A fuzzy multidimensional multiple-choice knapsack model for project portfolio selection using an evolutionary algorithm

Madjid Tavana, Kaveh Khalili-Damghani & Amir-Reza Abtahi

Annals of Operations Research

ISSN 0254-5330
Volume 206
Number 1

Ann Oper Res (2013) 206:449-483
DOI 10.1007/s10479-013-1387-3

Ann Oper Res (2013) 206:449–483
DOI 10.1007/s10479-013-1387-3

A fuzzy multidimensional multiple-choice knapsack model for project portfolio selection using an evolutionary algorithm

Madjid Tavana · Kaveh Khalili-Damghani · Amir-Reza Abtahi

Published online: 10 May 2013

Abstract Project portfolio selection problems are inherently complex problems with multiple and often conflicting objectives. Numerous analytical techniques ranging from simple weighted scoring to complex mathematical programming approaches have been proposed to solve these problems with precise data. However, the project data in real-world problems are often imprecise or ambiguous. We propose a fuzzy Multidimensional Multiple-choice Knapsack Problem (MMKP) formulation for project portfolio selection. The proposed model is composed of an Efficient Epsilon-Constraint (EEC) method and a customized multi-objective evolutionary algorithm. A Data Envelopment Analysis (DEA) model is used to prune the generated solutions into a limited and manageable set of implementable alternatives. Statistical analysis is performed to investigate the effectiveness of the proposed approach in comparison with the competing methods in the literature. A case study is presented to demonstrate the applicability of the proposed model and exhibit the efficacy of the procedures and algorithms.

Keywords Multidimensional multiple-choice knapsack · Multi-objective evolutionary algorithm · Efficient epsilon-constraint · Data envelopment analysis

M. Tavana (✉)
Business Systems and Analytics, Lindback Distinguished Chair of Information Systems and Decision Sciences, La Salle University, Philadelphia, PA 19141, USA
e-mail: tavana@lasalle.edu
url: http://tavana.us/

K. Khalili-Damghani
Department of Industrial Engineering, Islamic Azad University South-Tehran Branch, Tehran, Iran
e-mail: kaveh.khalili@gmail.com
url: http://kaveh-khalili.webs.com

A.-R. Abtahi
Department of Knowledge Engineering and Decision Sciences, Faculty of Economic Institutions Management, University of Economic Sciences, Tehran, Iran
e-mail: amir_abtahi@yahoo.com

1 Introduction

The knapsack problem is best described by a problem where a hitchhiker is searching for a combination of different items that maximizes the total value of all chosen items in his/her knapsack. The multiple objective knapsack problem is a well-known combinatorial optimization problem with a wide range of business applications in capital budgeting (Bas 2011) and production planning (Camargo et al. 2012; Moreno et al. 2010) among others. A good review of the single knapsack problem and its associated exact and heuristic algorithms is available in Martello and Toth (1990).

The Multidimensional Multiple-choice Knapsack Problem (MMKP) (as a generalization of the single knapsack problem) is one of the most complex forms of the knapsack problem family. Given a set of knapsacks with limited resources and some disjoint groups of items, the MMKP aims to fill the knapsacks by picking exactly one item from each group, such that the total profit value of the collected items is maximized and none of the resource constraints are violated (Crévits et al. 2012; Han et al. 2010; Ghasemi and Razzazi 2011). The MMKPs are NP-hard optimization problems where the computing time grows exponentially with problem size (Kellerer et al. 2004). Several heuristic and meta-heuristic approaches have been proposed to solve MOMDKP (Florios et al. 2010; Shah and Reed 2011; Aghezzaf and Naimi 2009; Alves and Almeida 2007).

Data Envelopment Analysis (DEA) has been used to compare project efficiencies in multi-project environments (Cook and Green 2000; Farris et al. 2006; Vitner et al. 2006; El-Mashaleh et al. 2010; Ghapanchi et al. 2012). However, the inability to clearly discriminate among the outputs often hinders its popularity and wide-spread use in project selection problems. Alternatively, the knapsack formulation could be used to formulate the problem of maximizing profit under constrained resources (Lin and Yao 2001; Lin 2008). A drawback of the knapsack problem is its inability to fully take into consideration the ranking of the projects according to their objective values because of the additivity in the objective function and the constraint equations in the integer programming model (Cook and Green 2000; Mavrotas et al. 2008). The real-world project selection problems aim at selecting a portfolio of projects that achieves organizational objectives without exceeding limited capital resources. Generally speaking, the environment is assumed to be uncertain and the project is assumed to possess imprecise or ambiguous data. The hybrid frameworks composed of DEA and the knapsack formulation are most suitable for solving these optimization problems (Chang and Lee 2012).

The project portfolio selection and capital budgeting problems are often formulated as knapsack problems. Khalili-Damghani et al. (2011) proposed a modular decision support system for optimum investment selection in the presence of uncertainty. They used a combination of fuzzy mathematical programming and a fuzzy rule based system to solve the investment problem based on the knapsack problem. Khalili-Damghani et al. (2012) proposed an integrated multi-objective framework for solving multi-period project selection problems. The proposed multi-objective model was a variant of the multi-objective multi-dimensional multi-period knapsack problem. Khalili-Damghani and Sadi-Nezhad (2013) proposed a decision support system for solving fuzzy multi-objective multi-period sustainable project selection problems. The proposed multi-objective model which was a variant of the multi-objective multi-dimensional multi-period knapsack problem was solved by a new fuzzy goal programming method. More recently, Khalili-Damghani et al. (2013c) compared the performance of a partial bound enumeration algorithm and an efficient variant of the epsilon-constraint method which was used to solve several benchmark instances of the fuzzy multi-dimensional multiple-choice knapsack problems. They compared the performance of the two methods statistically with respect to a set of simulated benchmark cases using different diversity and accuracy metrics.

In this paper, the knapsack formulation, an Efficient Epsilon-Constraint (EEC) method, and a customized multi-objective evolutionary algorithm are integrated into a structured but yet flexible model for solving the project selection problems in fuzzy environments. The proposed integrated model is not subject to the abovementioned drawbacks of DEA and the knapsack formulation, and also selects a portfolio of projects with fuzzy input and output data with respect to organizational objectives and resource constraints. Moreover, the current MMKPs in the literature are formulated to fill the knapsacks by picking exactly one item from each group. However, this assumption may be difficult to quantify in real world problems where the membership of an item to a group is not precise. In the proposed model we represent the membership function with fuzzy numbers rather than precise numbers.

The remainder of the paper is organized as follows: A review of the relevant literature on knapsack problems and DEA models is presented in Sect. 2. The details of the proposed framework including the problem definition, the solution procedures, and the pruning method are presented in Sect. 3. A real-world case study is presented in Sect. 4 to demonstrate the applicability of the proposed framework and exhibit the efficacy of the procedures and algorithms. The results of the experiment are presented in Sect. 5. Finally, we end the paper with conclusions and future research directions in Sect. 6.

2 Literature review

A multi-objective decision making model is composed of a vector of decision variables, objective functions, and constraints. The Decision Makers (DMs) are expected to consider the available resources and constraints while optimizing multiple objective functions. Let us consider the following definitions for multi-objective decision making problems (Hwang and Masud 1979):

Definition 1.1 x^* is said to be a *complete optimal solution*, if and only if there exists $x^* \in X$ such that $f_i(x^*) \leq f_i(x)$, $i = 1, \ldots, k$, for all $x \in X$. Also, the *ideal solution*, *superior solution*, or *utopia point* are equivalent terms indicating a complete optimal solution. In general, such a complete optimal solution that simultaneously minimizes all of the objective functions does not always exist when the objective functions conflict with each other.

Definition 1.2 x^* is said to be a *Pareto optimal solution*, if and only if there does not exist another $x \in X$ such that $f_i(x) \leq f_i(x^*)$ for all i and $f_j(x) < f_j(x^*)$ for at least one j. The *Pareto optimal solution* is also named differently by different disciplines: *non-dominated solution*, *non-inferior solution*, *efficient solution*, and *non-dominate solution*.

Definition 1.3 x^* is said to be a *weak Pareto optimal solution*, if and only if there does not exist another $x \in X$ such that $f_i(x) \leq f_i(x^*)$, $i = 1, \ldots, k$.

Let X^{CO}, X^P, and X^{WP} denote complete optimal, Pareto optimal, and weak Pareto optimal solution sets, respectively. We can easily obtain $X^{CO} \subseteq X^P \subseteq X^{WP}$ from the above definitions.

Multi-objective decision making problems often do not have a unique solution and therefore the DMs are required to choose a solution from the set of efficient solutions (Hwang and Masud 1979). The process may also require the articulation of the DMs' preferences. Depending upon the kind of preference information needed, Hwang and Masud (1979) have classified the multi-objective decision making techniques into the following four classes: (a) no articulation of preference information; (b) a priori articulation of preference information; (c) progressive articulation of preference information or interactive methods; and (d) a posterior articulation of preference information or non-dominated solution generation methods. The EEC method is a non-dominated solutions generation method proposed by Chankong and Haimes (1983).

2.1 Epsilon-constraint method

A special kind of multi-objective decision making problems, including linear programming, produces the entire efficient set. These methods can provide a representative subset of the Pareto set which in most cases is adequate. In the Epsilon-constrained method, the DM chooses one objective out of n to be optimized. For a multi-objective decision making problem with minimization objective functions, the remaining objectives are constraints as they are assumed to be less than or equal to some given target values. In mathematical terms, by selecting $f_j(x)$, $j \in \{1, \dots, k\}$, as the objective function to be optimized, we have the following problem $P(\varepsilon_j)$, $j \in \{1, \dots, k\}$:

$$\min\bigl\{f_j(x),\ j \in \{1, \dots, k\};\ f_i(x) \le \varepsilon_i,\ \forall i \in \{1, \dots, k\},\ i \neq j; x \in S\bigr\}$$

where, S is the feasible solution space.

One advantage of the Epsilon-constraint method is that it is able to generate efficient points in a non-convex Pareto curve. Therefore, the DM can vary the upper bounds ε_i to obtain weak Pareto optima. Clearly, this is also a drawback of this method since the DM has to choose appropriate upper bounds for the ε_i values. Moreover, the method is not particularly efficient if the number of objective functions increase. Several studies have focused on improving the Epsilon-constraint method (Mavrotas 2009). The traditional Epsilon-constraint methods are designed to find the efficient solutions to the problems by means of parametrical variation in the right-hand-side of the constraints. More formally, the implementation of the Epsilon-constraint methods requires attention to: (a) the calculation of the range of the objective functions over the efficient set; (b) the guarantee of efficiency of the obtained solution; and (c) the increased solution time for problems with several (more than two) objective functions. Mavrotas (2009) proposed a novel Epsilon-constraint method to address these issues.

2.2 Non-dominated sorting genetic algorithm (NSGA-II)

The Non-dominated Sorting Genetic Algorithm II (NSGA-II) is a well-known multi-objective evolutionary algorithm first introduced by Deb et al. (2002). The NSGA-II has been successfully applied in a wide range of engineering, management, and combinatorial optimization problems because it contains appropriate properties, such as elitism, fast non-dominated sorting and diversity maintenance along the Pareto-optimal front. A brief description of the mechanism of NSGA-II is provided here.

Each chromosome in a population is sorted based on non-domination into each front. The first front contains only non-dominant chromosomes among all chromosomes; the chromosomes in the second front are dominated by the chromosomes in the first front as this pattern is repeated. Each chromosome in each front is assigned a ranking in accordance to the front where it belongs to. Chromosomes in the first front are given a fitness value of 1, the chromosomes in the second front are assigned a fitness value of 2, and so forth.

A measure, called crowding distance, is calculated for each individual in the NSGA-II population. The crowding distance is a measure of how close an individual is to its neighbors. A large average crowding distance will result in an increased diversity in the population. Parents are selected from the population by using a binary tournament selection procedure based on their rank and crowding distance. The fitness of an individual, which belongs to a given rank, is less than others in the same rank with a greater crowding distance. The selected population generates offspring through crossover and mutation operators. All individuals, including the current population and the offsprings, are sorted again based on non-domination. Only the best N individuals are selected, where N is the population size. The selection is based on the rank and crowding distance on the last front. The details of the NSGA-II can be found in Deb et al. (2002).

NSGA-II is one of the most successful evolutionary computation algorithms. There are several successful applications of the NSGA-II algorithm in recent years. Khalili-Damghani et al. (2013b) proposed a hybrid fuzzy rule-based multi-criteria framework for sustainable project portfolio selection. They used the NSGA-II algorithm to generate non-dominated sets of fuzzy rules for an interpretable-accurate fuzzy rule-based system with three conflicting objectives. Khalili-Damghani et al. (2013a) proposed a new multi-objective particle swarm optimization method for solving the reliability redundancy allocation problems. They used the NSGA-II algorithm to assess the performance of a new multi-objective particle swarm optimization model based on benchmark cases.

2.3 Brief review of the DEA literature

The DEA is a mathematical programming approach that was initially developed as an efficiency analysis tool by Charnes et al. (1978). DEA uses the production metaphor of an entity, referred to as a Decision Making Unit (DMU), using inputs to produce outputs. DEA is able to compare DMUs using multiple criteria with different units of measurement while utilizing all inputs and outputs simultaneously. DEA is capable of: (1) identifying the best alternative; (2) ranking the alternatives; or (3) establishing a shortlist of the better alternatives for detailed review (Cook and Green 2000).

DEA has been used in solving project selection problems. Sowlati et al. (2005) presented a new model within the DEA for prioritizing information system projects. They used the inputs and outputs of the model as the criteria that judge the importance of the projects. A set of sample projects was created for which the criteria and priority scores were assigned by the DMs who compared each project to the set of defined projects. The new model was tested on a real case of prioritizing information system projects at a large financial institution. Their model provided a fair and equitable ranking and a new project could be prioritized at any time without affecting the priority of already assessed projects.

Vitner et al. (2006) investigated the possibility of using the DEA method for evaluating the performances of projects in a multi-project environment where each project was a one-time non-recurring event. Projects were evaluated by the earned value management system and multidimensional control system methods. They also proposed a method to reduce the number of inputs and outputs because it was usually necessary for the number of inputs and outputs not to exceed the number of projects.

Eilat et al. (2006) proposed a methodology for the construction and analysis of efficient, effective and balanced portfolios of projects with interactions. Their methodology was based on an extended DEA model that quantified some of the qualitative concepts embedded in the balanced scorecard approach. Their methodology included: a resource allocation scheme, an evaluation of individual projects, a screening of projects based on their relative values, and a construction and evaluation of portfolios. Their model evaluated individual projects and alternative project portfolios. They applied a branch-and-bound algorithm and an accumulation function that accounted for possible interactions among projects to generate portfolio alternatives. Eilat et al. (2008) developed a multi-criteria approach for evaluating research and development projects in different stages of their life cycle. They integrated the balanced scorecard with the DEA model through a hierarchical structure of constraints and proposed an extended DEA model. They illustrated the applicability of their approach with a case study involving an industrial research laboratory that selected dozens of projects every year.

Several authors have proposed integrated knapsack and DEA frameworks in the literature. Cook and Green (2000) used DEA and knapsack frameworks in a model to decide the project ranking in research and development selections. They select the optimal bundle of projects which possesses the best efficiency under resource constraints by treating each subset of the projects that could feasibly be selected within the resource constraints as a single, composite project. These composite projects are then evaluated by DEA against a 'production technology' defined by the available projects. They combine the evaluation and selection processes in a single model by placing the DEA model within a mixed-binary linear programming framework.

Mavrotas et al. (2008) proposed a portfolio optimization method for project selection where the final selection was guided by two aspects: satisfaction of certain logical constraints and assurance that the individual evaluation of the projects was valid to the maximum degree. Their unique approach compared the project portfolios without any special concerns on validating the project rankings. The entire process was implemented in two phases. The projects were first ranked through a multi-criteria approach. The obtained project rankings were then used in an integer programming module to effectively drive the final selection that satisfied the logical constraints. The innovative aspect of their approach was the way it overcame the well-known bias towards low cost projects which was caused by the knapsack formulation commonly used in the integer programming phase. Their proposed method improved the compatibility between the final selection of projects obtained from the integer programming model and the ranking obtained from the multi-criteria approach.

One limitation of the conventional DEA methods is the need for accurate measurement of the input and the output data. However, due to imprecise or ambiguous information over the project planning horizon, the model inputs for the project management decisions are often imprecise in practice (Xu et al. 2012). Imprecise evaluations may be the result of unquantifiable, incomplete and non-obtainable information. Numerous fuzzy methods have been proposed to deal with this impreciseness and ambiguity in DEA since the original study by Sengupta (1992). In general, fuzzy DEA methods can be classified into four categories, namely, the tolerance approach (Sengupta 1992), the α-level based approach (Kao and Liu 2003; Saati et al. 2002; Hatami-Marbini et al. 2010), the fuzzy ranking approach (Guo and Tanaka 2001) and the possibility approach (Lertworasirikul et al. 2003). An exhaustive review and taxonomy of various fuzzy DEA models can be found in Hatami-Marbini et al. (2011).

Fuzzy sets have also been used in the knapsack formulations. The three variables of knapsack capacity, item weights and profits in the conventional 0–1 knapsack formulations are assumed to be integers. However, many real-world knapsack problems involve imprecise data that could not be reasonably set to crisp integers. In the project selection problem, resource consumption and utility achievement levels of each project are often estimates. In addition, the quantities of the constraint resources are often dynamic. These conditions drive the need to introduce fuzzy set theory into knapsack models.

The literature consists of a few fuzzy DEA and knapsack formulations. Bas (2011) defuzzified a multidimensional 0–1 knapsack problem with fuzzy parameters using triangular norm fuzzy relations and developed surrogate relaxation models of the defuzzified problems. He then proposed a methodology for multi-attribute project portfolio selection and considered the optimal solutions from the original defuzzified problems as well as the near-optimal solutions from the surrogate relaxation models. He used a simple and effective fuzzy simple additive weighting method to manage the aggregation of evaluation results. Chang and Lee (2012) discussed the specific problem of selecting a portfolio of projects that achieved organizational objectives without exceeding limited capital resources in a fuzzy environment. They proposed an integrated DEA-knapsack formulation with fuzzy data and applied their model to a real-world problem in the engineering-procurement-construction industry. They applied three constraint handling techniques to transform their constrained optimization problem into an unconstrained problem, and for the first time used an artificial bee colony algorithm for the solutions.

3 The proposed framework

The hybrid framework proposed in this study is depicted in Fig. 1. The overall framework is comprised of three distinct but inter-related phases including: problem definition, solution procedures and pruning procedure.

As shown in Fig. 1, the first phase consists of modeling the new fuzzy knapsack problem and defining the new fuzzy project selection problem. The second phase consists of two solution procedures proposed to solve the multi-period project selection problem. The first procedure is an EEC method and the second procedure is a multi-objective evolutionary algorithm. A DEA model is used in the last phase of the proposed framework to prune the solutions found in the previous phase.

3.1 Phase I: Problem definition

In this phase, a new multi-objective knapsack problem and a multi-period project selection problem are defined as follows:

3.1.1 New fuzzy multidimensional multiple-choice knapsack model

We use the notations presented in Table 1 to model the new fuzzy knapsack problem.

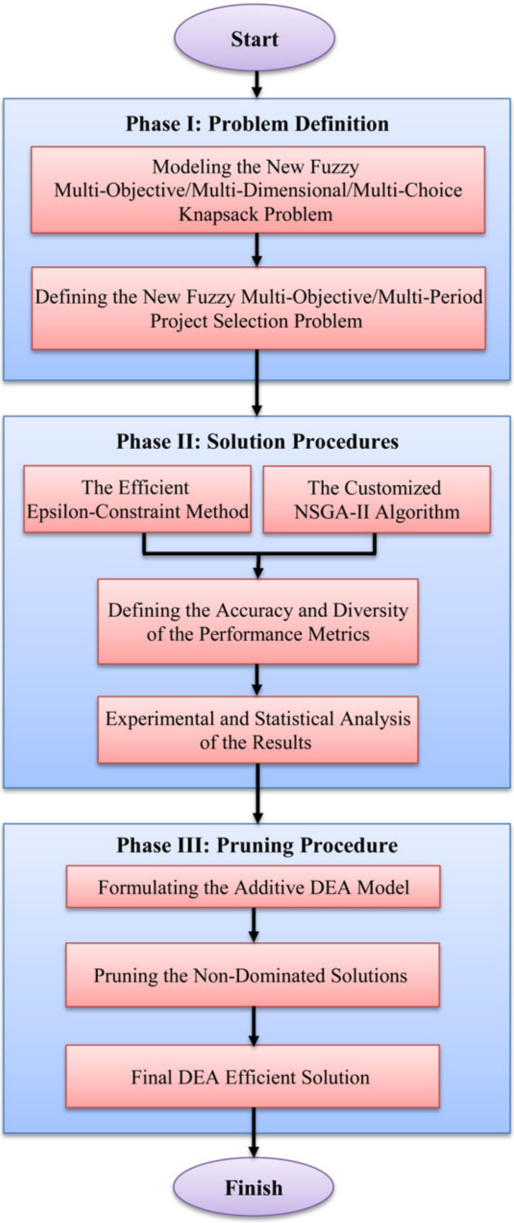
Let us also denote G_j, $j = 1, \dots, k$; to represent the set of objects with membership in group j. If each object i belongs to just one group, where μ_{ij} is the membership value of object i in group j, we have:

$$\mu_{ij} = \begin{cases} 1 & \text{if } i \in G_j, \\ 0 & \text{if } i \notin G_j, \end{cases} \qquad i = 1, \dots, n;\ j = 1, \dots, k \tag{1}$$

Then, the MMKP can be re-formulated as follows:

$$\begin{aligned} \max \sum_{i=1}^{n} p_i^l x_i, \qquad & l = 1, \dots, o \\ \sum_{i=1}^{n} w_i^d x_i \le W_d, \quad & d = 1, \dots, m \\ \sum_{i=1}^{n} \mu_{ij}.x_i = 1, \qquad & j = 1, \dots, k \\ x_i \in \{0, 1\}, \qquad & i = 1, \dots, n \end{aligned} \tag{2}$$

Fig. 1 A schematic view of the proposed hybrid framework



Most often in real-world problems, the assumption that each object belongs to one group may not be realistic since DMs cannot specify crisp boundaries for the groups. As a result an object could end up belonging to multiple groups with different membership degrees. Consequently, Eq. (1) may have multiple possibilities in real-world problems. Modeling this phenomenon can be useful in analyzing real-world problems.

Moreover, solving a fuzzy knapsack model may result in some undesirable cases wherein the sum of the membership degrees of the selected objects in the group is close to 1 but

Table 1 The parameters used in the proposed fuzzy MMKP

Sets		
G_j	Set of items in group j	
G_j^{α}	Set of objects with sufficient relations in group j	
Indices		
i	Index of items (objects)	$i = 1, 2, \ldots, n$
j	Index of groups	$j = 1, 2, \ldots, k$
l	Index of objective functions	$l = 1, 2, \ldots, o$
d	Index of dimensions	$d = 1, 2, \ldots, m$
Parameters		
n	Number of items (objects)	
k	Number of groups	
o	Number of objective functions	
m	Number of dimensions	
P_i^l	The profit of item i in objective l	$i = 1, 2, \ldots, n$; $l = 1, 2, \ldots, o$
w_i^d	Weight of object i in dimension d	$i = 1, 2, \ldots, n$; $d = 1, 2, \ldots, m$
W_d	Total capacity of dimension d of knapsack	$d = 1, 2, \ldots, m$
μ_{ij}	Membership value of object i in group j	$i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, k$
$\cong$	The sign of soft (fuzzy) equal	
Decision variables		
X_i	Binary decision variable: Equals to 1 if item i is selected; and 0 otherwise	$i = 1, 2, \ldots, n$

the membership degree of each selected object in the group is low. In order to avoid these undesirable circumstances, the selected objects in a group should have sufficient relations. An object has sufficient relations in a group if its membership degree in that group is equal to or greater than a pre-defined parameter. In other words, the final solution must include at least one object with sufficient relations. Therefore, the new fuzzy version of Model (2) can be formulated as follows:

$$\begin{aligned}
\max \sum_{i=1}^{n} p_i^l x_i, \qquad & l = 1, \ldots, o \\
\sum_{i=1}^{n} w_i^d x_i \le W_d, \quad & d = 1, \ldots, m \\
\sum_{i=1}^{n} \mu_{ij}.x_i \cong 1, \quad & j = 1, \ldots, k \\
\sum_{i \in G_j^{\alpha}} x_i \ge 1, \quad & j = 1, \ldots, k \\
x_i \in \{0, 1\}, \quad & i = 1, \ldots, n
\end{aligned} \tag{3}$$

where, G_j^{α} is the set of objects with sufficient relation in group j and is represented as:

$$G_j^{\alpha} = \{ i \mid \mu_{ij} \ge \alpha \}, \quad j = 1, \ldots, k \tag{4}$$

The sign $\cong$ represents the fuzzy membership degree of the items in each group (the associated constraint is written for each group). Since objects in the classical multidimensional multiple-choice 0–1 knapsack problems belong to one group, the number of selected objects in the final solution is equal to the number of groups (i.e. k). However, in Model (3),

Table 2 The notations used in the proposed fuzzy MMKP

Indices		
j	Index of groups of projects	$j = 1, \dots, k$
i	Index of projects	$i = 1, \dots, n$
o	Types of resources	$o = 1, \dots, z$
t	The planning horizon	$t = 1, \dots, T$
Parameters		
k	Number of groups of projects	
n	Number of projects	
z	Number of type of resources	
T	Number of planning horizons	
R_{ot}	Maximum available resource of type o in time period t	$o = 1, \dots, z; t = 1, \dots, T$
r_{oit}	Required resource of type o in project i in time period t	$o = 1, \dots, z; i = 1, \dots, n; t = 1, \dots, T$
B_{it}	Maximum available budget for project i in time period t	$i = 1, \dots, n; t = 1, \dots, T$
C_{ot}	Unit cost of resource of type o in time period t	$o = 1, \dots, z; t = 1, \dots, T$
p_{it}	Total net profit of project i in time period t	$i = 1, \dots, n; t = 1, \dots, T$
d_{it}	Duration of project i in time period t	$i = 1, \dots, n; t = 1, \dots, T$
μ_{ijt}	Membership degree of project i in group j in time period t	$i = 1, \dots, n; j = 1, \dots, k; t = 1, \dots, T$
Decision variables		
x_{ijt}	$\begin{cases} 1 & \text{if project } i \text{ is selected in group } j \text{ for investment in period } t \\ 0 & \text{otherwise} \end{cases}$	

the objects have fuzzy membership degrees and therefore the number of objects in the final solution may be not equal to k.

3.1.2 *Fuzzy multi-objective multi-period project selection problem*

A multi-period project selection problem is defined in order to demonstrate the real-world applications of the proposed fuzzy Model (3). Suppose an investor is facing several manufacturing and service investment opportunities in a finite planning horizon. The final output of these investment opportunities is a mixture of goods and service. In other words, an investment opportunity is not purely a production or service endeavor. In that case, the investor is interested in selecting a portfolio of investment opportunities with multiple objective functions such as profit, cost and time subject to a series of resource and planning horizon constraints. This problem can be modeled as a multi-period version of Model (3) considering additional constraints where the investment opportunities are goods and services. The fact that the outcome of an investment opportunity is a mixture of goods and services represents the fuzzy membership of each investment opportunity. The notations used to model the new multi-objective multi-period project selection problem are presented in Table 2.

Considering Model (3), we propose a mathematical programming model for fuzzy MMKP. The objective in the proposed model is to select a portfolio of projects with maximum total profit, minimum total cost and minimum total time. The constraints are available resources, the aforementioned rough project classification (i.e., goods and services) and the planning horizon constraints. Furthermore, additional constraints and variables may be supplied according to the nature of the project portfolio selection.

$$
\begin{aligned}
&\operatorname{Max} \Phi = \sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{k} p_{it} \times x_{ijt} \\
&\operatorname{Min} \omega = \sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{k} x_{ijt}\left(\sum_{o=1}^{z} r_{oit}.C_{ot}\right) \\
&\operatorname{Min} \theta = \sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{k} d_{jt} \times x_{ijt} \\
&\text{s.t.} \\
&\quad \sum_{i=1}^{n}\sum_{j=1}^{k} r_{oit} x_{ijt} \le R_{ot}, \qquad o = 1,2,\ldots,z;\ t = 1,2,\ldots,T, \\
&\quad \sum_{j=1}^{k}\left(\sum_{o=1}^{z} r_{oit}.C_{ot}\right) x_{ijt} \le B_{it}, \qquad i = 1,2,\ldots,n;\ t = 1,2,\ldots,T, \\
&\quad \sum_{i=1}^{n} \mu_{ijt} x_{ijt} \cong 1, \qquad t = 1,2,\ldots,T;\ j = 1,2,\ldots,k, \\
&\quad \sum_{i \in G_j^{\alpha}} x_{ijt} = 1, \qquad j = 1,\ldots,k;\ t = 1,2,\ldots,T, \\
&\quad \sum_{j=1}^{k} x_{ijt} = 1, \qquad i = 1,\ldots,n;\ t = 1,2,\ldots,T, \\
&\quad \sum_{t=1}^{T} x_{ijt} = 1, \qquad i = 1,\ldots,n;\ j = 1,2,\ldots,k, \\
&\quad \sum_{t=1}^{T} (t + d_{it}) \times x_{ijt} \le T + 1, \qquad i = 1,2,\ldots,n;\ j = 1,\ldots,k, \\
&\quad x_{ijt} \in \{0,1\}, \qquad i = 1,2,\ldots,n;\ j = 1,\ldots,k;\ t = 1,2,\ldots,T.
\end{aligned}
\tag{5}
$$

The first objective function maximizes the total profit of the selected projects in the service and production groups for all planning horizons. The second objective function minimizes the total cost of the selected projects in the service and production groups for all planning horizons. The third objective function minimizes the implementation period of the selected projects in the service and production groups for all planning horizons. Finding a set of non-dominated solutions for the problem is a challenging task due to the conflicting nature of these objective functions. The first set of constraints guarantees that the selected projects in the service and production groups meet the resource limitations in each planning horizon. The first set of constraints is written for all resource types in all planning horizons. The second set of constraints ensures that the cost of the selected portfolio is less than or equal to the available budget for each project in each time period of the planning horizon. The second set of constraints is written for all projects in all planning horizons. The third set of constraints guarantees that the sum of the fuzzy membership values of all selected projects in each group (i.e., goods or services) for each time period in the planning horizon is approximately equal to unit. On the other hand, a project can be selected from the product group or even the service group but the sum of the selected projects should be equal to unity for each time period. The fourth set of constraints guarantees that only the project with the minimum allowable number of relations in each group is selected in the final portfolio. The fourth set of constraints is written for all projects and all time periods of the planning horizons. The fifth set of constraints guarantees that each project is selected in one group in each time period of the planning horizon. The sixth set of constraints guarantees that each project in each group is selected in one time period during the planning horizon. The seventh set of constraints guarantees that each project will be selected and implemented within a restricted window of time in the planning horizon. The eighth set of constraints defines the binary state of the decision variables in the model.

Model (5) is a real-world and extended multi-period version of Model (3). In real-world multi-objective problems such as Model (5), generating a set of non-dominated solutions on the Pareto front is preferred.

3.2 Phase II: Solution procedures

In Phase II, an EEC method and a customized multi-objective evolutionary algorithm procedure are proposed to solve Model (5).

3.2.1 The efficient Epsilon-constraint method

We apply the EEC method proposed by Mavrotas (2009) to Model (5) and construct the following model:

$$
\begin{aligned}
&\text{Max}\ \Phi - \beta \times (S_2/r_2 + S_3/r_3)\\
&\text{s.t.}\\
&\quad \omega + S_2 = \varepsilon_2, \qquad \varepsilon_2 \in \left[\omega^-, \omega^+\right]\\
&\quad \theta + S_3 = \varepsilon_3, \qquad \varepsilon_3 \in \left[\theta^-, \theta^+\right]\\
&\quad X \in S
\end{aligned}
\tag{6}
$$

where, r_i, $i = 1, 2$, represent the range of objective i and are calculated using the lexicographic payoff table in Model (5) (i.e., using the ideal or "*the most desirable*" and nadir or "*the least desirable*" values of the cost and time objectives). The values of the slack variables in the second and the third objective functions are different because the scales of these two objective functions are dissimilar and should be put into a common scale. On the other hand, r_i, $i = 1, 2$, helps in resolving the problem of different scales in conflicting objective functions. The ω and θ have the same definitions as in Model (5). ω^+ and ω^- are the ideal and nadir values of the single cost optimization problem (5). Similarly, θ^+ and θ^- are the ideal and nadir values of the single quality optimization problem (5). $X \in S$ is the feasible region of Model (5) and β is a small positive value. The term $(S_2/r_2 + S_3/r_3)$ helps in generating possible strong non-dominated solutions on the Pareto front in Model (5). This term is subtracted from the objective function to act as a penalty for the objective function in Model (6). In other words, when the objective function in Model (6) is to be maximized, $(S_2/r_2 + S_3/r_3)$ should approach zero. r_i, $i = 1, 2$, are ranges of the objective functions and are always positive values. Therefore, the term $(S_2/r_2 + S_3/r_3)$ can be equal to zero if and only if the slack variables of the second and the third objective functions (i.e., S_2, and S_3) are equal to zero. In other words, the associated constraints in Model (6) have slack variables equal to zero (i.e., $\omega = \varepsilon_2$, and $\theta = \varepsilon_3$). Therefore, Model (6) can generate solutions in which one of the objective functions is maximized and two of the objective functions are on the boundary of the solution space. This is a promising condition for generating strong non-dominated solutions via changing ε_2, and ε_3.

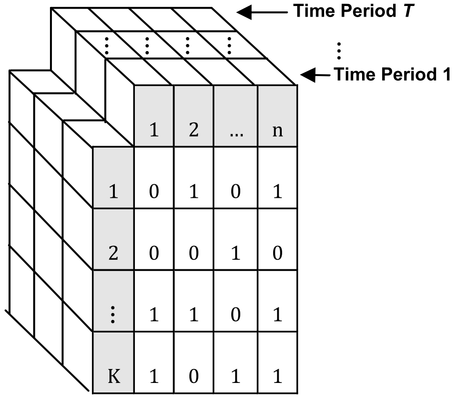


Fig. 2 The cubic binary-coded chromosome for the MMKP

3.2.2 *The customized multi-objective evolutionary algorithm procedure (NSGA-II algorithm)*

The reproduction operator mechanism of the NSGA-II, including crossover, mutation, fast-ranking, crowding distance and selection were discussed in the literature (Deb et al. 2002). Therefore, we only focus on the main properties of the modified NSGA-II here.

Considering k $(j = 1, 2, \ldots, k)$ different groups and n $(i = 1, 2, \ldots, n)$ projects, Fig. 2 presents a schematic 3-D view of a binary coded chromosome for representing a solution of Model (5). The dimensions of each chromosome are associated with k groups as length, n projects as width and t time periods as height, respectively.

In the modified NSGA-II, a heuristic method is used in the initial seeding to avoid a large number of infeasible solutions in the algorithm. The initial heuristic seeding procedure utilizes pseudo random methods to handle the third, fifth, and sixth sets of constraints in Model (5). In order to make a better exploration, half of the initial population is seeded using the heuristic procedure and the other half is seeded randomly.

A single point cut was selected as the crossover operator. According to the crossover probability, called P_c, two parents are selected and the crossover is accomplished. Figure 3 presents a schematic 2-D view of the proposed crossover. All binary values of a column in a given chromosome are changed based on a uniform mutation probability, called P_m. Figure 4 presents a schematic 2-D view of the mutation operator.

P_c and P_m are determined dynamically using linear equations which relate the crossover and mutation rate to the iteration number of the algorithm. P_c is set to a value close to 1 in the first iteration of the algorithm and then it is decreased as the iterations proceed. P_m is set to a small value in the first iteration and then it is increased as the iterations proceed. Dynamic parameter tuning causes better exploration and exploitation. The fitness function for a given rectangular binary chromosome is calculated based on the objective functions in Model (5).

A combination of strategies was also considered in the modified NSGA-II to handle the constraints in Model (5). The chromosomes, in which the multi-choice constraints are

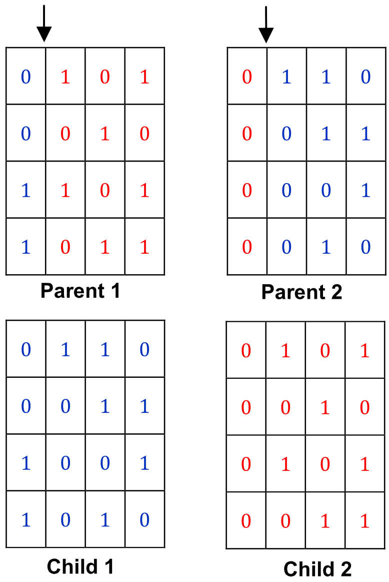


Fig. 3 The crossover for the MMKP

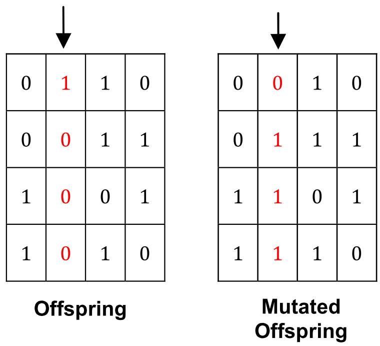


Fig. 4 The mutation for the MMKP

not met, are forbidden. On the other hand, the design where $\sum_{i=1}^{n}\sum_{j=1}^{k}\mu_{ijt}x_{ijt} \cong 1$, $t = 1, 2, \ldots, T$, are not met, are modified as follows:

$$\sum_{i=1}^{n}\sum_{j=1}^{k}\mu_{ijt}x_{ijt} = \begin{cases} 1, & \text{if } \sum_{i=1}^{n}\sum_{j=1}^{k}\mu_{ijt}x_{ijt} > 1,\ t = 1, 2, \ldots, T \\ \sum_{i=1}^{n}\sum_{j=1}^{k}\mu_{ijt}x_{ijt}, & \text{if } \sum_{i=1}^{n}\sum_{j=1}^{k}\mu_{ijt}x_{ijt} \cong 1,\ t = 1, 2, \ldots, T \\ 1, & \text{if } \sum_{i=1}^{n}\sum_{j=1}^{k}\mu_{ijt}x_{ijt} < 1,\ t = 1, 2, \ldots, T, \end{cases} \tag{7}$$

This condition should be held during the population initialization and evolutionary operators such as crossover and mutation.

A second problem arises when $\sum_{i=1}^{n}\sum_{j=1}^{k}\mu_{ijt}x_{ijt}$ is not in the allowed interval. In order to determine the object to be pruned from the violated group in the selected multi-period portfolio, a heuristic method is utilized in the modified NSGA-II. A cost-benefit ratio is calculated for each project in Model (5) where the profit of the project in a given time period is divided by the sum of the costs and resource requirements of the project in that time period as follows:

$$\text{Ratio}_{it} = \frac{P_{it}}{\sum_{o=1}^{z} r_{oit} \times C_{ot}}, \quad i = 1, 2, \ldots, n;\ t = 1, 2, \ldots, T. \tag{8}$$

where, r_{oit}, C_{ot}, and P_{it} have the same definition as Model (5). The projects with lower Ratio_{it} have larger priority to be pruned if required. So, the values of Ratio_{it} are listed in an ascending order in each time period and the redundant projects in that time period are pruned from the top of the sorted list.

The violation value of each chromosome (candidate solution) in the population is calculated, considering the dimension constraints of Model (5), as

$$o_{i1} = \text{Max}\left\{0, \sum_{j=1}^{k}\left(\sum_{o=1}^{z} r_{oit}.C_{ot}\right)x_{ijt} - B_{it}\right\} \tag{9}$$

$$o_{i2} = \text{Max}\left\{0, \sum_{i=1}^{n}\sum_{j=1}^{k} r_{oit}x_{ijt} - R_{ot}\right\} \tag{10}$$

where, o_{i1} is the violation value of the cost dimension for chromosome i in the population, and o_{i2} is the violation value of the resource dimension for chromosome i in the population.

A dynamic self-adaptive penalty strategy where the iteration of the algorithm, the situation of a single chromosome, and the overall situation of the population are concurrently considered is provided. In this strategy, the violation value of each chromosome (candidate solution) in the population is calculated using Eqs. (9)–(10). Then, the dynamic self-adaptive penalty functions for the profit, cost, and time objectives are calculated for each chromosome in the population as follows:

$$\text{Penalty} = \lambda\left[\left(\frac{o_{i1}}{o_{i1}^{\min}}\right)^{\alpha} \times t^{\beta}\right] + (1-\lambda)\left[\left(\frac{o_{i2}}{o_{i2}^{\min}}\right)^{\alpha} \times t^{\beta}\right] \tag{11}$$

$$\Phi_i' = \Phi_i - \text{Penalty} \tag{12}$$

$$\omega_i' = \omega_i + \text{Penalty} \tag{13}$$

$$\theta_i' = \theta_i + \text{Penalty} \tag{14}$$

where, Φ_i', ω_i', and θ_i' are the modified values for profit, cost, and time objectives for the violated chromosome i (solution) in the population. o_{i1} and o_{i2} have the same definition of (9) and (10), respectively. $o_{i1}^{\min} = \min_i\{\varepsilon + o_{i1}\}$ and $o_{i2}^{\min} = \min_i\{\varepsilon + o_{i2}\}$ are the minimum violation values for each chromosome in the population considering the cost and resource constraints in Model (5), respectively. The ε is a small positive value used to avoid division by zero, t is the iteration number, and α, β, and λ are the control parameters.

This type of dynamic self-adaptive penalty function provides a more efficient search method in the solution space by ensuring that the violation in the final steps of the algorithm will be penalized to a greater degree. Moreover, the penalty values are considered using a minimum violation of the population. The latter property will produce a penalty value so that the best chromosomes are considered in the population. Those chromosomes that do not meet the time window constraint in the planning horizon (i.e., $\sum_{t=1}^{T}(t + d_{it}).x_{ijt} \leq T + 1, i = 1, 2, \ldots, n;\ j = 1, \ldots, k$) are eliminated. This may occur during the population initialization and the evolutionary operators.

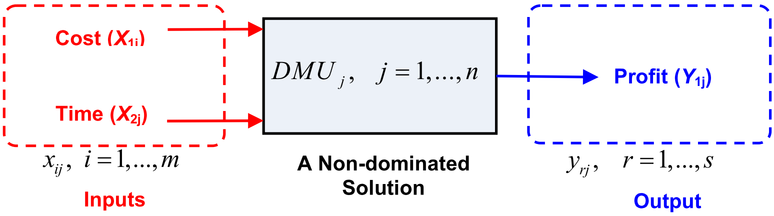


Fig. 5 The DMU representation of a non-dominated solution

3.3 Phase III: Pruning procedure

The selection of the final solution from the resultant non-dominated solutions is a difficult task. In the previous section, both methods were used to generate a set of non-dominated solutions. In circumstances where the DM has no knowledge of the properties of the objectives and the posterior articulation of the DM's preferences, a systematic approach is needed to facilitate the final selection process. A pruning method such as Data Envelopment Analysis (DEA) can be used to decrease the final number of non-dominated solutions. DEA is a methodology for evaluating and measuring the relative efficiencies of a set of decision making units (DMUs) that use multiple inputs to produce multiple outputs. The DEA method is based on the economic notion of Pareto optimality, which states that a DMU is considered to be inefficient if some other DMUs can produce at least the same amount of output with less of same input and not more of any other inputs. Otherwise, a DMU is considered to be Pareto efficient. Due to its solid mathematical basis and wide applications to real-world problems, much effort has been devoted to the DEA models since the pioneering work of Charnes et al. (1978). Each non-dominated solution of the modified NSGA-II algorithm is considered as a DMU with two inputs (i.e. cost and time) and one output (i.e. profit). Figure 5 presents the schematic view of a DMU.

Let us assume that there are a set of n DMUs (DMU_j, $j = 1, \ldots, n$) producing s outputs (y_{rj}, $r = 1, \ldots, s$) by consuming m inputs (x_{ij}, $i = 1, \ldots, m$). We use the following additive model proposed by Charnes et al. (1982) to check whether a given DMU (i.e., a non-dominated solution) is efficient.

$$\begin{aligned}
&\max \sum_{i=1}^{m} S_{ip}^{-} + \sum_{r=1}^{s} S_{rp}^{+} \\
&\text{s.t.} \\
&\quad \sum_{j=1}^{n} \lambda_j x_{ij} + S_{ip}^{-} = x_{ip}, \qquad i = 1, 2, \ldots, m \\
&\quad \sum_{j=1}^{n} \lambda_j y_{rj} - S_{rp}^{+} = y_{rp}, \qquad r = 1, 2, \ldots, s \\
&\quad \sum_{j=1}^{n} \lambda_j = 1 \\
&\quad \lambda_j, S_{ip}^{-}, S_{rp}^{+} \geq 0, \qquad j = 1, 2, \ldots, n, \ i = 1, 2, \ldots, m, \ r = 1, 2, \ldots, s
\end{aligned} \tag{15}$$

where, S_{ip}^{-} and S_{rp}^{+} are the input and output slacks and DMU_p is efficient under the additive Model (15) if and only if the optimal value of its objective function is zero.

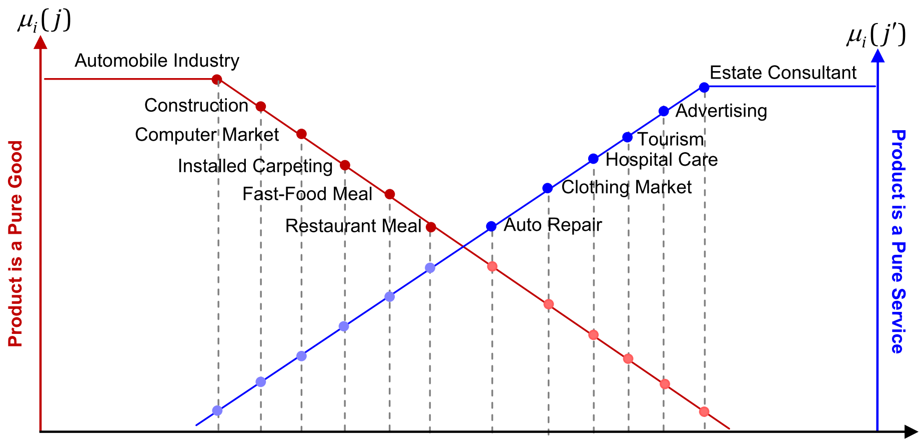


Fig. 6 A fuzzy spectrum of goods and services

4 The real-world case study

Peoples Bank,[1] the largest local bank in the State of East Virginia, is considering 12 different investment projects. The Chief Financial Officer (CFO) of the bank needs to select a non-dominated portfolio of projects by minimizing their costs and time and maximizing their profits over the next 10 years with respect to some known resource and cost constraints.

The projects considered by the Peoples Bank are investment opportunities with predetermined profit, cost and resource requirements. The bank considers two categories (i.e., goods and services) for grouping their projects; each project belongs to both groups. Fuzzy sets are used to represent the dependency of these projects on a spectrum of goods and services presented in Fig. 6. This property highlights the novelty of the proposed Models (5)–(6).

As shown in Fig. 6, a spectrum of goods and services is considered in our case study. For instance, both the fast-food meal and restaurant meal are categorized in similar groups. However, the amount of services provided with restaurant meals by far exceeds the amount of services provided with the fast-food meals. Similarly, consider goods and services provided to the customers in the automobile industry. The amount of goods provided in the automobile industry by far exceeds the amount of services. Consequently, the automobile industry belongs to both the service category and the goods category with different degrees of membership. In contrast, consider the case of the state consultant in which the amount of provided services is much higher than the amount of goods.

The 12 investment projects considered in this study all belong to both goods and services groups with different membership values. The available costs, profits and times of these 12 projects over the next 10 years are presented in Table 3.

Tables 4 and 5 present the available and required resources for each project during the next 10 years.

Table 6 presents the unit cost of the resources for each project during the next 10 years. Table 7 presents the membership value of each project in both goods and services groups during the next 10 years.

[1] The name is changed to protect the anonymity of the bank.

Table 3 The cost, profit, and time data for the investment projects

B_{jt} ($10 million)										
Project	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Automobile industry	6	8	8	9	6	8	6	9	6	7
Construction	2	4	2	2	4	3	4	3	3	4
Computer market	2	2	1	2	2	1	2	1	1	1
Installed carpeting	2	2	1	1	1	2	1	2	1	1
Fast-food meal	2	1	2	1	1	1	1	2	1	1
Restaurant meal	1	1	1	1	2	2	1	2	2	2
Clothing market	4	4	4	3	3	4	4	2	2	3
Tourism	13	14	14	11	11	12	14	11	10	10
Auto repair	4	3	3	5	2	3	2	4	4	4
Hospital care	2	1	3	2	1	1	1	3	2	2
Advertising	2	2	2	1	2	2	1	1	2	1
Estate consulting	2	3	2	3	3	2	3	3	2	2
P_{jt} ($10 million)										
Project	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Automobile industry	12	16	16	18	12	16	12	18	12	14
Construction	6	12	6	6	12	9	12	9	9	12
Computer market	10	10	5	10	10	5	10	5	5	5
Installed carpeting	4	4	2	2	2	4	2	4	2	2
Fast-food meal	12	6	12	6	6	6	6	12	6	6
Restaurant meal	4	4	4	4	8	8	4	8	8	8
Clothing market	20	20	20	15	15	20	20	10	10	15
Tourism	130	140	140	110	110	120	140	110	100	100
Auto repair	32	24	24	40	16	24	16	32	32	32
Hospital care	6	3	9	6	3	3	3	9	6	6
Advertising	30	30	30	15	30	30	15	15	30	15
Estate consulting	10	15	10	15	15	10	15	15	10	10
d_{jt} (years)										
Project	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Automobile industry	1	1	2	3	5	3	2	5	3	5
Construction	1	1	1	5	5	2	3	4	3	3
Computer market	0.5	0.5	0.3	0.5	0.5	0.4	0.4	0.2	0.1	0.4
Installed carpeting	0.8	0.4	0.6	0.3	0.5	0.8	0.7	0.4	0.2	0.7
Fast-food meal	0.3	0.7	0.3	0.5	0.8	0.8	1	0.2	0.8	0.2
Restaurant meal	0.2	0.5	1	0.9	1	0.7	1	0.5	0.8	0.9
Clothing market	0.5	0.3	0.8	0.3	0.8	0.6	0.9	1	0.2	0.2
Tourism	1	3	1	1	3	1	3	1	2	1
Auto repair	0.4	0.8	1.1	0.7	0.3	0.2	1	0.9	0.9	0.7
Hospital care	0.8	0.2	0.4	0.2	0.2	0.2	0.4	0.6	0.6	0.6
Advertising	2	1	2	3	1	2	3	2	2	2
Estate consulting	1	2	3	1	2	3	3	3	1	3

Table 4 The available resources for the investment projects

Available resources	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Human (man/time period) R_{1t}	921	841	561	722	824	673	912	737	947	758
Machine (quantity/time period) R_{2t}	55	688	726	938	783	541	992	729	998	563
Material (m^3/time period) R_{3t}	5248	5472	5032	5479	5120	5078	5008	5640	5462	5614

The data in Tables 3–7 were obtained from the historical data for similar projects, knowledge of the experts, research of the marketing unit, strategic planning documents and some forecasting data concerning the industry and customer behavior.

5 The experimental results

In this section, we present the results from the EEC method and the modified NSGA-II method for the 12 investment projects at the Peoples Bank.

5.1 Software-hardware implementation

The proposed EEC method was coded in LINGO 11.0 and VBA for MS-Excel 12.0. The modified NSGA-II algorithm was coded in VBA for MS-Excel 12.0. All codes ran on a PIV Pentium portable PC with MS-Windows XP Professional, 1 GB of RAM, and 2.0 GHz Core 2 Due CPU.

5.2 Results

Table 8 presents the Lexicographic payoff table for the different objectives. Considering the calculated values in Table 8, the range of each objective function can be easily found. Table 9 presents the parameters of the modified NSGA-II method. Ten non-dominated solutions were generated using both methods to demonstrate their impacts in the model. Table 10 presents the objective values of the generated non-dominated solutions with the EEC and the modified NSGA-II methods. The structure of the solution vector for both methods is presented in Table 11 and some of the generated non-dominated solutions for both methods are plotted in Fig. 7.

5.3 Parameter tuning and estimation of the Pareto front through the reference set

The results from the implementation of EEC and the modified NSGA-II methods revealed their effectiveness in generating non-dominated solutions. Next, statistical analysis was performed to find the best fitted parameters in the algorithms. In order to perform an acceptable statistical analysis, we first fitted the essential parameters given in Table 12 for both methods.

Next, we ran both Algorithms 10 times to re-generate the Pareto front dependently using their fitted parameters. The best solutions for these 10 runs, presented in Section (a) and Section (b) of Fig. 8, formed the re-generated Pareto fronts for each method. Figure 9 presents the objective function values for both approaches. Table 13 shows the upper-bound and the lower-bound of the generated non-dominated solutions.

Table 5 The required resources for the investment projects

r_{1ot} (human)

Project	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Automobile industry	67	67	97	77	57	50	69	77	97	88
Construction	71	70	67	69	75	98	96	80	54	57
Computer market	88	74	66	79	66	80	68	99	55	53
Installed carpeting	68	90	51	91	80	60	54	92	91	69
Fast-food meal	52	59	54	90	57	86	98	59	91	89
Restaurant meal	74	73	63	68	76	54	95	77	86	94
Clothing market	99	91	74	82	54	83	92	86	71	73
Tourism	79	80	96	81	54	80	86	56	77	84
Auto repair	50	61	75	57	98	84	62	67	60	66
Hospital care	77	75	62	81	65	70	71	92	52	95
Advertising	71	69	97	93	52	92	71	63	75	87
Estate consulting	76	76	75	55	86	97	94	57	78	54

r_{2ot} (machine)

Project	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Automobile industry	9	6	5	7	7	8	5	9	5	5
Construction	9	9	6	8	5	9	8	9	8	9
Computer market	6	8	8	5	5	9	7	5	9	8
Installed carpeting	8	5	9	5	6	6	8	8	6	5
Fast-food meal	8	5	9	7	7	6	8	8	7	5
Restaurant meal	5	9	5	5	7	7	7	6	7	8
Clothing market	7	6	6	5	5	5	5	6	6	9
Tourism	8	7	7	7	6	5	6	6	6	9
Auto repair	5	8	9	6	7	7	7	6	6	7
Hospital care	7	9	9	7	5	8	9	6	7	6
Advertising	5	8	5	8	7	6	9	8	5	7
Estate consulting	8	7	7	7	8	5	7	9	7	8

r_{3ot} (material)

Project	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Automobile industry	588	587	584	503	505	564	543	545	592	504
Construction	547	505	558	573	577	566	567	587	540	552
Computer market	588	510	549	556	531	541	558	540	511	576
Installed carpeting	541	548	560	594	584	537	571	578	514	547
Fast-food meal	523	582	517	533	552	588	585	579	566	552
Restaurant meal	593	508	594	526	557	584	571	597	562	545
Clothing market	545	505	538	579	583	598	591	583	535	509
Tourism	528	567	535	528	569	504	523	554	593	506
Auto repair	589	537	505	537	515	534	540	588	565	561
Hospital care	552	531	580	567	596	548	566	541	574	587
Advertising	545	554	593	532	557	550	510	569	505	578
Estate consulting	564	539	521	540	541	566	505	552	571	533

Table 6 The unit costs of the resources for the investment projects

Resources	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Labor $(C_{1t})^*$	24	11	17	24	26	17	23	17	19	12
Machine $(C_{2t})^{**}$	1130	1077	1093	1148	1182	1129	1080	1187	1083	1114
Material $(C_{3t})^{***}$	7	8	7	9	8	9	7	6	5	9

*The unit cost of labor has been calculated per hour ($10/hr)—Approximately four clusters are considered for labor types

**The unit cost of machines have been calculated per hour ($1000/hr)—Approximately four clusters are considered for machine types

***The unit cost of materials have been calculated per m^3 ($500/m^3)—Approximately four clusters are considered for material types

Table 7 The membership values of the investment groups

Investment category	G_1	G_2
Automobile industry	0.70	0.30
Construction	0.30	0.70
Computer market	0.70	0.30
Installed carpeting	0.00	1.00
Fast-food meal	0.90	0.10
Restaurant meal	0.70	0.30
Clothing market	0.30	0.70
Tourism	0.40	0.60
Auto repair	0.70	0.30
Hospital care	0.50	0.50
Advertising	0.90	0.10
Estate consulting	0.20	0.80

5.4 Accuracy measures

In lieu of the fact that the real Pareto front of a problem is not known in advance, a Reference Set (RS) could be used for comparison purposes. The RS only contains the non-dominated solutions for all runs for both methods. The comparison metrics for accuracy and diversity for both methods were calculated using the accuracy and diversity metrics proposed by Yu and Gen (2010). Figure 10 presents the RS results for the two methods. The RS contains 49 non-dominated solutions which were selected from the generated solutions for both methods in all runs.

Figure 11 presents a comparison between the EEC and the RS and between the modified NSGA-II and the RS. These figures show the parity of the EEC and the modified NSGA-II methods to the RS method in different dimensions.

In order to evaluate the performance of the EEC and the modified NSGA-II methods, the following accuracy and diversity metrics proposed by Yu and Gen (2010) were used in this case study.

- *Number of non-dominated solutions* (*NNS*): This metric represents the number of non-dominated solutions found by each method. A fast sorting procedure was used to deter-

Table 8 The ideal and nadir payoff values of Model (9)

Ideal values			
	Profit	Cost	Time
Profit	6.24E+8	229265.0	26.70
Cost	2.48E+08	82362	14.00
Time	1.14E+08	105242.0	4.000
Nadir values			
	Profit	Cost	Time
Profit	6.2E+7	105467.0	12.00
Cost	5.18E+08	275440	25.50
Time	4.58E+08	247483.0	47.30

Table 9 The modified NSGA-II and EEC parameters

Initializing the modified NSGA-II parameters		Initializing the EEC parameters	
Pop. no.	20	Project no.	12
Archive size	35	Time period no.	10
Maximum iteration no.	1000	Group no.	2
Mutation rate	Change linearly in [0.01, 0.05]	Objective functions no.	3
Cross rate	Change linearly in [0.9, 0.7]	Maximum cost allowed	247483
Alpha	2	Maximum profit allowed	6.24E+8
Beta	3	Minimum time allowed	4.0

mine whether a candidate solution for a given method is non-dominated or not. The higher this metric, the more the method has converged towards the real Pareto front.

- *Error ratio* (*ER*): ER measures the non-convergence of the methods towards the real Pareto front. The definition of the ER is given as:

$$\mathrm{ER} = \frac{\sum_{i=1}^{N} e_i}{N} \tag{16}$$

where N is the number of non-dominated solutions found, and

$$e_i = \begin{cases} 0 & \text{if the solution i belongs to the Pareto front} \\ 1 & \text{otherwise} \end{cases}$$

The closer this metric is to 1, the less the solution has converged toward RS.

- *Generational distance* (*GD*): This metric calculates the distance between RS and the solution set. The definition of this metric is given as:

$$\mathrm{GD} = \frac{\sum_{i=1}^{N} d_i}{N} \tag{17}$$

where, $d_i = \min_{p\in PF}\{\sqrt{\sum_{k=1}^{m}(z_k^i - z_k^p)^2}\}$ is the minimum Euclidean distance between solution i and RS where $|m|$ is the number of objective functions.

Table 10 The objective values of the sample solutions

	The Epsilon values									
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
EEC										
Φ	624000000	624000000	624000000	624000000	624000000	624000000	624000000	624000000	624000000	624000000
ω	228661	219608	222022	218642	227610	214152	215615	215615	239251	217411
θ	26.9	25.7	23	24.1	24.8	23.5	28	28	22.5	23.9
Modified NSGA-II										
Φ	539000000	529000000	526000000	519000000	512000000	505000000	502000000	499000000	497000000	496000000
ω	196747	199596	196993	191612	181418	185001	178767	187147	180772	164378
θ	23	19.7	21.4	25.5	22.5	20.5	22.3	18.4	21.2	23.9

Table 11 The solution structure

EEC										
Epsilon	Time periods									
	1	2	3	4	5	6	7	8	9	10
0	12	1, 2, 6, 11	2, 4, 10	4, 12	3, 7	1, 8, 9, 11	3	6, 8	5, 7, 9	5, 10
0.1	12	1, 2, 6, 11	2, 10	4, 12	3, 7	1, 8, 9, 11	3	6, 8	5, 7, 9	5, 10
0.2	12	1, 2, 11	2, 4	4, 12	3, 7	1, 8, 9, 11	3	6, 8	5, 7, 9	5, 10
0.3	12	1, 2, 11	2, 5	3, 4, 12	1, 7, 10	6, 8, 9, 11	3	8	5, 7, 9	6, 10
0.4	1, 2	1, 11, 12	2	12	3, 7, 10	8, 9, 11	3	8	5, 7, 9	5, 10
0.5	1, 7	1, 2, 11	2	12	5	8, 11, 12	3	5	5, 7, 9	9
0.6	1, 7	8, 9, 11	2	12	10	11, 12	7	8	5	9
0.7	11	11	2	12	9	8	7	12	8	9
0.8	3, 10	2, 4, 11	2, 5, 12	4, 7	3, 7	1, 8, 9, 11	1, 6	6, 8	9, 12	5, 10
0.9	3, 10, 12	2, 11, 12	2	4, 5	1, 7, 10	1, 8, 9, 11	3, 4, 6	8	7, 9	5, 6

Modified NSGA-II										
Solution	Time periods									
	1	2	3	4	5	6	7	8	9	10
1	4	2, 9, 11	1, 12	5, 12	3, 7	2, 9, 11	1, 4	6, 8	3, 8	5, 6
2	1, 3, 4	2, 9	3	5, 10	5, 11	8	1, 2	8, 9	7, 12	6, 7, 10
3	4	2, 9, 11	1, 12	5, 12	3, 7	2, 9, 11	4, 10	6, 8	3, 8	5, 6
4	1, 6	11	5, 7	4, 6	4, 8	1, 3, 12	2, 7, 11	8, 9	9, 12	3
5	1, 3	1, 8, 11	11, 12	5, 12	4, 7	3, 8	2, 10	7, 9	5	9
6	3, 6	1, 2	11, 12	3, 4, 5	1, 9	8	5	8, 9	7, 12	6, 7
7	2, 6	1, 9	11, 12	5, 10	1, 4	8	10, 11	8, 9	7, 12	3, 7
8	1, 7	11	6, 10	4, 6	4, 8	3, 5, 11	1, 10	8, 9	9, 12	3
9	4	3, 7, 11	1, 12	6, 9	10, 12	6, 8	1, 2	8, 9	5	5, 7
10	4	1, 9	11, 12	2, 5	1, 10	8	3	8, 9	7, 12	6, 7

Table 12 The fitted parameters for the EEC and modified NSGA-II methods

EEC		Modified NSGA-II	
Archive size	30	Chromosome no.	20
Upper bound of cost	277068	Archive size	30
Upper bound of time	47.8	Maximum iteration no.	1000
Lower bound of cost	0	Cross rate	Change linearly in [0.8, 0.7]
Lower bound of time	0	Mutation rate	Change linearly in [0.02, 0.04]
Step size of cost	0.1	Alpha	1
Step size of time	0.1	Beta	5

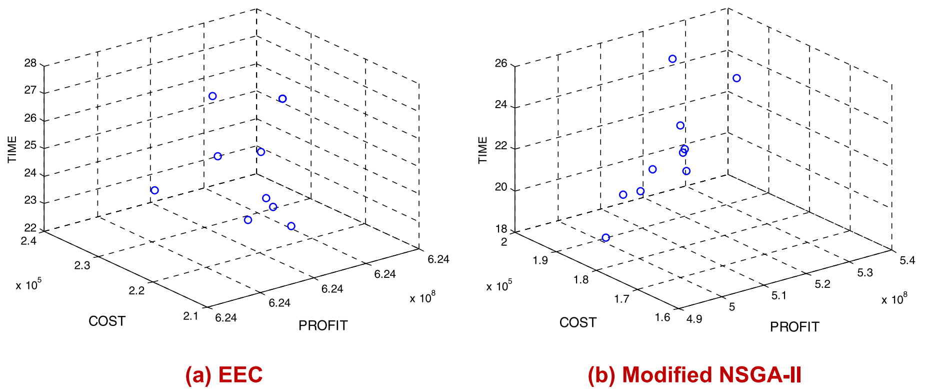


Fig. 7 The 3D-view of the non-dominated solutions for EEC and modified NSGA-II

Table 13 The upper-bound and the lower-bound values for the generated non-dominated solutions

EEC			
	Profit	Cost	Time
Upper bound	624000000	275440	28
Lower bound	234000000	82065	4.7
Modified NSGA-II			
	Profit	Cost	Time
Upper bound	550000000	208972	24.9
Lower bound	397000000	150096	17.1

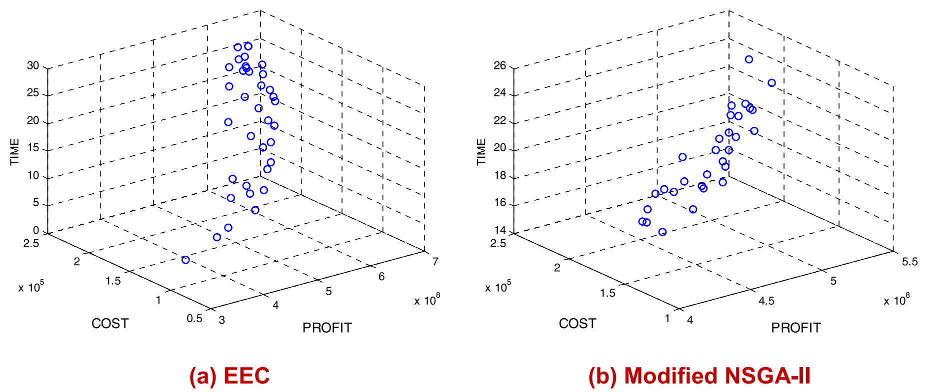
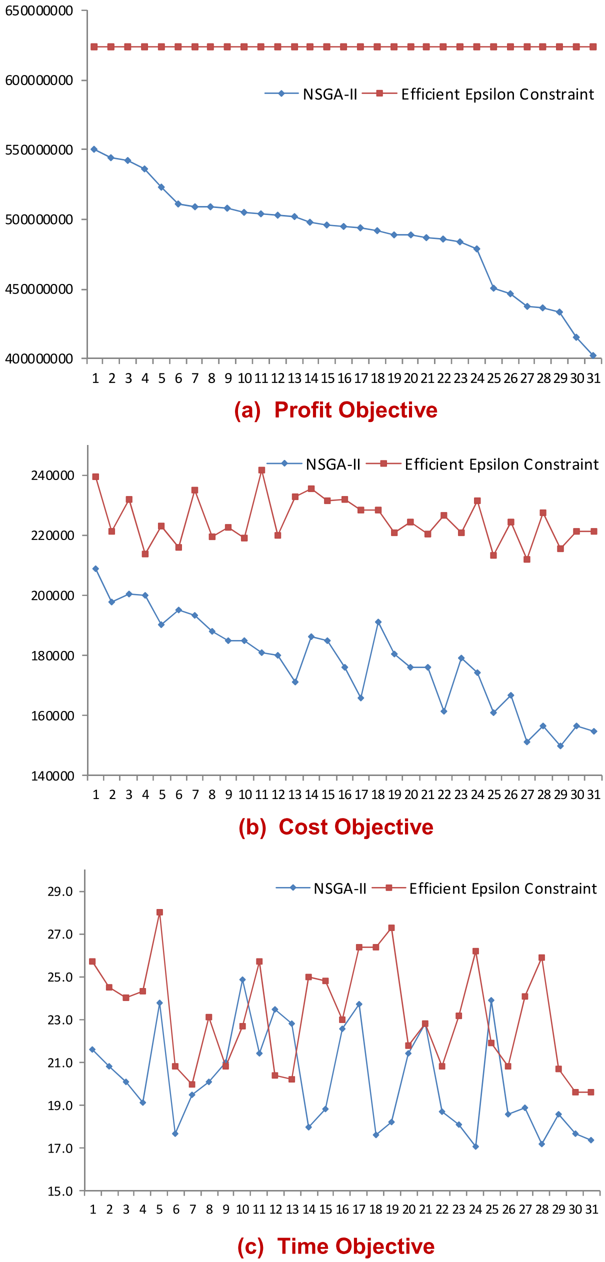


Fig. 8 The re-generated 3-D Pareto front for EEC and modified NSGA-II

5.5 Diversity measures

- *Spacing metric* (*SM*): SM measures the uniformity of the spread of the points in the solution set. In order to calculate SM, $\bar{d}$, the mean value of all the d_i, is calculated as

Fig. 9 The objective values of EEC and modified NSGA-II



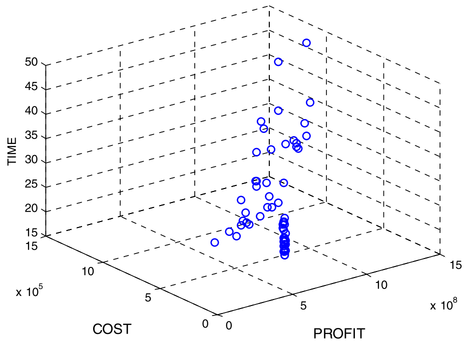
follows:

$$\bar{d} = \frac{\sum_{i=1}^{N} d_i}{N} \tag{18}$$

Next, SP, the standard deviation of the closest distances, is calculated as follows:

$$\mathrm{SP} = \sqrt{\left(\frac{\sum_{i=1}^{N} (\bar{d} - d_i)^2}{N-1}\right)} \tag{19}$$

Fig. 10 The reference set



- *Diversification metric* (*DM*): DM measures the spread of the solution set and is calculated as follows:

$$\mathrm{DM} = \left[\sum_{i=1}^{N} \max\bigl(\|x_i - y_i\|\bigr)\right]^{\frac{1}{2}} \tag{20}$$

where, $\|x_i - y_i\|$ is the Euclidean distance between the non-dominated solution x_i and the non-dominated solution y_i. The aforementioned metrics as well as the CPU run time were calculated for both methods in 10 different runs. The results are represented in Table 14 and Table 15.

As shown in Tables 14 and 15, the computation times for the modified NSGA-II method is larger than the computation times for the EEC method. The modified NSGA-II has more non-dominated solutions in comparison with the EEC method. However, the modified NSGA-II had a smaller error ratio in all runs representing the lower non-convergence of the NSGA-II method towards the RS. The modified NSGA-II method re-generated non-dominated solutions which were considerably closer to the RS than the EEC method. The modified NSGA-II method also re-generated non-dominated solutions with smaller average values for SM. This shows that the solutions generated by the modified NSGA-II method are more uniformly distributed throughout the RS in comparison with solutions generated by the EEC method. The average DM value for the modified NSGA-II method is much more promising in comparison with the average DM values for the EEC method. In other words, the modified NSGA-II method results in more scattered non-dominated solutions.

Although the conclusion that the non-dominated solutions produced by the modified NSGA-II method dominate the solutions produced by the EEC method sounds logical, statistical analysis is used next to further test this conclusion.

5.6 Statistical analysis and pruning the solutions

The Kolmogorov-Smirnov test was performed for different metrics to determine whether the results fit a normal distribution. The Kolmogorov-Smirnov test results for these metrics are presented in Fig. 12.

As shown in Fig. 12, there is not enough evidence to reject the null hypothesis of the Kolmogorov-Smirnov test that states the populations of these metrics follow normal distributions. Hence, considering these findings, we performed a parametric statistical test to check whether there is a significant difference between the means for different metrics. Analysis of Variance (ANOVA) was used and the following hypotheses of one-way ANOVA were formulated:

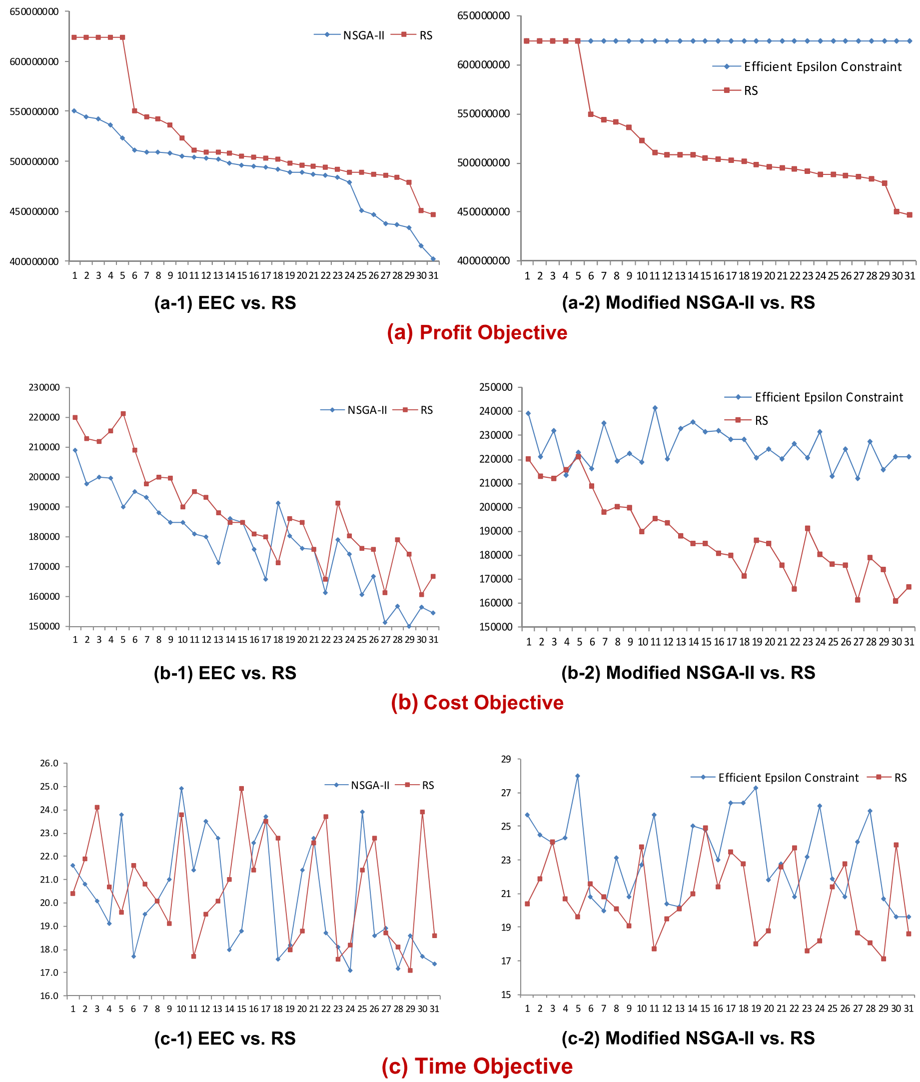


Fig. 11 The comparison between EEC and RS and between modified NSGA-II and RS

H_0: the means of population metrics are all equal
H_1: the means of population metrics are not all equal

The ANOVA results are presented in Table 16.

As shown in Fig. 16, there is not enough evidence to reject the null hypothesis. Therefore, the previous interpretations of Table 14 are statistically proven. Finally, we pruned the non-dominated solutions generated by the modified NSGA-II method using the additive DEA Model (15). Model (15) produced 10 non-dominated solutions for the modified NSGA-II method presented in Table 17 where 4 of the 10 solutions (DMU_4, DMU_6, DMU_8 and DMU_{10}) were determined as efficient. This pruning procedure was intended to facilitate the portfolio selection process.

Table 14 The computational results of the accuracy metrics for the EEC and modified NSGA-II methods

Run	NNS		ER		GD	
	Modified NSGA-II	EEC	Modified NSGA-II	EEC	Modified NSGA-II	EEC
1	28	6	0.09677	0.8064516	967.74	4165.718
2	29	6	0.06451	0.8064516	645.16	4165.718
3	26	6	0.16129	0.8064516	1612.90	4165.718
4	29	6	0.06451	0.8064516	645.16	4165.718
5	31	6	0	0.8064516	0.00	4165.718
6	23	6	0.25806	0.8064516	2580.65	4165.718
7	23	5	0.25806	0.8387097	2580.65	4493.5532
8	28	5	0.09677	0.8387097	967.74	4493.5532
9	28	6	0.09677	0.8064516	967.74	4165.718
10	31	6	0	0.8064516	0.00	4165.718
Average	27.6	5.8	0.10967	0.8129032	1096.77	4231.285
Standard deviation	2.83627	0.421637	0.09149	0.0136012	914.926	138.22749

Table 15 The computational results of the diversity metrics and time for the EEC and modified NSGA-II methods

Run	SM		DM		CPU time (s)	
	Modified NSGA-II	EEC	Modified NSGA-II	EEC	Modified NSGA-II	EEC
1	3071.01	5056.481	2.2E+12	136797.72	376.453	96.160409
2	2533.84	5056.481	2.2E+11	136797.72	383.766	90.664732
3	3976.60	5056.481	2.2E+10	136797.72	378.672	91.984035
4	2457.40	5056.481	2.2E+12	136797.72	364.828	92.130627
5	0	5056.481	2.2E+12	136797.72	363.703	96.560206
6	4465.15	5056.481	2.2E+11	136797.72	361.250	94.011733
7	4465.15	4662.9979	2.2E+12	136798.03	361.250	95.467984
8	3095.20	4662.9979	2.2E+13	136798.03	364.953	91.752884
9	3095.20	5056.481	2.2E+13	136798.03	373.953	95.651284
10	0	5056.481	2.2E+12	136798.03	379.359	93.429075
Average	2715.95	4977.7844	2.2E+12	136797.85	370.81875	93.781297
Standard deviation	1597.74	165.90704	1.0E+12	0.1579288	8.4836899	2.0998342

6 Conclusions and future research directions

In the classic 0–1 MMKP, each object belongs to one and only one group. However, this assumption is often not valid in the real-world applications as each object may belong to multiple groups with different degrees of membership. We have applied fuzzy set theory to model this problem as a new fuzzy MMKP. An EEC and a modified NSGA-II method were proposed and customized to solve the fuzzy MMKP.

The proposed EEC and modified NSGA-II methods were applied to a real-world project portfolio selection problem. Different sets of non-dominated solutions were generated us-

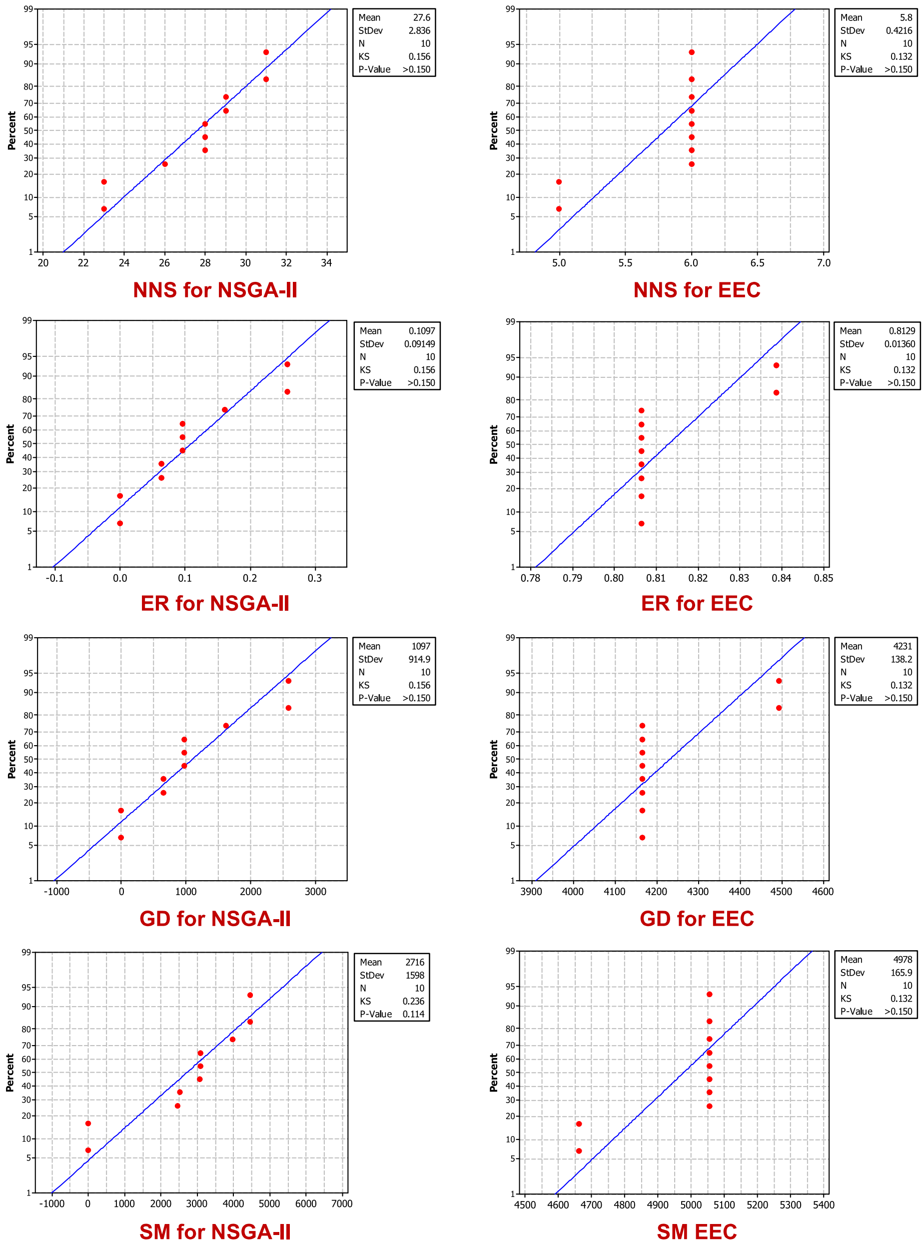


Fig. 12 The Kolmogorov-Smirnov normality test results for the performance metrics

ing the proposed EEC and the modified NSGA-II methods. The performance of the two proposed methods was compared with the results from the project portfolio selection problem. The accuracy and diversity metrics proposed by Yu and Gen (2010) were used in this case study to evaluate the performance of the EEC and the modified NSGA-II methods. Ac-

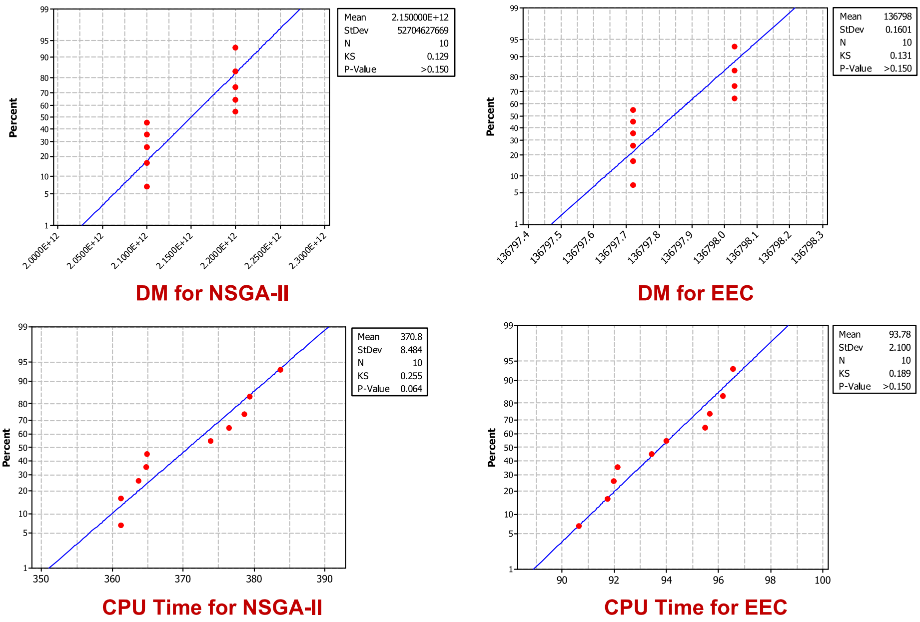


Fig. 12 (*Continued*)

cording to the accuracy and diversity metrics, the non-dominated solutions produced by the modified NSGA-II method dominate the solutions produced by the EEC method.

In order to further test this conclusion, ANOVA was used to test whether there is a significant difference between the means for different metrics. The results from the ANOVA test validated our earlier conclusion concerning the relative preference of the modified NSGA-II method over the EEC method. Finally, the non-dominated solutions generated with the modified NSGA-II method were pruned using an additive DEA model to ease the portfolio selection process. The additive DEA model found a set of efficient solutions with lower cost, shorter time and higher profit on the Pareto front. This pruning process provided the DMs with a smaller set of non-dominated solutions for further consideration and selection.

Risk is another paradigm in project management. Risk is naturally assumed to have multiple relations with profit, cost, and time in project management. The risk of a project can be modeled with utility theory and included in the model proposed in this study. The effect of time value of money and inflation on the optimal portfolio of projects can also be an interesting area for future research. The inclusion of risk, time value of money, and inflation in the proposed fuzzy MMKP could be an interesting path of further work.

The parameters of a project such as profit, cost, time, and resource requirements are usually assumed to be varied during the planning horizon. This can affect the tradeoff problem. Some kind of uncertainty modeling such as supplying interval data, fuzzy sets or even probabilistic programming can also be interesting for future research. Another possible area of extension is to use Markov decision process (Beltrami et al. 1985; Durinovic et al. 1986) for modeling the uncertainty of the project parameters.

Acknowledgements The authors would like to thank the anonymous reviewers and the editor for their insightful comments and suggestions.

Table 16 The ANOVA results for the comparison metrics

I. NNS

Source	Degree of freedom	Sum of square	Mean square	F	P-value
Factor	1	2376.2	2376.2	577.99	0
Error	18	74	4.11		
Total	19	2450.2			

$S = 2.028$, R-Sq = 96.98 %, R-Sq (adj) = 96.81 %

II. ER

Source	Degree of freedom	Sum of square	Mean square	F	P-value
Factor	1	2.4106	2.4106	712.8	0
Error	18	0.06087	0.00338		
Total	19	2.47148			

$S = 0.05815$, R-Sq = 97.54 %, R-Sq (adj) = 97.40 %

III. GD

Source	Degree of freedom	Sum of square	Mean square	F	P-value
Factor	1	48969072	48969072	116.02	0
Error	18	7597073	422060		
Total	19	56566145			

$S = 649.7$, R-Sq = 86.57 %, R-Sq (adj) = 85.82 %

IV. SM

Source	Degree of freedom	Sum of square	Mean square	F	P-value
Factor	1	19026979	19026979	33.3	0
Error	18	10283917	571329		
Total	19	29310897			

$S = 755.9$, R-Sq = 64.91 %, R-Sq (adj) = 62.97 %

V. DM

Source	Degree of freedom	Sum of square	Mean square	F	P-value
Factor	1	2.19E+25	2.19E+25	877804.9	0
Error	18	4.50E+20	2.50E+19		
Total	19	2.19E+25			

$S = 5000000000$, R-Sq = 100.00 %, R-Sq (adj) = 100.00 %

VI. CPU time

Source	Degree of freedom	Sum of square	Mean square	F	P-value
Factor	1	383748.8	383748.8	10046.7	0
Error	18	687.5	38.2		
Total	19	384436.3			

$S = 6.180$, R-Sq = 99.82 %, R-Sq (adj) = 99.81 %

Table 17 Pruning the non-dominated solutions with the additive Model (24)

DMU	Inputs		Output	S_{ip}^{-}		S_{rp}^{+}	Objective value
	Cost	Time	Profit	S_{1p}^{-}	S_{2p}^{-}	S_{1p}^{+}	
DMU_1	196747.00	23.00	539000000.00	196724.50	184977.10	43000000.00	43381701.60
DMU_2	199596.00	19.70	529000000.00	0.00	0.00	27909090.91	27909090.91
DMU_3	196993.00	21.40	526000000.00	199573.50	178743.10	30000000.00	30378316.60
DMU_4*	191612.00	25.50	519000000.00	0.00	0.00	0.00	0.00
DMU_5	181418.00	22.50	512000000.00	196970.50	187123.10	16000000.00	16384093.60
DMU_6*	185001.00	20.50	505000000.00	0.00	0.00	0.00	0.00
DMU_7	178767.00	22.30	502000000.00	191589.50	180748.10	6000000.00	6372337.60
DMU_8*	187147.00	18.40	499000000.00	0.00	0.00	0.00	0.00
DMU_9	180772.00	21.20	497000000.00	181395.50	164354.10	1000000.00	1345749.60
DMU_{10}*	164378.00	23.90	496000000.00	0.00	0.00	0.00	0.00

*Efficient solutions

References

Aghezzaf, B., & Naimi, M. (2009). The two-stage recombination operator and its application to the multiobjective 0/1 knapsack problem: a comparative study. *Computers & Operations Research*, *36*, 3247–3262.

Alves, M. J., & Almeida, M. (2007). MOTGA: a multiobjective Tchebycheff based genetic algorithm for the multidimensional knapsack problem. *Computers & Operations Research*, *34*, 3458–3470.

Bas, E. (2011). A capital budgeting problem for preventing workplace mobbing by using analytic hierarchy process and fuzzy 0–1 bidimensional knapsack model. *Expert Systems with Applications*, *38*(10), 12415–12422.

Beltrami, E., Katehakis, M. N., & Durinovic, S. (1985). Multiobjective Markov decision process in urban modeling. *Mathematical Modelling*, *6*(4), 333–338.

Camargo, V. C. B., Mattiolli, L., & Toledo, F. M. B. (2012). A knapsack problem as a tool to solve the production planning problem in small foundries. *Computers & Operations Research*, *39*(1), 86–92.

Chang, P. T., & Lee, J. H. (2012). A fuzzy DEA and knapsack formulation integrated model for project selection. *Computers & Operations Research*, *39*(1), 112–125.

Chankong, V., & Haimes, Y. (1983). *Multi-objective decision making theory and methodology*. New York: Elsevier.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429–444.

Charnes, A., Cooper, W. W., Seiford, L. M., & Stutz, J. (1982). A multiplicative model for efficiency analysis. *Socio-Economic Planning Sciences*, *16*, 223–224.

Cook, W. D., & Green, R. H. (2000). Project prioritization—a resource constrained data envelopment analysis approach. *Socio-Economic Planning Sciences*, *34*(2), 85–99.

Crévits, I., Hanafi, S., Mansi, R., & Wilbaut, C. (2012). Iterative semi-continuous relaxation heuristics for the multiple-choice multidimensional knapsack problem. *Computers & Operations Research*, *39*(1), 32–41.

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast an elitist multi-objective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, *6*, 182–197.

Durinovic, S., Lee, H. M., Katehakis, M. N., & Filar, J. A. (1986). Multiobjective Markov decision process with average reward criterion. *Large Scale Systems*, *10*(3), 215–226.

Eilat, H., Golany, B., & Shtub, A. (2006). Constructing and evaluating balanced portfolios of R&D projects with interactions: a DEA based methodology. *European Journal of Operational Research*, *172*(3), 1018–1039.

Eilat, H., Golany, B., & Shtub, A. (2008). R&D project evaluation: an integrated DEA and balanced scorecard approach. *Omega*, *36*(5), 895–912.

El-Mashaleh, M. S., Rababeh, S. M., & Hyari, K. H. (2010). Utilizing data envelopment analysis to benchmark safety performance of construction contractors. *International Journal of Project Management*, *28*(1), 61–67.

Farris, J. A., Groesbeck, R. L., Aken, E. M. V., & Letens, G. (2006). Evaluating the relative performance of engineering design projects: a case study using data envelopment analysis. *IEEE Transactions on Engineering Management*, *53*(3), 471–482.

Florios, K., Mavrotas, G., & Diakoulaki, D. (2010). Solving multiobjective, multiconstraint knapsack problems using mathematical programming and evolutionary algorithms. *European Journal of Operational Research*, *203*, 14–21.

Ghapanchi, A. H., Tavana, M., Khakbaz, M. H., & Low, G. (2012). A methodology for selecting portfolios of projects with interactions and under uncertainty. *International Journal of Project Management*. doi:10.1016/j.ijproman.2012.01.012.

Ghasemi, T., & Razzazi, M. (2011). Development of core to solve the multidimensional multiple-choice knapsack problem. *Computers & Industrial Engineering*, *60*(2), 349–360.

Guo, P., & Tanaka, H. (2001). Fuzzy DEA: a perceptual evaluation method. *Fuzzy Sets and Systems*, *119*(1), 149–160.

Han, B., Leblet, J., & Simon, G. (2010). Hard multidimensional multiple choice knapsack problems, an empirical study. *Computers & Operations Research*, *37*(1), 172–181.

Hatami-Marbini, A., Saati, S., & Tavana, M. (2010). An ideal-seeking fuzzy data envelopment analysis framework. *Applied Soft Computing*, *10*(4), 1062–1070.

Hatami-Marbini, A., Emrouznejad, A., & Tavana, M. (2011). A taxonomy and review of the fuzzy data envelopment analysis literature: two decades in the making. *European Journal of Operational Research*, *214*, 457–472.

Hwang, C. L., & Masud, A. S. M. (1979). *Multiple objective decision making methods and applications*. New York: Springer.

Kao, C., & Liu, S. T. (2003). A mathematical programming approach to fuzzy efficiency ranking. *International Journal of Production Economics*, *86*(2), 145–154.

Kellerer, H., Pferschy, U., & Pisinger, D. (2004). *Knapsack problems*. Berlin: Springer.

Khalili-Damghani, K., & Sadi-Nezhad, S. (2013). A decision support system for fuzzy multi-objective multi-period sustainable project selection. *Computers & Industrial Engineering*, *64*, 1045–1060.

Khalili-Damghani, K., Sadi-Nezhad, S., & Aryanezhad, M. B. (2011). A modular decision support system for optimum investment selection in presence of uncertainty: combination of fuzzy mathematical programming and fuzzy rule based system. *Expert Systems with Applications*, *38*, 824–834.

Khalili-Damghani, K., Tavana, M., & Sadi-Nezhad, S. (2012). An integrated multi-objective framework for solving multi-period project selection problems. *Applied Mathematics and Computation*, *219*, 3122–3138.

Khalili-Damghani, K., Abtahi, A.-R., & Tavana, M. (2013a). A new multi-objective particle swarm optimization method for solving reliability redundancy allocation problems. *Reliability Engineering & Systems Safety*, *111*, 58–75.

Khalili-Damghani, K., Sadi-Nezhad, S., Hosseinzadeh Lotfi, F., & Tavana, M. (2013b). A hybrid fuzzy rule-based multi-criteria framework for sustainable project portfolio selection. *Information Sciences*, *220*, 442–462.

Khalili-Damghani, K., Tavana, M., & Nojavan, M. (2013c). Solving fuzzy multidimensional multiple-choice knapsack problems: the multi-start partial bound enumeration method versus the efficient epsilon-constraint method. *Applied Soft Computing*, *13*, 1627–1638.

Lertworasirikul, S., Fang, S. C., Joines, J. A., & Nuttle, H. L. W. (2003). Fuzzy data envelopment analysis (DEA): a possibility approach. *Fuzzy Sets and Systems*, *139*(2), 379–394.

Lin, F. T. (2008). Solving the knapsack problem with imprecise weight coefficients using genetic algorithm. *European Journal of Operational Research*, *185*(1), 133–145.

Lin, F. T., & Yao, J. S. (2001). Using fuzzy numbers in knapsack problems. *European Journal of Operational Research*, *135*(1), 158–176.

Martello, S., & Toth, P. (1990). *Knapsack problems: algorithms and computer implementations*. *Wiley-Interscience series in discrete mathematics and optimization*. Chichester: Wiley.

Mavrotas, G. (2009). Effective implementation of the ε-constraint method in multi-objective mathematical programming problems. *Applied Mathematics and Computations*, *213*, 455–465.

Mavrotas, G., Diakoulaki, D., & Kourentzis, A. (2008). Selection among ranked projects under segmentation, policy and logical constraints. *European Journal of Operational Research*, *187*(1), 177–192.

Moreno, E., Espinoza, D., & Goycoolea, M. (2010). Large-scale multi-period precedence constrained knapsack problem: a mining application. *Electronic Notes in Discrete Mathematics*, *36*(1), 407–414.

Saati, S., Memariani, A., & Jahanshahloo, G. R. (2002). Efficiency analysis and ranking of DMUs with fuzzy data. *Fuzzy Optimization and Decision Making*, *1*, 255–267.

Sengupta, J. K. (1992). A fuzzy systems approach in data envelopment analysis. *Computers and Mathematics with Applications*, *24*(8–9), 259–266.

Shah, R., & Reed, P. (2011). Comparative analysis of multiobjective evolutionary algorithms for random and correlated instances of multiobjective d-dimensional knapsack problems. *European Journal of Operational Research*, *211*, 466–479.

Sowlati, T., Paradi, J. C., & Suld, C. (2005). Information systems project prioritization using data envelopment analysis. *Mathematical and Computer Modeling*, *41*(11–12), 1279–1298.

Vitner, G., Rozenes, S., & Spreggett, S. (2006). Using data envelop analysis to compare project efficiency in a multi-project environment. *International Journal of Project Management*, *24*(4), 323–329.

Xu, J., Zheng, H., Zeng, Z., Wu, S., & Shen, M. (2012). Discrete time–cost–environment trade-off problem for large-scale construction systems with multiple modes under fuzzy uncertainty and its application to Jinping-II Hydroelectric Project. *International Journal of Project Management.* doi:10.1016/j.ijproman.2012.01.019.

Yu, X., & Gen, M. (2010). *Introduction to evolutionary algorithms*. London: Springer.