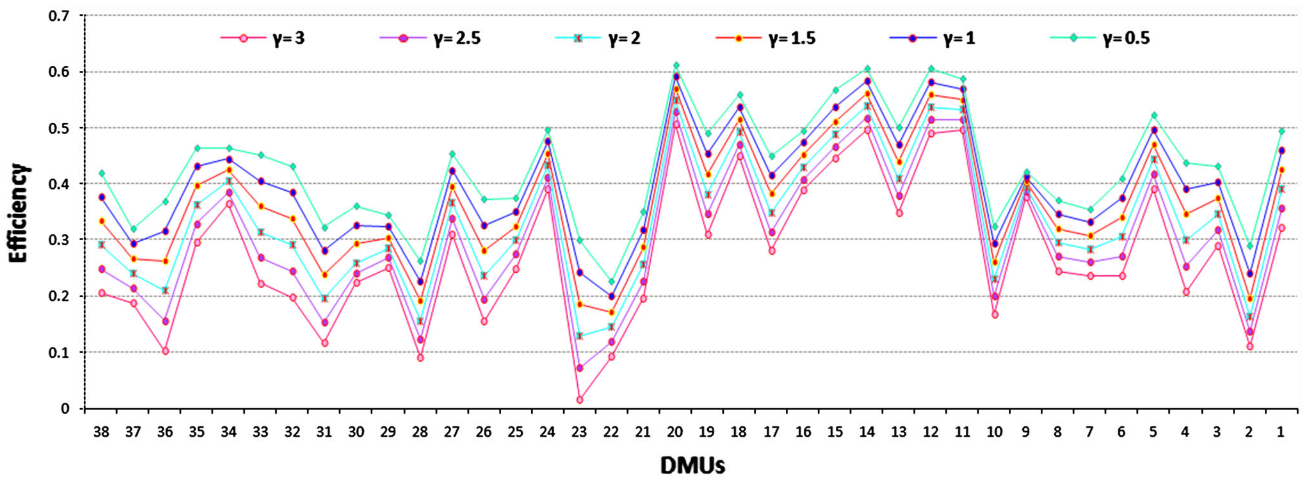

Fig. 1 Comparison of objective function under $\lambda = 0.8$ and different values of γ

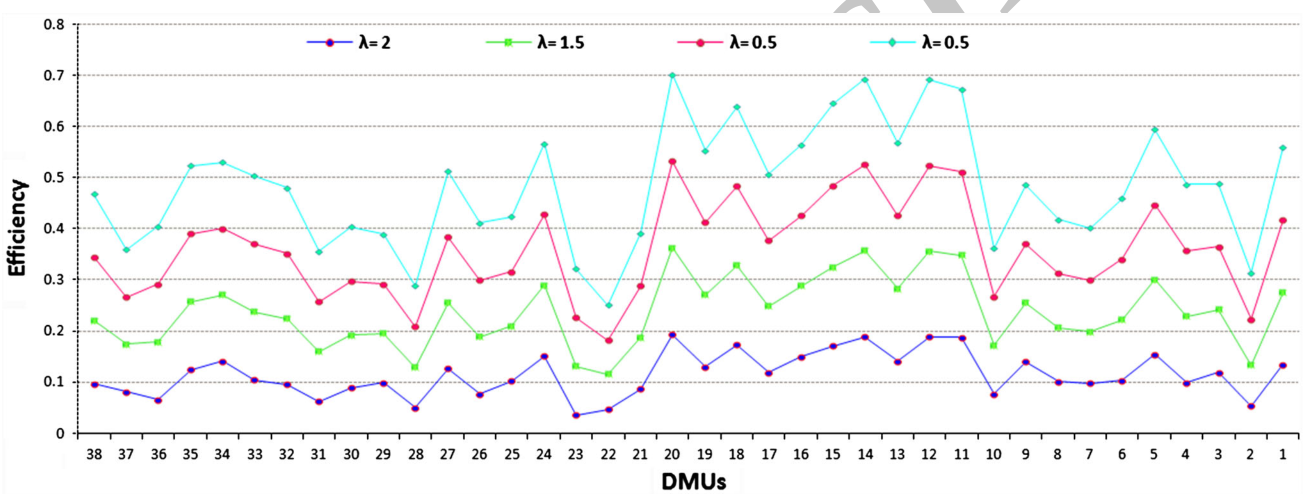

Fig. 2 Comparison of objective function under $\gamma = 0.8$ and different values of λ

expected value and variance of DMUs' efficiency under probable scenarios, and a tradeoff between solution and model robustness. We implemented the results of the proposed model using data gathered from an Iranian energy organization when optimistic, medium, and pessimistic scenarios exist about their reported data. Our preliminary results indicate that our robust DEA approach can provide analysts with more decision criteria under uncertain condition. Moreover, sensitivity analyses on parameters λ and γ can derive a spectrum of solutions that may be useful for managerial tradeoffs.

Although our model is restricted to input-oriented model and constant return to scale, one can easily generalize it to output-oriented model and variable return to scale. Our model considers statics situation, however, it can be readily developed to dynamic situation to adapt the model in multi-period real problems. There are also other directions and suggestions for future research. First, we assumed that all outputs are desirable; however, the real problem may be the undesirable output data, decreasing the amount of which is favorable. Extending our robust DEA model for both desirable and undesirable data is interesting. Second, the robust DEA model can be developed into the two-stage or network DEA models where output of a DMU becomes input data of other DMU(s). Eventually, the parameters of the proposed model may be changed during the planning horizon. In this situation, we can expand the suggested model into the Malmquist model in dynamic condition. Since the return to scale models is linear, if the discrete uncertainties have been observed in data, one can utilize our approach to analyze the efficiency of DMUs. It means that all four models of CCR, BCC, CCR-BCC, and BCC-

Table 2 All terms of objective function of the Robust DEA model under $\lambda = 0.8$ and $\gamma = 3$

DMU	Pessimistic efficiency		Medium efficiency		Optimistic efficiency		Expected value of efficiency		Variance measure		Infeasibility penalty measure		Objective function		Expected value-**λ** *(variance)	
	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank
1	0.7137	8	0.7512	9	0.8038	11	0.7550	9	0.2827	9	0.0686	11	0.4556	13	0.5289	9
2	0.3551	37	0.3738	37	0.3999	37	0.3756	37	0.1406	37	0.0504	23	0.1778	35	0.2631	37
3	0.6205	20	0.6532	21	0.6989	21	0.6564	21	0.2458	21	0.0564	19	0.4058	17	0.4598	21
4	0.6525	16	0.6868	17	0.7349	18	0.6902	17	0.2584	17	0.0917	4	0.3294	25	0.4835	17
5	0.7133	9	0.7794	7	0.8610	7	0.7833	7	0.2932	7	0.0526	22	0.5282	7	0.5487	7
6	0.5773	22	0.6296	22	0.6945	22	0.6328	22	0.2369	22	0.0687	10	0.3482	21	0.4432	22
7	0.5029	28	0.5385	28	0.5848	29	0.5412	28	0.2026	28	0.0474	26	0.3317	24	0.3791	28
8	0.5230	26	0.5621	26	0.6123	26	0.5649	26	0.2115	26	0.0500	24	0.3446	22	0.3957	26
9	0.5713	23	0.6131	23	0.6671	23	0.6161	23	0.2307	23	0.0181	38	0.4854	11	0.4316	23
10	0.4797	30	0.5049	31	0.5403	32	0.5074	31	0.1900	31	0.0623	15	0.2577	31	0.3555	31
11	0.8159	4	0.8589	4	0.9190	4	0.8631	4	0.3231	4	0.0362	35	0.6475	4	0.6046	4
12	0.8212	2	0.8924	2	0.9813	2	0.8968	2	0.3357	2	0.0456	27	0.6486	3	0.6282	2
13	0.7170	7	0.7547	8	0.8075	10	0.7585	8	0.2840	8	0.0608	16	0.4818	12	0.5313	8
14	0.8301	1	0.8923	3	0.9724	3	0.8968	3	0.3357	3	0.0441	29	0.6532	2	0.6282	3
15	0.7763	5	0.8172	6	0.8744	6	0.8213	6	0.3075	6	0.0434	30	0.5891	6	0.5753	6
16	0.6795	11	0.7153	12	0.7654	12	0.7189	12	0.2691	12	0.0384	34	0.5159	9	0.5036	12
17	0.6305	18	0.6870	16	0.7573	14	0.6905	16	0.2585	16	0.0677	13	0.4017	18	0.4837	16
18	0.7711	6	0.8260	5	0.8975	5	0.8302	5	0.3108	5	0.0441	28	0.5947	5	0.5815	5
19	0.6876	10	0.7493	10	0.8260	8	0.7530	10	0.2819	10	0.0727	8	0.4414	14	0.5275	10
20	0.8198	3	0.9009	1	1.0000	1	0.9054	1	0.3390	1	0.0424	32	0.6658	1	0.6342	1
21	0.4910	29	0.5373	29	0.5943	27	0.5400	29	0.2022	29	0.0603	17	0.2920	29	0.3783	29
22	0.3303	38	0.3601	38	0.3970	38	0.3619	38	0.1355	38	0.0537	20	0.1558	37	0.2535	38
23	0.4712	32	0.5054	30	0.5498	30	0.5079	30	0.1902	30	0.1136	1	0.1040	38	0.3558	30
24	0.6769	12	0.7359	11	0.8095	9	0.7395	11	0.2769	11	0.0425	31	0.5204	8	0.5180	11
25	0.5206	27	0.5678	25	0.6264	25	0.5707	25	0.2136	25	0.0499	25	0.3501	19	0.3997	25
26	0.5253	25	0.5530	27	0.5917	28	0.5558	27	0.2081	27	0.0776	7	0.2541	32	0.3893	27
27	0.6504	17	0.6846	18	0.7326	19	0.6881	18	0.2576	18	0.0573	18	0.4306	15	0.4820	18
28	0.3764	36	0.4096	36	0.4510	36	0.4116	36	0.1541	36	0.0658	14	0.1631	36	0.2883	36
29	0.4793	31	0.5045	32	0.5398	33	0.5070	32	0.1898	32	0.0346	36	0.3404	23	0.3552	32
30	0.4341	34	0.4570	35	0.4889	35	0.4592	35	0.1719	35	0.0322	37	0.3057	28	0.3217	35
31	0.4323	35	0.4706	34	0.5183	34	0.4729	34	0.1770	34	0.0715	9	0.2030	34	0.3313	34
32	0.6228	19	0.6800	19	0.7507	16	0.6834	19	0.2558	19	0.0936	3	0.3178	27	0.4787	19
33	0.6710	13	0.7063	14	0.7557	15	0.7098	14	0.2657	14	0.0912	5	0.3482	20	0.4972	14
34	0.6541	15	0.6885	15	0.7367	17	0.6919	15	0.2590	15	0.0398	33	0.4867	10	0.4847	15
35	0.6688	14	0.7085	13	0.7625	13	0.7121	13	0.2666	13	0.0678	12	0.4203	16	0.4988	13
36	0.5701	24	0.6001	24	0.6421	24	0.6031	24	0.2258	24	0.1063	2	0.2094	33	0.4225	24
37	0.4510	33	0.4929	33	0.5447	31	0.4954	33	0.1855	33	0.0530	21	0.2750	30	0.3470	33
38	0.6037	21	0.6582	20	0.7259	20	0.6615	20	0.2477	20	0.0856	6	0.3227	26	0.4634	20

CCR can be appropriately altered by the robust optimization approach.

References

Bar-Shira Z, Finkelshtain I (1999) Two-moments decision models and utility-representable preferences. J Econ Behave Organ 38:237–244. doi:10.1016/S0167-2681(99)00008-6

Bashiri M, Moslemi A (2013) The analysis of residuals variation and outliers to obtain robust response surface. J Ind Eng Int 9(1):1–10

Bashiri M, Farshbaf-Geranmayeh A, Mogouie H (2013) A neuro-data envelopment analysis approach for optimization of uncorrelated multiple response problems with smaller the better type controllable factors. J Ind Eng Int 9(1):30

Ben-Tal A, Nemirovski A (2000) Robust solutions of linear programming problems contaminated with uncertain data. Math Program 88(3):411–424

Bertsimas D, Sim M (2003) Robust discrete optimization and network flows. Math Progr 98:49–71. doi:10.1007/s10107-003-0396-4

Bertsimas D, Sim M (2004) The price of robustness. Oper Res 52:35–53. doi:10.1287/opre.1030.0065

Bertsimas D, Thiele A (2006) A robust optimization approach to inventory theory. Oper Res 54:150–168. doi:10.1287/opre.1050.0238

Bertsimas D, Pachamanova D, Sim M (2004) Robust linear optimization under general norms. Oper Res Lett 32:510–516. doi:10.1016/j.orl.2003.12.007

Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2:429–444. doi:10.1016/0377-2217(78)90138-8

Charnes A, Cooper WW, Seiford L, Stutz J (1983) Invariant multiplicative efficiency and piecewise Cobb-Douglas envelopments. Oper Res Lett 2:101–103. doi:10.1016/0167-6377(83)90014-7

Foroughi AA, Aouni B (2012) Ranking units in DEA based on efficiency intervals and decision-maker's preferences. Int Trans Oper Res 19:567–579. doi:10.1111/j.1475-3995.2011.00834.x

Goto M, Tsutsui M (1998) Comparison of productive and cost efficiencies among Japanese and US electric utilities. Omega-Int J Manage S26:177–194. doi:877a01e2e3e3d0aa067a8d136704a38b

Hatami-Marbini A, Tavana M, Saati S, Agrell PJ (2012) Positive and normative use of fuzzy DEA-BCC models: a critical view on NATO enlargement. Int T Oper Res 20:411–433. doi:10.1111/j.1475-3995.2012.00871.x

Hosseini MB, Tarokh MJ (2011) Interval type-2 fuzzy set extension of DEMATEL method. In: Computational Intelligence and Information Technology, vol 250. Springer, Berlin, Heidelberg, pp 157–165

Izadi A, Kimiagari AM (2014) Distribution network design under demand uncertainty using genetic algorithm and Monte Carlo simulation approach: a case study in pharmaceutical industry. J Ind Eng Int 10(1):1

Jamasb T, Pollitt M (2000) Benchmarking and regulation: international electricity experience. Utilities Policy 9:107–130. doi:272cd35c9c5e488b5e93b23f124f38b8

Khalaj M, Khalaj F, Khalaj A (2013) A novel risk-based analysis for the production system under epistemic uncertainty. J Ind Eng Int 9(1):35

Monfared MAS, Safi M (2013) Network DEA: an application to analysis of academic performance. J Ind Eng Int 9(1):1–10

Morita H (2003) Analysis of economies of scope by data envelopment analysis: comparison of efficient frontiers. Int Trans Oper Res 10:393–402. doi:10.1111/1475-3995.00415

Mulvey JM, Vanderbei RJ, Zenios SA (1995) Robust optimization of large scale system. Operres 43:264–281. doi:10.1287/opre.43.2.264

Roghanian E, Foroughi A (2010) An empirical study of Iranian regional airports using robust data envelopment analysis. Int J Ind Engine Comput 1:65–72. doi:10.5267/j.ijiec.2010.01.006

Sadjadi SJ, Omrani H (2008) Data envelopment analysis with uncertain data: an application for Iranian electricity distribution companies. Energy Policy 36:4247–4254. doi:10.1016/j.enpol.2008.08.004

Saranga H, Phani BV (2009) Determinants of operational efficiencies in the Indian pharmaceutical industry. Int Trans Oper Res 16:109–130. doi:10.1111/j.1475-3995.2009.00668.x

Satapathy S, Mishra P (2013) A customer oriented systematic framework to extract business strategy in Indian electricity services. J Ind Eng Int 9(1):33

Shad Z, Roghanian E, Mojibian F (2014) Integration of QFD, AHP, and LPP methods in supplier development problems under uncertainty. J Ind Eng Int 10(1):1–9

Shokouhi AH, Hatami-Marbini A, Tavana M, Saati S (2010) A robust optimization approach for imprecise data envelopment analysis. Comput Ind Eng 59:387–397. doi:10.1016/j.cie.2010.05.011

Soyster AL (1973) Technical Note—convex programming with set-inclusive constraints and applications to inexact linear programming. Oper res 21:1154–1157. http://www.jstor.org/stable/168933

Von Neumann J, Morgenstern O (2007) Theory of games and economic behavior (commemorative edition). Princeton university press, New Jersey