CEJOR
https://doi.org/10.1007/s10100-018-0549-4

ORIGINAL PAPER

Selecting most efficient information system projects in presence of user subjective opinions: a DEA approach

Mehdi Toloo[1] · **Soroosh Nalchigar**[2] · **Babak Sohrabi**[3]

© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract Information System (IS) project selection is a critical decision making task that can significantly impact operational excellence and competitive advantage of modern enterprises and also can involve them in a long-term commitment. This decision making is complicated due to availability of numerous IS projects, their increasing complexities, importance of timely decisions in a dynamic environment, as well as existence of multiple qualitative and quantitative criteria. This paper proposes a Data Envelopment Analysis approach to find most efficient IS projects while considering subjective opinions and intuitive senses of decision makers. The proposed approach is validated by a real world case study involving 41 IS projects at a large financial institution as well as 18 artificial projects which are defined by the decision makers.

Keywords Information systems · Project management · Data envelopment analysis · Performance evaluation · Selection-based problems

✉ Mehdi Toloo
mehdi.toloo@vsb.cz
http://homel.vsb.cz/~tol0013/

Soroosh Nalchigar
soroosh@cs.toronto.edu

Babak Sohrabi
bsohrabi@ut.ac.ir

1 Department of Systems Engineering, Technical University of Ostrava, Sokolská tř. 33, 702 00 Ostrava, Czech Republic

2 Department of Computer Science, University of Toronto, Toronto, Canada

3 Department of Information Technology Management, University of Tehran, Tehran, Iran

Published online: 14 May 2018

1 Introduction

Digital economy has converted Information Technology (IT) management to one of the critical organizational positions. Today, IT managers have many responsibilities (e.g., data centers, staff management, telecommunication, servers, workstations, websites, user support, regulatory compliance, disaster recovery) and interact with various departments within the enterprise. In many organizations, these managers can have a direct impact on strategic direction of the company (Holtsnider and Jaffe 2012).

A critical task of IT managers is the decision making by which the most proper IS projects are selected from a set of competing proposals (Asosheh et al. 2010; Badri et al. 2001). This decision making is difficult and complicated due to availability of numerous IS projects, their increasing complexities, importance of timely decisions in a dynamic environment (Deng and Wibowo 2008), as well as existence of various qualitative and quantitative criteria (Chen and Cheng 2009; Yang et al. 2013). Selecting the best IS projects is a critical strategic resource allocation decision that can involve the enterprise in a long-term commitment (Badri et al. 2001). In these contexts, establishing a systematic IS project selection approach is of great importance for today's organizations (Yang et al. 2013).

The IS project selection is a Multi-Criteria Decision Making (MCDM) problem (Karsak and Özogul 2009; Lee and Kim 2001; Yeh et al. 2010) that has received lots of attention from both academic researchers as well as industrial practitioners. While various methods have been proposed in the literature, approaches that consider decision makers' subjective opinions have gained less attention. Moreover, those existing methods that accommodate subjective opinions often require decision makers to provide criteria weights (either regarding exact numeric values or fuzzy linguistic terms), which could be difficult and impossible is some cases. In other words, while these methods can be useful in supporting IS project selection, they suffer from requiring a significant amount of input from the decision makers which often proves to be tough to obtain. In many decision situations, it is easier for decision makers to provide samples of *good* and *bad* alternatives rather than defining weights for decision criteria and calculating the utility of alternatives (Sowlati et al. 2005). This paper extends a novel Data Envelopment Analysis (DEA) approach for evaluation and selection of IS projects from a set of competing proposals. The approach takes into account the subjective and intuitive opinions of decision makers regarding artificial projects that are representative of good or bad alternatives.

Charnes et al. (1978) introduced the first DEA model (CCR model) under constant returns to scale (CRS) assumption and Banker et al. (1984) extended a new DEA model (BCC model) with the aim of considering variable returns to scale (VRS) assumption. DEA is a non-parametric Linear Programming (LP) based technique for measuring the relative efficiency of a set of homogeneous units, usually referred to as Decision Making Units (DMUs). The basic idea of DEA is that the relative efficiency of a DMU is determined by its ability to convert inputs into desired outputs. Due to its successful applications and case studies, DEA has received an enormous amount of attention by researchers. Efficiency analysis of organizational investments in IT (Shafer and Byrd 2000), evaluation of data mining algorithms (Nakhaeizadeh and Schnabl 1997), examining bank efficiency (Paradi and Zhu 2013), modeling environmental performance and energy efficiency (Arabi et al. 2016; Zhou et al. 2008), assessment of company's financial statements (Edirisinghe and Zhang 2007), performance evaluation of Research and Development (R&D) active firms (Khoshnevis and Teirlinck 2018), examining Spanish and Portuguese construction companies (Kapelko 2018), and ranking of countries in Summer Olympic Games 2016 (Jablonsky 2018) are among applications of DEA in various areas.

One of the purposes of DEA in practice is to provide the prioritization among DMUs. However, DEA models partition all the DMUs into two sets: efficient and inefficient, where an efficient and inefficient DMU respectively have a score of 1 and less than 1. Hence, these models fail to provide more information about the efficient DMUs. To tackle this issue, some ranking methods are developed which enable us to discriminate between efficient DMUs. The supper-efficiency and cross-efficiency are two well-known ranking methods which are originated by Andersen and Petersen (1993) and Doyle and Green (1994), respectively. The super-efficiency model compares the unit under evaluation with a linear combination of all other units. As a matter of fact, the super-efficiency score is obtained by eliminating the data on the unit under consideration from the solution set, and hence the efficient DMU may get a score greater than one. Nonetheless, the supper-efficiency models may suffer from infeasibility issue (Adler et al. 2002). Cross-efficiency is based on the concept of peer review and on the efficiencies determined for each DMU by using optimal weighting from other DMUs (Sexton et al. 1986). It is harder to have ties in cross efficiency than in traditional DEA; however, cross-efficiency uses a fixed weighting scheme for all DMUs for single input and multiple outputs, which eliminates the flexibility of each DMU to have its own weighting scheme (Sowlati et al. 2005).

In some real-world problems, known as *selection-based* problems, selecting a single efficient unit is concerned rather than ranking all DMUs: For instance, in DEA applications such as robot selection (Baker and Talluri 1997), flexible manufacturing system selection (Shang and Sueyoshi 1995), enterprise resource planning system selection (Karsak and Özogul 2009; Lall and Teyarachakul 2006), media selection (Farzipoor Saen 2011), recommender system selection (Sohrabi et al. 2015), athlete selection (Masoumzadeh et al. 2016; Ramón et al. 2012; Toloo and Tavana 2017), and facility layout design problem (Ertay et al. 2006; Toloo 2015). This paper presents a new DEA approach for selecting most efficient IS projects in the presence of user subjective opinions. The paper illustrates the application of the approach to a real-world case in a financial institution, and also compares its results with previous methods. The approach can be applied in situations where decision maker(s) need to rank and select from alternatives that are in competition for limited resources. While this paper focuses on IS project selection as the main application case, the approach could be used in future for ranking alternatives in other domains, where the goal is to rank alternatives given a set of criteria and managerial judgments.

The rest of this paper is structured as follows. Section 2 reviews previously proposed methods for ranking IS projects. Section 3 reviews baseline DEA models that are related to this study. Section 4 presents our proposed DEA approach. In order to validate the proposed approach and to illustrate its characteristics and advantages, the penultimate section utilizes a real data set involving 41 real IS projects along with 18 artificial projects which have been defined by the decision makers. It also compares the new approach with some other approaches. This paper ends with Sect. 6 which provides some concluding remark and directions for future research.

2 Related works

Numerous methods have been proposed in the literature for evaluating and ranking IT and IS projects. For example, Schniederjans and Santhanam (1993) proposed a zero–one goal programming model for selecting IS projects and discussed the importance of using multi-objective, constrained resource modeling in IS project selection decision making. Han et al. (1998) used quality function development technique to propose a method for determining IS development projects priority. Their approach takes alignment between business strategy and IS into consideration as a part of the evaluation process and then the approach was applied to a real-world case in Korea. Santhanam and Kyparisis (1995) synthesized project selection models of various disciplines (e.g., R&D, capital budgeting) and formulated a nonlinear 0–1 Goal Programming (GP) model for IS project selection. Schniederjans and Wilson (1991) combined Analytic Hierarchy Process (AHP) within GP modeling to propose an IS project selection methodology. The authors demonstrated the applicability of their approach on a numerical example and showed that hybrid approaches have advantages from using these techniques separately.

Lee and Kim (2001) discussed that previous methods disregard the interdependencies among decision criteria and alternative projects and also lack a group decision making approach. To overcome these drawbacks, they synthesized Delphi technique, Analytic Network Process (ANP), and zero–one GP as an integrated approach for IS project selection. Shafer and Byrd (2000) proposed a framework based on DEA for evaluating the efficiency of organizational IT investments and applied their method to data from 209 large organizations. Badri et al. (2001) discussed that multiple factors affect the IS project selection decision and there is a lack of a single model that includes all necessary factors. To tackle this drawback, the author proposed a zero–one GP project selection model that includes a comprehensive set of factors derived from other disciplines. Real-world data of IS projects were used to validate their approach. Sowlati et al. (2005) proposed a DEA method for ranking IS projects. Their approach needs decision makers to define a set of artificial projects to which each real project is compared and receives a ranking score. They tested their approach on real data of IS projects at a large financial institution. Later in this paper, we will refer to this study and show the application of our approach on the data set presented in their paper, along with a comparison analysis with other similar methods.

Sarkis and Sundarraj (2006) proposed a two-stage methodology for evaluation of enterprise IT technologies. The first stage uses ANP to produce utility weights for each alternative and the second stage uses integer programming to select alternative(s) subject to managerial and cost constraints. Kengpol and Tuominen (2006) proposed an integrated approach of ANP, Delphi, and Maximise Agreement Heuristic (MAH) method for reaching a group consensus and selecting IT proposals. The approach was applied in a real case of five logistics firms in Thailand also some of its limitations were discussed. Deng and Wibowo (2008) developed a decision support system for assisting selection of the multi-criteria analysis method in solving the IS project selection problem. The proposed system includes a knowledge base of IF–THEN rules that represent the effect of the characteristics of decision problem and requirements on the decision analysis technique. They demonstrated the applicability of the proposed approach in a real case of selecting a supply chain management IS project at a steel mill company in Taiwan. Karsak and Özogul (2009) proposed an approach which integrates Quality Function Deployment (QFD), fuzzy linear regression and zero–one goal programming to deal with Enterprise Resource Planning (ERP) system selection. Gao et al. (2008) proposed a fuzzy approach based on the Technique for Order Performance by Similarity to Ideal Solution (TOPSIS) for IS project selection. The authors illustrated the approach on a case study of IS selection in a Chinese university.

Yeh et al. (2010) presented a fuzzy multi-criteria decision making approach for selecting IS projects. The proposed approach handles subjectiveness and imprecision of the human decision making process and uses triangular fuzzy numbers to characterize linguistic terms. Nalchigar and Nasserzadeh (2009) proposed a DEA model for finding efficient IS project in the presence of imprecise data and illustrated its application on real data of eight competing IS project proposals in Iran Ministry of Commerce. Later, Asosheh et al. (2010) extended their DEA approach and combined it with Balanced Score Card (BSC) to propose a new approach for IT project selection. The approach used BSC as a framework for defining the set of evaluation criteria and DEA for ranking the alternatives. Hou (2011) proposed a grey multi-criteria decision model for IT/IS project selection. The grey theory was used to deal with the uncertainty and fuzziness of IT/IS project selection contexts. Bai and Zhan (2011) proposed a fuzzy ANP method for IT project selection and illustrated its application in an oils and food importer and exporter company in China. Yang et al. (2013) proposed a hybrid decision model for IS project selection based on three categories of criteria (critical, quantitative, and qualitative) that were derived from literature review and interviews. The authors showed the applicability of their approach for cargo IS selection of an airline company. The approach presented in this paper is different from previous works in the sense that it combines decision makers' subjective opinions regarding a set of artificial alternatives and finds the most efficient IS project using a new integrated DEA approach.

3 DEA models

DEA is a non-parametric, LP-based technique for measuring and assessing the relative efficiency of a set of similar entities. The original DEA models (CCR and BCC) define efficiency of a DMU as the maximum ratio of weighted outputs to weighted inputs, subject to the constraint that the same proportion for all DMUs must be less than or equal to one. The outcome of these models is a categorization of all DMUs as either *efficient* or *inefficient*, and hence these models fail to discriminate the efficient DMUs. However, in many applications, the decision makers need to find a single most efficient DMU among a given set of alternatives. To solve this problem, some DEA models have been proposed in the literature. Ertay et al. (2006) extended minimax DEA model to identify a single most efficient DMU and used that to evaluate layout design of manufacturing systems. Amin and Toloo (2007) improved their work and proposed Model (1) for finding the most efficient DMU, given a set of units.

$$
\begin{aligned}
& \min d_{max} \\
& s.t. \\
& d_{max} - d_j \geq 0 && j = 1, \ldots, n \\
& \sum_{i=1}^{m} v_i x_{ij} \leq 1 && j = 1, \ldots, n \\
& \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + d_j - \beta_j = 0 && j = 1, \ldots, n \\
& \sum_{j=1}^{n} d_j = n - 1 \\
& 0 \leq \beta_j \leq 1 && j = 1, \ldots, n \\
& d_j \in \{0, 1\} && j = 1, \ldots, n \\
& v_i, u_r \geq \varepsilon && i = 1, \ldots, m; r = 1, \ldots, s
\end{aligned}
\tag{1}
$$

This model is a Mixed Binary Linear Programming (MBLP) model in which it is assumed there are n DMUs with multiple inputs and multiple outputs: j: index for DMUs ($j = 1, \ldots, n$), i: index for inputs ($i = 1, \ldots, m$), r: index for outputs ($r = 1, \ldots, s$), x_{ij}: ith input of DMU$_j$, y_{rj}: rth input of DMU$_j$, v_i: weight for ith input, u_r: weight for rth output, d_j: deviation of the DMU$_j$ from efficiency frontier. d_{max}: maximum deviation from efficiency ($d_{max} = \max \{d_j : j = 1, \ldots, n\}$), ε: the non-Archimedean infinitesimal for forestalling weights to be equal to zero.

The objective function of Model (1) is to minimize the maximum deviation from efficiency. DMU$_j$ is most efficient unit if and only if $d_j^* = 0$. The constraint $\sum_{j=1}^{n} d_j = n - 1$ forces among all the DMUs for only a single (known as most efficient) unit. The model uses optimal Common Set of Weights (CSW) for all DMUs and hence it needs to be solved only once in order to find the most efficient unit. In general, common weight models have several advantages over the traditional DEA models; (i) the optimal set of weights is obtained by solving only a single integrated problem, (ii) there is no need to solve corresponding individual LP problem for evaluating all efficiencies, (iii) these models render more discriminating power among efficient DMUs. For more details about the common weight approach and its benefits, we refer the readers to Cook et al. (1990), Roll et al. (1991).

Model (1) inherits all aforementioned advantages of common weight approaches. It assumes the constant returns to scale technology in order to identify the most CCR-efficient unit. Hence, it is not applicable to cases in which DMUs operate in variable returns to scale. Later, Toloo and Nalchigar (2009) extended this model and proposed a new DEA model for identifying the most BCC-efficient unit. Also, Amin (2009) argued that Model (1) may result in more than one efficient DMU and suggested a non-linear model which has been linearized by Toloo et al. (2017).

Toloo and Nalchigar (2011) continued this line of research and proposed a new DEA model for finding the most efficient unit while the data of inputs and outputs of alternatives are imprecise (e.g., cardinal, ordinal, or interval). They illustrated the application of their model in a supplier selection setting, where the goal was to find the best supplier. Toloo (2012b) found some problems in the Toloo and Nalchigar's

(2009) model and proposed a new integrated (MBLP-DEA) model for finding the most BCC-efficient DMU. His approach includes two steps. The first step recognizes a set of candidate DMUs for being the most BCC-efficient unit. The second step determines the single most BCC-efficient unit among the candidates. The author proposed following LP model to be used in the first step:

$$\begin{aligned} &\min d_{max} \\ &\text{s.t.} \\ &d_{max} - d_j \geq 0 && j = 1, \ldots, n \\ &\sum_{i=1}^{m} v_i x_{ij} \leq 1 && j = 1, \ldots, n \\ &\sum_{r=1}^{s} u_r y_{rj} + u_0 - \sum_{i=1}^{m} v_i x_{ij} + d_j = 0 && j = 1, \ldots, n \\ &d_j \geq 0 && j = 1, \ldots, n \\ &v_i, u_r \geq \varepsilon && i = 1, \ldots, m; r = 1, \ldots, s \end{aligned} \tag{2}$$

The free variable u_0 is added to the model to have a variable returns to scale envelop. Here DMU$_j$ is efficient if and only if $d_j^* = 0$ (or equivalently, $\frac{\sum_{r=1}^{s} u_r^* y_{rj} + u_0^*}{\sum_{i=1}^{m} v_i^* x_{ij}} = 1$). It is plain to verify that there can be more than one efficient DMU and hence DMU$_j$ is a *candidate* for being most BCC-efficient DMU if and only if $d_j^* = 0$. In order to enforce finding a single efficient DMU, Toloo (2012b) imposed a new set of auxiliary binary variables along with some additional constraints into Model (1) and formulated following MBLP-DEA integrated model:

$$\begin{aligned} &\min d_{max} \\ &\text{s.t.} \\ &d_{max} - d_j \geq 0 && j = 1, \ldots, n \\ &\sum_{i=1}^{m} v_i x_{ij} \leq 1 && j = 1, \ldots, n \\ &\sum_{r=1}^{s} u_r y_{rj} + u_0 - \sum_{i=1}^{m} v_i x_{ij} + d_j = 0 && j = 1, \ldots, n \\ &\sum_{j=1}^{n} \theta_j = n - 1 \\ &d_j \leq M\theta_j && j = 1, \ldots, n \\ &\theta_j \leq N d_j && j = 1, \ldots, n \\ &\theta_j \in \{0, 1\} && j = 1, \ldots, n \\ &u_r \geq \varepsilon && r = 1, \ldots, s \\ &v_i \geq \varepsilon && i = 1, \ldots, m \end{aligned} \tag{3}$$

where M and N are large numbers and θ_j is a binary variable. In this model, if $\theta_j = 0$, then clearly the constraint $\theta_j \leq N d_j$ is redundant and the constraint $d_j \leq M\theta_j$ forces that d_j is equal to zero. Otherwise, if $\theta_j = 1$, then $d_j \leq M\theta_j$ is a redundant constraint and $\theta_j \leq N d_j$ insures d_j to be positive. These imply that in this model:

$$d_j \begin{cases} = 0 & \theta_j = 0 \\ > 0 & \theta_j = 1 \end{cases}$$

Toloo (2013) continued this line of research by formulating a new DEA model to determine the best efficient DMU between the several efficient ones without explicit inputs. More recently, Toloo (2014a) proposed an epsilon-free DEA approach for finding most efficient units. His approach excludes the non-Archimedean infinitesimal epsilon and is computationally more efficient than previous ones; however, it is applicable only to constant returns to scale situation. Toloo and Ertay (2014) and Toloo (2016) dealt with vendor selection problem under certain and uncertain input prices assumptions. The authors, in order to illustrate the potential application of their approaches, utilized a case study of an automotive company located in Turkey. Toloo and Kresta (2014) developed a method to select the best alternative for asset financing and applied their approach to a real data set involving 139 different alternatives for long-term asset financing provided by Czech banks and leasing companies. Interested readers are referred to those references for formulation of these models.

Although these models can support decision making in organizations, they have a conceptual view of the decision problem context and do not accommodate user's subjective opinions and judgments. In other words, these models are not well aligned with real-world organizational decision making contexts and do not represent any subjective information from decision makers. This paper fills this gap by proposing a new DEA approach that finds the most efficient DMU while considering decision makers subjective opinions.

Sowlati et al. (2005) created an LP model within DEA framework for ranking DMUs. Their model accommodates decision makers' intuitive sense and produces a priority score for each DMU that allows them to be ranked. The intuitive, subjective opinions of decision makers are imported into the approach in terms of a set of artificial DMUs that are representative of good or bad alternatives. Their approach requires decision makers to provide a set of sample DMUs, called *artificial* DMUs, and to define the value of each criterion and a priority score for each of them. Then, the model compares each real DMU with the defined set of sample/artificial DMUs and assigns a priority score to it. Subsequently, the DMUs are prioritized based on their score.

Sowlati et al. (2005) proposed the following model:

$$
\begin{array}{ll}
\max \sum_{r=1}^{s} u_r y_{r0} + u_0 & \\
\text{s.t.} & \\
\sum_{i=1}^{m} v_i x_{io} = 1 & \\
\sum_{r=1}^{s} u_r y_{rJ} + u_0 - \sum_{i=1}^{m} v_i p_J x_{iJ} \leq 0 \; J = 1, \ldots, N & \\
\sum_{r=1}^{s} u_r y_{ro} + u_0 - \sum_{i=1}^{m} v_i x_{io} \leq 0 & \\
u_r \geq \varepsilon & r = 1, \ldots, s \\
v_i \geq \varepsilon & i = 1, \ldots, m
\end{array}
\tag{4}
$$

where it is assumed that there are N artificial DMUs defined by decision makers with multiple inputs and multiple outputs: J: index for artificial DMUs ($J = 1, \ldots, N$),

x_{iJ}: ith input of artificial DMU$_J$ ($r = 1, \ldots, m$), y_{rJ}: rth input of artificial DMU$_J$ ($r = 1, \ldots, s$), p_J: user's assigned priority score of the artificial DMU$_J$, x_{io}: ith input of DMU$_o$ (DMU under consideration), y_{ro}: rth output of DMU$_o$.

This model should be solved once for each real DMU to rank them all. Model (4) compares each real DMU to the set of artificial DMUs and hence assessing the priority of a new added DMU would not affect the priority of already assessed ones. Sowlati et al. (2005) illustrated that using the Assurance Region method (Thompson et al. 1986), as any DEA model, Model (4) can be extended to consider managerial opinions and judgments about the relative importance of ranking criteria. In other words, imposing some suitable restrictions can control the factor weights along with the manager's opinion (see Sowlati et al. 2005, p. 1287). Traditional multiplier DEA models contain $n + 1$ constraints; however, there are $N + 2$ constraints in Model (4).

The non-Archimedean epsilon ε plays an important role in the DEA models (see Amin and Toloo 2004; Toloo 2014b). It seems Sowlati et al. (2005) practically ignored the role of epsilon in their research; because on one hand, the authors did not provide any approach to find a suitable value for the epsilon in their approach and, on the other hand, the authors presented some DEA models with $\varepsilon = 0$. For instance, consider the following traditional BCC input-oriented model which evaluates the performance of DMU$_o$ relative to n (real) units, i.e., DMU$_j$; $j = 1, \ldots, n$:

$$
\begin{aligned}
& \max \sum_{r=1}^{s} u_r y_{r0} + u_0 \\
& \text{s.t.} \\
& \sum_{i=1}^{m} v_i x_{io} = 1 \\
& \sum_{r=1}^{s} u_r y_{rj} + u_0 - \sum_{i=1}^{m} v_i p_J x_{ij} \le 0 \quad j = 1, \ldots, n \\
& u_r \ge \varepsilon \quad r = 1, \ldots, s \\
& v_i \ge \varepsilon \quad i = 1, \ldots, m
\end{aligned}
\tag{5}
$$

The dual form of Model (5) is expressed as bellow (see Sowlati et al. 2005, p. 1297) which is not a correct formulation based on the primal–dual relations in linear programming:

$$
\begin{aligned}
& \min \theta \\
& \text{s.t.} \\
& \sum_{j=1}^{n} \lambda_j x_{ij} \le \theta x_{io} \quad i = 1, \ldots, m \\
& \sum_{j=1}^{n} \lambda_j y_{rj} \ge y_{ro} \quad r = 1, \ldots, s \\
& \sum_{j=1}^{n} \lambda_j = 1 \\
& \lambda_j \ge 0 \quad j = 1, \ldots, n
\end{aligned}
\tag{6}
$$

As a matter of fact, the following model is the correct dual form (see Ali and Seiford 1993, p. 291)

$$\begin{aligned}
&\min \theta - \varepsilon\left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right) \\
&\text{s.t.} \\
&\sum_{j=1}^{n} \lambda_j x_{ij} - s_i^- = \theta x_{io} \quad i = 1, \ldots, m \\
&\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro} \quad r = 1, \ldots, s \\
&\sum_{j=1}^{n} \lambda_j = 1 \\
&s_i^- \geq 0 \qquad i = 1, \ldots, m \\
&s_r^+ \geq 0 \qquad r = 1, \ldots, s \\
&\lambda_j \geq 0 \qquad j = 1, \ldots, n
\end{aligned} \tag{7}$$

As inspection makes it clear, Model (6) is the dual of Model (5) if and only if $\varepsilon = 0$; however, it shows that the role of epsilon is disregarded in these models.

4 Proposed approach

The approach of Sowlati et al. (2005) addresses several drawbacks in its previous approaches (e.g., requiring less amount of input from the decision makers and hence simplicity). However, their approach does not treat all the DMUs on an equal footing since it includes solving one LP model for each alternative. This assumption could be problematic as in many real cases, IS project proposals are competing against each other and hence should be treated equally. In this paper, we propose a new DEA approach for addressing this drawback. The proposed approach finds most efficient units in the presence of managerial subjective judgments such as artificial DMUs with assigned priority scores. We illustrate that our approach evaluates all the DMUs on an equal footing and we show that our approach requires less amount of computation.

The first step of our approach finds a set of candidate DMUs to be the most efficient unit by solving the following model:

$$\begin{aligned}
&\min d_{max} \\
&\text{s.t.} \\
&d_{max} - d_j \geq 0 \qquad j = 1, \ldots, n \\
&\sum_{i=1}^{m} v_i x_{ij} \leq 1 \qquad j = 1, \ldots, n \\
&\sum_{r=1}^{s} u_r y_{rJ} + u_0 - \sum_{i=1}^{m} v_i p_J x_{iJ} \leq 0 \quad J = 1, \ldots, N \\
&\sum_{r=1}^{s} u_r y_{rj} + u_0 - \sum_{i=1}^{m} v_i x_{ij} + d_j = 0 \quad j = 1, \ldots, n \\
&d_j \geq 0 \qquad j = 1, \ldots, n \\
&u_r \geq \varepsilon \qquad r = 1, \ldots, s \\
&v_i \geq \varepsilon \qquad i = 1, \ldots, m
\end{aligned} \tag{8}$$

Here d_j is the deviation from efficiency of DMU_j and $d_{max} = \max\{d_j{:}j = 1, \ldots, n\}$. In order to measure the performance of $\text{DMU}_j(j = 1, \ldots, n)$, in the presence of user defined artificial $\text{DMU}_J(J = 1, \ldots, N)$, Model (8) minimizes the maximum deviation from efficiency. Suppose that the model is solved and the optimal solution $(\boldsymbol{v}^*, \boldsymbol{u}^*, u_o^*, \boldsymbol{d}^*, d_{max}^*)$ is at hand, where $\boldsymbol{v}^* = \left(v_1^*, \ldots, v_m^*\right) \in \mathbb{R}^m$, $\boldsymbol{u}^* = \left(u_1^*, \ldots, u_s^*\right) \in \mathbb{R}^s$, and $\boldsymbol{d}^* = \left(d_1^*, \ldots, d_n^*\right) \in \mathbb{R}^n$. Taking the common set of weights into consideration, the efficient frontier can be defined as $\boldsymbol{u}^*\boldsymbol{y} - u_0^* - \boldsymbol{v}^*\boldsymbol{x} = 0$ and subsequently DMU_k having minimum value of d_j is the closest unit into the efficient frontier. Therefore, the most efficient unit candidate is defined as follows:

Definition 1 DMU_k is a candidate for being most efficient unit if and only if $d_k^* = \min\{d_j^*{:}j = 1, \ldots, n\}$.

Indeed, Model (8) is a modified version of Model (2) which makes users able to import their intuitive sense by defining a set of artificial DMUs. This model is an aggregated model which evaluates all the DMUs using a common set of weights. Therefore, this model needs to be solved only once, to compare all the real DMUs with the artificial ones, and then finds a set of candidates for being the most efficient one. It should be mentioned here that, unlike Model (2), in this model d_k^* is not necessarily equal to zero and would have higher values. It is notable that if solving this model results in $d_k^* = 0$, it implies that DMU_k has a better performance than all the user defined artificial DMUs.

In order to find a suitable value for ε in Model (8), we propose the following LP model, which is an extended version of epsilon models suggested by Toloo (2012b) and Toloo and Nalchigar (2009):

$$
\begin{aligned}
&\varepsilon_1^* = \max \varepsilon \\
&\text{s.t.} \\
&\sum_{i=1}^{m} v_i x_{ij} \le 1 && j = 1, \ldots, n \\
&\sum_{r=1}^{s} u_r y_{rJ} + u_0 - \sum_{i=1}^{m} v_i p_J x_{iJ} \le 0 && J = 1, \ldots, N \\
&\sum_{r=1}^{s} u_r y_{rj} + u_0 - \sum_{i=1}^{m} v_i x_{ij} \le 0 && j = 1, \ldots, n \\
&\varepsilon - u_r \le 0 && r = 1, \ldots, s \\
&\varepsilon - v_i \le 0 && i = 1, \ldots, m \\
&\varepsilon \ge 0
\end{aligned}
\tag{9}
$$

Now we provide some interesting properties of Model (9).

Theorem 1 *Model* (9) *is always feasible.*

Proof A simple computation clarifies that $\left(\varepsilon^o, \boldsymbol{v^o}, \boldsymbol{u^o}, u_0^o\right) = (0, 0_{\boldsymbol{m}}, 0_{\boldsymbol{s}}, 0)$ is a feasible solution for Model (9) where $0_{\boldsymbol{m}} = (0, \ldots, 0) \in \mathbb{R}^m$. □

Theorem 2 *The optimal objective value of Model* (9) *is bounded.*

Proof From the first set of constraints in Model (9), i.e. $\sum_{i=1}^{m} v_i x_{ij} \leq 1$, for each feasible solution $\left(\varepsilon^o, \boldsymbol{v^o}, \boldsymbol{u^o}, u_0^o\right)$ we have $\min\{v_1^o, \ldots, v_m^o\} < \infty$. Without loss of generality suppose $v_1^o = \min\{v_1^o, \ldots, v_m^o\}$. Now, the first constraint in the last set of constraints of Model (9), i.e., $\varepsilon - v_1 \leq 0$, implies that $\varepsilon^* < \infty$. □

The following theorem verifies the feasibility of Model (8):

Theorem 3 *Model* (8) *is feasible for* $0 \leq \varepsilon \leq \varepsilon_1^*$.

Proof Suppose that $\left(\varepsilon^*, \boldsymbol{v^*}, \boldsymbol{u^*}, u_0^*, \boldsymbol{d}^*\right)$ is an optimal solution of Model (9). Note that such solution exist by Theorem 1. Let $d_{max} = \max\{d_1^*, \ldots, d_n^*\}$ and select $0 \leq \varepsilon \leq \varepsilon^*$ as a suitable value for the non-Archimedean epsilon value. As inspection makes clear, $(d_{max}, \boldsymbol{v^*}, \boldsymbol{u^*}, u_0^*, \boldsymbol{d}^*)$ is a feasible solution for Model (8) which completes the proof. □

Theorem 4 *Model* (8) *is infeasible for* $\varepsilon \in (\varepsilon_1^*, \infty)$.

Proof Suppose, contrary to our claim, that $(\bar{d}_{max}, \bar{\boldsymbol{v}}, \bar{\boldsymbol{u}}, \bar{u}_0, \bar{\boldsymbol{d}})$ is a feasible solution for Model (8) with $\bar{\varepsilon} > \varepsilon_1^*$. A trivial verification shows that $(\bar{\varepsilon}, \bar{\boldsymbol{v}}, \bar{\boldsymbol{u}}, \bar{u}_0, \bar{\boldsymbol{d}})$ is a feasible solution for Model (9) with an objective value, $\bar{\varepsilon}$, which is larger than the optimal objective value, ε_1^*, which is impossible. □

Definition 2 Let $E = \left\{k : d_k^* = \min d_j^* (j = 1, \ldots, n)\right\}$ resulted from solving Model (8). DMU$_k$ for $k \in E$ is a candidate for being most efficient DMU in presence of artificial, user defined DMUs.

If $|E_j| = 1$, then the most efficient DMU is identified. Otherwise, following model is proposed as second step of our approach for further analysis of candidate DMUs and finding the most efficient unit:

$$
\begin{array}{lr}
\min d_{max} & \\
\text{s.t.} & \\
d_{max} - d_j \geq 0 & j = 1, \ldots, n \\
\sum_{i=1}^{m} v_i x_{ij} \leq 1 & j = 1, \ldots, n \\
\sum_{r=1}^{s} u_r y_{rJ} + u_0 - \sum_{i=1}^{m} v_i p_J x_{iJ} \leq 0 & J = 1, \ldots, N \\
\sum_{r=1}^{s} u_r y_{rj} + u_0 - \sum_{i=1}^{m} v_i x_{ij} + d_j = 0 & j = 1, \ldots, n \\
\sum_{j=1}^{n} \theta_j = n - 1 & \\
\left(d_j - d_{min}\right) \leq M\theta_j & j = 1, \ldots, n \\
\theta_j \leq M\left(d_j - d_{min}\right) & j = 1, \ldots, n \\
\theta_j \in \{0, 1\} & j = 1, \ldots, n \\
d_j \geq 0 & j = 1, \ldots, n \\
u_r \geq \varepsilon & r = 1, \ldots, s \\
v_i \geq \varepsilon & i = 1, \ldots, m
\end{array}
\tag{10}
$$

where M is a large positive number, and the variable d_{min}, as will be explained shortly, equals min $\{d_j\colon j=1, \ldots, n\}$. This model is indeed based on Model (8) and includes additional constraints and variables to assure finding a single most efficient unit.

Theorem 5 *Model* (10) *finds the most efficient DMU in the presence of artificial user defined DMUs.*

Proof Without loss of generality, suppose that $k \in E$ and $\theta_k = 0$, then the constraint $(d_k - d_{min}) \leq M\theta_k$ leads to $d^*_{min} = d^*_k = \min d^*_j$, whereas the constraint $\theta_k \leq M(d_k - d_{min})$ is redundant. On the other hand, if $j \in E, j \neq k$, then the constraint $\theta_j \leq M\left(d_j - d_{min}\right)$ forces d_j to take a greater value than d_{min}, meanwhile the constraint $\left(d_j - d_{min}\right) \leq M\theta_j$ is redundant (for a large enough value of M). To summarize, in Model (10):

$$d_j \begin{cases} = d_{min}, & if\theta_j = 0 \\ > d_{min}, & if\theta_j = 1 \end{cases}$$

Referencing the constraint $\sum_{j=1}^{n} \theta_j = n-1$, there is a single DMU which is considered as the most efficient unit. □

The existence of alternative optimal solution is an issue which should be considered. To verify that whether there is another alternative optimal solution to Model (10), referencing to Toloo (2012b), we add a new constant $\theta_K = 1$ to the model and resolve the resulting model. If the optimal value increases, then there is no alternative optimal solution.

It is clear on inspection that the following model identifies the maximum non-Archimedean epsilon value for Model (10):

$$\begin{aligned}
&\varepsilon_2^* = \max \varepsilon \\
&\text{s.t.} \\
&\sum_{i=1}^{m} v_i x_{ij} \leq 1 && j = 1, \ldots, n \\
&\sum_{r=1}^{s} u_r y_{rJ} + u_0 - \sum_{i=1}^{m} v_i p_J x_{iJ} \leq 0 && J = 1, \ldots, N \\
&\sum_{r=1}^{s} u_r y_{rj} + u_0 - \sum_{i=1}^{m} v_i x_{ij} + d_j = 0 && j = 1, \ldots, n \\
&\sum_{j=1}^{n} \theta_j = n - 1 \\
&\left(d_j - d_{min}\right) \leq M\theta_j && j = 1, \ldots, n \\
&\theta_j \leq M\left(d_j - d_{min}\right) && j = 1, \ldots, n \\
&\varepsilon - u_r \leq 0 && r = 1, \ldots, s \\
&\varepsilon - v_i \leq 0 && i = 1, \ldots, m \\
&\theta_j \in \{0, 1\} && j = 1, \ldots, n \\
&d_j \geq 0 && j = 1, \ldots, n \\
&\varepsilon \geq 0
\end{aligned} \tag{11}$$

The following theorems can be proved in much the same way as Theorems 1–3.

Theorem 6 *Model* (11) *is always feasible.*

Theorem 7 *The optimal objective value of Model* (11) *is bounded.*

Theorem 8 *Model* (10) *is feasible for* $0 \leq \varepsilon \leq \varepsilon_2^*$.

Theorem 9 *Model* (10) *is infeasible for* $\varepsilon \in (\varepsilon_2^*, \infty)$.

The following theorem verifies the relationship between the optimal objective value of Models (9) and (11):

Theorem 10 $0 \leq \varepsilon_2^* \leq \varepsilon_1^*$.

Proof Let S_1 and S_2 be the feasible region of Models (9) and (11), respectively. In other words:

$$S_1 = \left\{ (\varepsilon, \boldsymbol{v}, \boldsymbol{u}, u_0) \left| \begin{array}{l} \boldsymbol{u}\boldsymbol{y}_J + u_0 - p_J \boldsymbol{v}\boldsymbol{x}_J \leq 0\,(\forall J)\,, \boldsymbol{u}\boldsymbol{y}_j + u_0 - \boldsymbol{v}\boldsymbol{x}_j \leq 0\,(\forall j) \\ \boldsymbol{v}\boldsymbol{x}_j \leq 1\,(\forall j)\,, \varepsilon - v_i \leq 0\,(\forall i)\,, \varepsilon - u_r \leq 0\,(\forall r) \end{array} \right. \right\}$$

$$S_2 = \left\{ (d_{min}, \boldsymbol{d}, \boldsymbol{\theta}, \boldsymbol{v}, \boldsymbol{u}, u_0) \left| \begin{array}{l} \boldsymbol{v}\boldsymbol{x}_j \leq 1\,(\forall j)\,, \boldsymbol{u}\boldsymbol{y}_J + u_0 - p_J \boldsymbol{v}\boldsymbol{x}_J \leq 0\,(\forall J)\,, \boldsymbol{u}\boldsymbol{y}_j + u_0 - \boldsymbol{v}\boldsymbol{x}_j + d_j = 0\,(\forall j) \\ 1_n \boldsymbol{\theta} = n - 1, \left(d_j - d_{min}\right) \leq M\theta_j\,(\forall j)\,, \theta_j \leq M\left(d_j - d_{min}\right)(\forall j) \\ v_i \geq \varepsilon\,(\forall i)\,, u_r \geq \varepsilon\,(\forall r)\,, d_j \geq 0\,(\forall j)\,, \theta_j \geq 0 (\forall j) \end{array} \right. \right\}$$

where $1_n = (1, \ldots, 1) \in \mathbb{R}^n$. In addition, let $\bar{\varepsilon}$ be the given (obtained) non-Archimedean epsilon and $(\bar{d}_{min}, \bar{\boldsymbol{d}}, \bar{\boldsymbol{\theta}}, \bar{\boldsymbol{v}}, \bar{\boldsymbol{u}}, \bar{u}_0)$ be a feasible solution for Model (11). An easy computation clears that $(\bar{\varepsilon}, \bar{\boldsymbol{v}}, \bar{\boldsymbol{u}}, \bar{u}_0)$ is a feasible solution for Model (9) and since the objective function of Models (9) and (11) are identical, and these models are the maximization type, we arrive at $\varepsilon_2^* \leq \varepsilon_1^*$. Note that the reverse is not always true which completes the proof. □

In general, we can summarize the merits of our new approach from both technical and computational points of view:

1. Technically: Sowlati et al.'s approach evaluates all DMUs by different sets of weights and hence could be problematic as in many real cases, IS project proposals are competing against each other and hence should be treated in an identical situation. However, our model treats all IS projects on an equal footing which makes the IS project evaluation more realistic. Moreover, the common weight models render more discriminating power among efficient DMUs (for more details see Cook et al. 1996).

2. Computationally: In our approach, there is no need to run a model for each DMU and hence it is simpler and more efficient than the Sowlati et al.'s approach. To be more specific, consider the following standard form of Model (4):

$$
\begin{array}{ll}
\max \sum_{r=1}^{s} u_r y_{r0} + u_0^+ - u_0^- & \\
\text{s.t.} & \\
\sum_{i=1}^{m} v_i x_{io} = 1 & \\
\sum_{r=1}^{s} u_r y_{rJ} + u_0^+ - u_0^- - \sum_{i=1}^{m} v_i p_J x_{iJ} - t_J = 0 & J = 1, \ldots, N \\
\sum_{r=1}^{s} u_r y_{ro} + u_0^+ - u_0^- - \sum_{i=1}^{m} v_i x_{io} - t = 0 & \\
u_r - s_r^y = \varepsilon & r = 1, \ldots, s \\
v_i - s_i^x = \varepsilon & i = 1, \ldots, m \\
u_r, s_r^y \geq 0 & r = 1, \ldots, s \\
v_i, s_i^x \geq 0 & i = 1, \ldots, m \\
t_J \geq 0 & J = 1, \ldots, N \\
t, u_0^+, u_0^- \geq 0 &
\end{array}
\tag{12}
$$

where there are $N+2(m+s)+3$ variables and $N+m+s+2$ constraints. According to Bazaraa et al. (2010), it is often empirically suggested that on the average in most instances, the simplex method for solving an LP with p decision variables and q constraints (in the standard from) requires roughly on the order of q to $3q$ iterations. Each iteration needs $q(p-q)+p+1$ multiplications and $q(p-q+1)$ additions. Analogously, Model (12) requires on the order of $N+m+s+2$ to $3(N+m+s+2)$ iterations and in each iteration it needs $(N+m+s+2)(m+s+1)+N+2(m+s)+4$ multiplications and $(N+m+s+2)(m+s+1)$ additions (see Toloo et al. 2015) for more details). As a result, a single run of the proposed MILP model (10) needs significantly less computations than n times running of model (4), i.e., one run for each real DMU.

5 Illustration

We consider a real case of ranking IS projects at a large financial institution which is adapted from Sowlati et al. (2005). The case study included 41 real IS projects to be ranked. Also, the decision makers had defined 18 artificial projects to be used for ranking the real ones. There are four inputs for each DMU: Time to market (I_1), Green dollar costs (I_2), Brown dollar costs (I_3), and Potential risks (I_4). Also, there are four outputs for each DMU: Breath of benefits (O_1), Intangible benefits (O_2), Green dollar benefits (O_3), and Brown dollar benefits (O_4). Readers are referred to Sowlati et al. (2005) for definition of each input and output. Tables 1 and 2 present the data of real and artificial IS projects, respectively. Using Model (4), Sowlati et al. (2005) ranked

Table 1 Data of 41 real IS projects Sowlati et al. (2005)

IS projects	DEA inputs				DEA outputs			
DMUs	I_1	I_2	I_3	I_4	O_1	O_2	O_3	O_4
1	20	1	30	60	50	100	1	100
2	10	1	30	40	100	60	1	60
3	40	1	20	1	50	100	1	20
4	60	1	20	1	100	100	1	1
5	40	30	50	40	100	100	1	80
6	90	1	70	60	75	100	1	100
7	60	1	20	40	90	80	1	20
8	50	10	20	40	90	100	1	20
9	60	1	40	20	75	60	1	50
10	50	1	20	1	80	40	1	20
11	20	20	30	60	80	100	1	40
12	20	20	10	1	25	100	1	20
13	30	30	20	40	100	60	1	40
14	40	70	30	1	100	100	1	80
15	40	20	30	60	50	80	1	60
16	60	1	30	40	90	60	1	20
17	90	1	20	1	100	40	1	10
18	60	1	30	20	50	100	1	10
19	60	30	30	40	50	100	1	60
20	80	1	80	80	75	100	1	80
21	40	20	20	20	75	60	1	20
22	40	20	20	20	75	60	1	20
23	90	1	10	1	25	40	1	20
24	90	1	20	40	75	40	1	20
25	90	1	20	60	100	1	1	30
26	30	1	30	20	75	20	1	1
27	90	1	30	10	50	60	1	20
28	90	1	20	20	50	40	1	20
29	100	1	20	1	75	20	1	10
30	100	1	50	1	75	40	1	1
31	60	10	20	40	75	20	1	20
32	90	10	20	60	75	60	1	20
33	60	20	20	40	75	40	1	20
34	100	1	30	1	75	20	1	1
35	100	1	20	1	50	20	1	1
36	60	10	20	60	20	40	1	20
37	100	10	30	1	50	20	1	1

Table 1 continued

IS projects	DEA inputs				DEA outputs			
DMUs	I_1	I_2	I_3	I_4	O_1	O_2	O_3	O_4
38	90	20	30	60	100	20	1	20
39	100	1	40	20	50	40	1	1
40	100	1	20	1	1	20	1	1
41	1	100	20	1	1	20	1	1

The data involves a constant output O_3. See "Appendix A" to see the role of a constant input/output in the original BCC model, i.e., without the user subjective opinions

the IS projects in Table 1 and found that DMU14 achieves the highest rank among all 41 competing proposals. Data of this section is used later in this paper to demonstrate application of our proposed approach.

Using GAMS operations research software,[1] we solve the proposed Models (8)–(11) for the data of real and artificial IS projects. The maximum value of non-Archimedean epsilon obtained from Model (9) is $\varepsilon_1^* = 0.004149$. Using this value and solving Model (8) for the dataset we get $d_1^* = d_2^* = d_{14}^* = \min\left\{d_j^*, j = 1, \ldots, 41\right\} = 1.0622$ or equivalently $E = \{1, 2, 14\}$, which implies that DMU_1, DMU_2, and DMU_{14} are suitable candidates for being most efficient IS projects. Because of $|E| > 1$, Model (8) fails to discriminate the most efficient IS project and Model (10) should be utilized.

Solving Model (10) with $\varepsilon_2^* = 5.176 \times 10^{-5}$ results in $d_{14}^* = d_{min}^* = 0.373$, and $\theta_{14}^* = 0$. This indicates that DMU_{14} is the most efficient unit. It is important to make sure that there is no alternative solution for Model (10) (for a deeper discussion of alternative solutions in DEA we refer the readers to Toloo 2012a). To verify that, we add a new constraint $\theta_{14} = 1$ to the model and repeat the calculation. The new result is $d_{max}{}^* = 0.391$, showing that the optimal objective value is increased and hence there is no alternative optimal solution.

Table 3 presents the complete results of Model (10) and compares it with some similar approaches: super-efficiency, cross-efficiency, Sowlati et al. (2005) and Toloo (2012b). It can be seen that DMU_{14} is also ranked as the first IS project in super-efficiency and Sowlati et al. (2005). Moreover, the super-efficiency model for DMU_6 is infeasible which shows one of the main issues of the method (For more details see Lawrence and Zhu 1999 and Lee et al. 2011). On the other hand, cross-efficiency picks DMU_2 because the method considers the average self-evaluated and peer-evaluated scores for each DMU. Toloo (2012b) selects a different IS project because its formulation does not capture the subjective opinions of decision makers. In comparison with Sowlati et al.'s approach, our approach treats all DMUs an equal footing. The proposed DEA models obtain a common set of optimal weights in the evaluation of all DMUs and hence treat them equally. Since in many IS project selection situations all the proposals are competing together to get approved and receive resources, it makes

1 Available for free at www.gams.com.

Table 2 Data of 18 artificial IS projects defined by decision makers (Sowlati et al. 2005)

Artificial IS projects	DEA inputs				DEA outputs				User assigned priority score
DMUs	I_1	I_2	I_3	I_4	O_1	O_2	O_3	O_4	
A	1	1	1	1	1	100	100	100	0.75
B	100	100	100	100	100	1	1	1	0.05
C	1	1	1	1	100	1	100	100	0.65
D	100	100	100	100	1	100	1	1	0.15
E	100	100	100	100	1	1	1	1	0.01
F	1	1	1	1	100	100	1	100	0.55
G	100	100	100	100	1	1	100	1	0.25
H	1	1	1	1	100	100	100	100	1
I	1	1	1	1	100	100	100	1	0.6
G	100	100	100	100	1	1	1	100	0.2
K	100	1	1	1	100	100	100	100	0.7
L	1	100	100	100	1	1	1	1	0.05
M	100	1	100	100	1	1	1	1	0.25
N	1	100	1	1	100	100	100	100	0.55
O	100	100	1	100	1	1	1	1	0.2
P	1	1	100	1	100	100	100	100	0.6
Q	100	100	100	1	1	1	1	1	0.1
R	1	1	1	100	100	100	100	100	0.5

Table 3 A comparison between five approaches

IS projects	Super-efficiency		Cross-efficiency		Sowlati et al. (2005)		Toloo (2012a, b)		New approach	
DMUs	Score	Rank	Score	Rank	Score	Rank	d_j^*	Rank	d_j^*	Rank
1	4.5	5	0.8928	2	0.534	15	0	1	0.473	2
2	18.4	3	0.9242	1	0.528	16	0.002	3	0.474	3
3	7.73	4	0.8604	4	0.548	3	0.133	6	0.475	4
4	35.06	2	0.8776	3	0.549	2	0.088	5	0.476	5
5	2.33	8	0.3096	31	0.018	40	0.080	4	0.478	6
6	Infeasible	?	0.6093	11	0.517	17	0.353	16	0.590	15
7	1.04	16	0.7467	6	0.501	20	0.294	11	0.606	19
8	1.06	15	0.411	25	0.054	29	0.207	9	0.576	11
9	1.15	14	0.654	8	0.471	23	0.314	13	0.545	9
10	3.1	7	0.8003	5	0.548	5	0.299	12	0.531	8
11	1.6	11	0.3447	28	0.027	33	0.206	8	0.597	17
12	2	9	0.4917	20	0.548	4	0.191	7	0.492	7
13	1.28	12	0.393	26	0.026	34	0.248	10	0.573	10
14	40	1	0.4218	24	0.549	1	0.001	2	0.373	1
15	0.7	39	0.2865	36	0.026	35	0.413	19	0.649	21
16	1	17	0.6132	10	0.501	21	0.419	20	0.649	23
17	4.39	6	0.7314	7	0.548	6	0.424	21	0.582	12
18	1	18	0.5764	15	0.538	14	0.378	18	0.602	18
19	0.71	38	0.2718	37	0.017	41	0.371	17	0.593	16
20	1	19	0.5289	17	0.507	19	0.520	23	0.705	29
21	0.86	32	0.3235	30	0.026	36	0.353	14	0.586	13
22	0.86	33	0.3235	29	0.026	37	0.353	14	0.586	14

Table 3 continued

IS projects	Super-efficiency		Cross-efficiency		Sowlati et al. (2005)		Toloo (2012a, b)		New approach	
DMUs	Score	Rank	Score	Rank	Score	Rank	d_j^*	Rank	d_j^*	Rank
23	1.95	10	0.6286	9	0.547	9	0.651	28	0.649	22
24	1	20	0.5786	14	0.461	25	0.647	27	0.726	33
25	1.2	13	0.5888	12	0.516	18	0.746	34	0.792	39
26	1	21	0.5013	19	0.470	24	0.519	22	0.662	25
27	1	22	0.5232	18	0.494	22	0.585	24	0.644	20
28	1	23	0.5335	16	0.446	26	0.668	30	0.692	26
29	1	24	0.5881	13	0.547	7	0.652	29	0.659	24
30	1	25	0.4421	23	0.547	8	0.731	32	0.693	27
31	0.85	34	0.2889	33	0.052	31	0.643	26	0.726	32
32	0.83	35	0.2909	32	0.053	30	0.684	31	0.778	35
33	0.83	36	0.2679	38	0.026	39	0.602	25	0.711	30
34	1	26	0.4782	21	0.547	10	0.731	32	0.694	28
35	1	27	0.4568	22	0.547	11	0.793	35	0.715	31
36	0.73	37	0.1824	40	0.052	32	0.871	37	0.841	41
37	1	28	0.168	41	0.389	28	0.872	38	0.741	34
38	0.67	40	0.216	39	0.026	38	0.830	36	0.828	40
39	1	29	0.3662	27	0.446	27	0.872	38	0.778	36
40	1	30	0.2888	35	0.547	12	0.997	40	0.782	37
41	1	31	0.2888	34	0.547	13	0.997	41	0.782	37

more sense to evaluate them in an identical condition. Moreover, our approach does not need to solve one individual LP problem for each of the alternatives. It is while the super-efficiency, cross-efficiency, and Sowlati et al. (2005) approaches require n LP models to be solved (i.e., treating alternatives differently), where n is total number of alternative projects ($n = 41$ in the case study). Our approach assist decision makers by a fairly less number of computations. Based on our computations, 158 iterations used by the CEPLEX solver of GAMS software to solve all the 41 LP models and the total number of multiplications and additions is 44,240 and 39,816, respectively.

6 Conclusion

IS project selection is of great importance in today's enterprises for their operational excellence, profitability, competitive advantage, and hence survivability in current dynamic environment. This paper proposed a new DEA approach for evaluation and selection of most efficient IS projects. The proposed approach makes users able to import their subjective opinions and intuitive senses regarding a set of artificial alternatives that represent good or bad IS projects. The proposed model then compares the real alternatives with the set of artificial ones and finds the most efficient project. Applicability of proposed approach is illustrated on a real-world case study and data set obtained from a previous study. Results indicate that the proposed approach provides a fair and equitable evaluation.

Future works can extend the proposed approach to fuzzy data environments and make it able to receive decision makers' judgments in terms of linguistic terms. Moreover, to utilize the proposed approach for full ranking of DMUs is another promising avenue for future research. It worth mentioning that although this paper focused on IS project selection problem, it contributes to the field of DEA by introducing a new integrated approach for finding most efficient units. Additionally, the proposed DEA approach could be applied in the broader area of enterprise decision making besides enterprise IS project. These extensions are left for future research. Recently, Toloo and Salahi (2018) formulated an interesting powerful discriminative approach for selecting the most efficient unit in DEA which can be extended for taking the subjective opinions and intuitive senses of decision makers into consideration. Finding the most efficient DMU under uncertainty in another interesting area of research (for a deeper discussion we refer the readers to Toloo et al. 2018).

Acknowledgements The research was supported by the European Social Fund (CZ.1.07/2.3.00/20.0296) and the Czech Science Foundation (GAČR 16-17810S). All support is greatly acknowledged.

Appendix A

Consider the following input-oriented BCC model:

$$
\begin{aligned}
&\theta_o^* = \max \sum_{r=1}^{s} u_r y_{r0} + u_0 \\
&\text{s.t.} \\
&\sum_{i=1}^{m} v_i x_{io} = 1 \\
&\sum_{r=1}^{s} u_r y_{rj} + u_0 - \sum_{i=1}^{m} v_i x_{ij} \le 0 \quad j = 1, \ldots, n \\
&u_r \ge \varepsilon \quad r = 1, \ldots, s \\
&v_i \ge \varepsilon \quad i = 1, \ldots, m
\end{aligned}
\tag{A.1}
$$

Without loss of generality, suppose $y_{sj} = c$ for $j = 1, \ldots, n$. Model (A.1) can be rewritten as:

$$
\begin{aligned}
&\theta_o^* = \max \sum_{r=1}^{s-1} u_r y_{r0} + u_s c + u_0 \\
&\text{s.t.} \\
&\sum_{i=1}^{m} v_i x_{io} = 1 \\
&\sum_{r=1}^{s-1} u_r y_{rj} + u_s c + u_0 - \sum_{i=1}^{m} v_i x_{io} \le 0 \quad j = 1, \ldots, n \\
&u_r \ge \varepsilon \quad r = 1, \ldots, s \\
&v_i \ge \varepsilon \quad i = 1, \ldots, m
\end{aligned}
\tag{A.2}
$$

Now, by making the change of variable $\bar{u}_0 = u_s c + u_0$ we obtain the following model which is equivalent to the BCC model (A.1):

$$
\begin{aligned}
&\theta_o^* = \max \sum_{r=1}^{s-1} u_r y_{r0} + \bar{u}_0 \\
&\text{s.t.} \\
&\sum_{i=1}^{m} v_i x_{io} = 1 \\
&\sum_{r=1}^{s-1} u_r y_{rj} + \bar{u}_0 - \sum_{i=1}^{m} v_i x_{io} \le 0 \quad j = 1, \ldots, n \\
&u_r \ge \varepsilon \quad r = 1, \ldots, s-1 \\
&v_i \ge \varepsilon \quad i = 1, \ldots, m
\end{aligned}
\tag{A.3}
$$

Hence, the following theorem is proved:

Theorem A.1 *A constant output in the input-oriented BCC model is redundant.*

Analogously, the following theorem can be proved:

Theorem A.2 *A constant input in the output-oriented BCC model is redundant.*

References

Adler N, Friedman L, Sinuany-Stern Z (2002) Review of ranking methods in the data envelopment analysis context. Eur J Oper Res 140(2):249–265

Ali A, Seiford L (1993) Computational accuracy and infinitesimals in data envelopment analysis. INFOR 31(4):290–297

Amin GR (2009) Comments on finding the most efficient DMUs in DEA: an improved integrated model. Comput Ind Eng 56(4):1701–1702

Amin GR, Toloo M (2004) A polynomial-time algorithm for finding epsilon in DEA models. Comput Oper Res 31(5):803–805

Amin GR, Toloo M (2007) Finding the most efficient DMUs in DEA: an improved integrated model. Comput Ind Eng 56(2):71–77

Andersen P, Petersen NC (1993) A procedure for ranking efficient units in data envelopment analysis. Manage Sci 39(10):1261–1264

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

Asosheh A, Nalchigar S, Jamporazmey M (2010) Information technology project evaluation: an integrated data envelopment analysis and balanced scorecard approach. Expert Syst Appl 37(8):5931–5938

Badri MA, Davis D, Davis D (2001) A comprehensive 0–1 goal programming model for project selection. Int J Project Manage 19(4):243–252

Bai H, Zhan Z (2011) An IT Project selection method based on fuzzy analytic network process. In: 2011 International conference on system science, engineering design and manufacturing information, vol 2, pp 275–279

Baker RC, Talluri S (1997) A closer look at the use of data envelopment analysis for technology selection. Comput Ind Eng 32(1):101–108

Banker RD, Charnes A, Cooper WW (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. Manage Sci 30(9):1078–1092

Bazaraa MS, Jarvis JJ, Sherali HD (2010) Linear programming and network flows, 4th edn. Wiley, New York

Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2:429–444

Chen C-T, Cheng H-L (2009) A comprehensive model for selecting information system project under fuzzy environment. Int J Project Manage 27(4):389–399

Cook WD, Roll Y, Kazakov A (1990) A DEA model for measuring the relative efficiency of highway maintenance patrols. INFOR 28(2):113–124

Cook WD, Kress M, Seiford LM (1996) Data envelopment analysis in the presence of both quantitative and qualitative factors. J Oper Res Soc 47(7):945–953

Deng H, Wibowo S (2008) Intelligent decision support for evaluating and selecting information systems projects. Eng Lett 16(3):412–418

Doyle J, Green RH (1994) Efficiency and cross-efficiency in DEA: derivations, meanings and uses. J Oper Res Soc 45(5):567–578

Edirisinghe NCP, Zhang X (2007) Generalized DEA model of fundamental analysis and its application to portfolio optimization. J Bank Finance 31(11):3311–3335

Ertay T, Ruan D, Tuzkaya U (2006) Integrating data envelopment analysis and analytic hierarchy for the facility layout design in manufacturing systems. Inf Sci 176(3):237–262

Farzipoor Saen R (2011) Media selection in the presence of flexible factors and imprecise data. J Oper Res Soc 62(9):1695–1703

Gao P, Feng J, Yang L (2008) Fuzzy TOPSIS algorithm for multiple criteria decision making with an application in information systems project selection. In: WiCOM'08 4th international conference on wireless communications, networking and mobile computing, 2008, pp 1–4

Han CH, Kim JK, Choi SH, Kim SH (1998) Determination of information system development priority using quality function development. Comput Ind Eng 35(1):241–244

Holtsnider B, Jaffe BD (2012) IT manager's handbook: getting your new job done. Morgan Kaufmann, Burlington

Hou G (2011) IT/IS project selection: a grey multi-criteria decision model approach. In: 2011 international conference on E-business and E-government (ICEE), pp 1–4

Jablonsky J (2018) Ranking of countries in sporting events using two-stage data envelopment analysis models: a case of Summer Olympic Games 2016. Cent Eur J Oper Res. https://doi.org/10.1007/s10100-018-0537-8

Kapelko M (2018) Measuring inefficiency for specific inputs using data envelopment analysis: evidence from construction industry in Spain and Portugal. CEJOR 26(1):43–66

Karsak EE, Özogul CO (2009) An integrated decision making approach for ERP system selection. Expert Syst Appl 36(1):660–667

Kengpol A, Tuominen M (2006) A framework for group decision support systems: an application in the evaluation of information technology for logistics firms. Int J Prod Econ 101(1):159–171

Khoshnevis P, Teirlinck P (2018) Performance evaluation of R&D active firms. Socio Econ Plann Sci 61:16–28

Lall V, Teyarachakul S (2006) Enterprise Resource Planning (ERP) system selection: a Data Envelopment Anaysis (DEA) approach. J Comput Inf Syst 47(1):123–127

Lawrence M, Zhu J (1999) Infeasibility of super-efficiency data envelopment analysis models. Inf Syst Oper Res 37(2):174–188

Lee JW, Kim SH (2001) An integrated approach for interdependent information system project selection. Int J Project Manage 19(2):111–118

Lee H-S, Chu C-W, Zhu J (2011) Super-efficiency DEA in the presence of infeasibility. Eur J Oper Res 212(1):141–147

Masoumzadeh A, Toloo M, Amirteimoori A (2016) Performance assessment in production systems without explicit inputs: an application to basketball players. IMA J Manag Math 27(2):143–156

Nakhaeizadeh G, Schnabl A (1997) Development of multi-criteria metrics for evaluation of data mining algorithms. In: Proceedings of the third international conference on knowledge discovery and data mining. AAAI Press

Nalchigar S, Nasserzadeh SMR (2009) Application of DEA for selecting most efficient information system project with imprecise data. In: 2009 IEEE international conference on industrial engineering and engineering management. IEEE Publisher, pp 1653–1657

Paradi JC, Zhu H (2013) A survey on bank branch efficiency and performance research with data envelopment analysis. Omega 41(1):61–79

Ramón N, Ruiz JL, Sirvent I (2012) Common sets of weights as summaries of DEA profiles of weights: with an application to the ranking of professional tennis players. Expert Syst Appl 39(5):4882–4889

Roll Y, Cook WD, Golany B (1991) Controlling factor weights in data envelopment analysis. IIE Trans 23(1):2–9

Santhanam R, Kyparisis J (1995) A multiple criteria decision model for information system project selection. Comput Oper Res 22(8):807–818

Sarkis J, Sundarraj RP (2006) Evaluation of enterprise information technologies: a decision model for high-level consideration of strategic and operational issues. IEEE Trans Syst Man Cybern Part C Appl Rev 36(2):260–273

Schniederjans MJ, Santhanam R (1993) A multi-objective constrained resource information system project selection method. Eur J Oper Res 70(2):244–253

Schniederjans MJ, Wilson RL (1991) Using the analytic hierarchy process and goal programming for information system project selection. Inf Manag 20(5):333–342

Sexton TR, Silkman RH, Hogan AJ (1986) Data envelopment analysis: critique and extensions. N Dir Program Eval 1986(32):73–105

Shafer SM, Byrd TA (2000) A framework for measuring the efficiency of organizational investments in information technology using data envelopment analysis. Omega 28(2):125–141

Shang J, Sueyoshi T (1995) A unified framework for the selection of a flexible manufacturing system. Eur J Oper Res 85(2):297–315

Sohrabi B, Toloo M, Moeini A, Nalchigar S (2015) Evaluation of recommender systems: a multi-criteria decision making approach. Iran J Manag Stud 8(4):589–605

Sowlati T, Paradi JC, Suld C (2005) Information systems project prioritization using data envelopment analysis. Math Comput Model 41(11):1279–1298

Thompson RG, Singleton FD, Thrall RM, Smith BA (1986) Comparative site evaluations for locating a high-energy physics lab in Texas. Interfaces 16(6):35–49

Toloo M (2012a) Alternative solutions for classifying inputs and outputs in data envelopment analysis. Comput Math Appl 63(6):1104–1110

Toloo M (2012b) On finding the most BCC-efficient DMU: a new integrated MIP-DEA model. Appl Math Model 36(11):5515–5520

Toloo M (2013) The most efficient unit without explicit inputs: an extended MILP-DEA model. Measurement J Int Measurement Confed 46(9):3628–3634

Toloo M (2014a) An epsilon-free approach for finding the most efficient unit in. Appl Math Model 38(13):3182–3192

Toloo M (2014b) The role of non-Archimedean epsilon in finding the most efficient unit: with an application of professional tennis players. Appl Math Model 38(21–22):5334–5346

Toloo M (2015) Alternative minimax model for finding the most efficient unit in data envelopment analysis. Comput Ind Eng 81:186–194

Toloo M (2016) A cost efficiency approach for strategic vendor selection problem under certain input prices assumption. Measurement 85:175–183

Toloo M, Ertay T (2014) The most cost efficient automotive vendor with price uncertainty: a new DEA approach. Measurement 52(1):135–144

Toloo M, Kresta A (2014) Finding the best asset financing alternative: a DEA-WEO approach. Measurement 55:288–294

Toloo M, Nalchigar S (2009) A new integrated DEA model for finding most BCC-efficient DMU. Appl Math Model 33(1):597–604

Toloo M, Nalchigar S (2011) A new DEA method for supplier selection in presence of both cardinal and ordinal data. Expert Syst Appl 38(12):14726–14731

Toloo M, Salahi M (2018) A powerful discriminative approach for selecting the most efficient unit in DEA. Comput Ind Eng 115:269–277

Toloo M, Tavana M (2017) A novel method for selecting a single efficient unit in data envelopment analysis without explicit inputs/outputs. Ann Oper Res 253(1):657–681

Toloo M, Masoumzadeh A, Barat M (2015) Finding an initial basic feasible solution for DEA models with an application on bank industry. Comput Econ 45(2):323–336

Toloo M, Tavana M, Santos-Arteaga FJ (2017) An integrated data envelopment analysis and mixed integer non-linear programming model for linearizing the common set of weights. Cent Eur J Oper Res. https://doi.org/10.1007/s10100-017-0510-y

Toloo M, Keshavarz E, Hatami-Marbini A (2018) Dual-role factors for imprecise data envelopment analysis. Omega 77:15–31

Yang C-L, Chiang S-J, Huang R-H, Lin Y-A (2013) Hybrid decision model for information project selection. Qual Quant 47(4):2129–2142

Yeh C-H, Deng H, Wibowo S, Xu Y (2010) Fuzzy multicriteria decision support for information systems project selection. Int J Fuzzy Syst 12(2):170–174

Zhou P, Ang BW, Poh KL (2008) A survey of data envelopment analysis in energy and environmental studies. Eur J Oper Res 189(1):1–18