

## Multi-objective portfolio optimization of mutual funds under downside risk measure using fuzzy theory

A. Alimi<sup>a\*</sup>, M. Zandieh<sup>b</sup> and M. Amiri<sup>c</sup>

<sup>a</sup>Department of Financial Engineering, Raja University of Qazvin, Qazvin, Iran

<sup>b</sup>Department of Industrial Management, Management and Accounting Faculty, Shahid Beheshti University, G.C., Tehran, Iran

<sup>c</sup>Department of Industrial management, Management and Accounting Faculty, Allameh Tabatabaeei University, Tehran, Iran

### ARTICLE INFO

#### Article history:

Received 6 April 2012  
Received in revised format  
26 April 2012  
Accepted May 24 2012  
Available online  
30 May 2012

#### Keywords:

Clustering  
Portfolio optimization  
Mean-semivariance  
Multi-objective non-linear  
programming  
Fuzzy technique programming  
Pareto optimal solution

### ABSTRACT

Mutual fund is one of the most popular techniques for many people to invest their funds where a professional fund manager invests people's funds based on some special predefined objectives; therefore, performance evaluation of mutual funds is an important problem. This paper proposes a multi-objective portfolio optimization to offer asset allocation. The proposed model clusters mutual funds with two methods based on six characteristics including rate of return, variance, semivariance, turnover rate, Treynor index and Sharpe index. Semivariance is used as a downside risk measure. The proposed model of this paper uses fuzzy variables for return rate and semivariance. A multi-objective fuzzy mean-semivariance portfolio optimization model is implemented and fuzzy programming technique is adopted to solve the resulted problem. The proposed model of this paper has gathered the information of mutual fund traded on NASDAQ from 2007 to 2009 and Pareto optimal solutions are obtained considering different weights for objective functions. The results of asset allocation, rate of return and risk of each cluster are also determined and they are compared with the results of two clustering methods.

© 2012 Growing Science Ltd. All rights reserved

## 1. Introduction

Mutual fund is one of the most popular methods for many people to invest their funds where a professional fund manager invests people's funds based on some special predefined objectives. There are literally various types of mutual funds in the world, which makes it difficult to choose the appropriate one. Therefore, we need to use an appropriate technique to make an assessment on different mutual funds and choose the most efficient ones. Murthi et al. (1997) and Basso & Funari (2001) proposed a data envelopment analysis model to measure the mutual fund performance and Deb & Banerjee (2009) proposed a downside risk analysis for equity mutual funds. Chang et al. (2010) suggested an extended TOPSIS method with different distance approaches.

\* Corresponding author. Tel:+983518228184  
E-mail: amiralimi2005@yahoo.com (A. Alimi)

Portfolio optimization is one of the main problems in modern investment theories. Markowitz (1952, 1959) argued that in portfolio optimization problem there are two criteria: 1) return of portfolio that should be maximized 2) risk of portfolio that should be minimized. In Markowitz mean-variance model there is a set of solutions called efficient frontier where an investor should choose one of them depending his/her risk/reward criteria. The mean-variance model is valid if the return has normal distribution and in case the return is not normally distributed downside risk measures is used instead of variance measure (Vercher et al., 2007).

In recent studies, researchers have used multi-objective portfolio optimization models. Chang et al. (2009) introduced a portfolio optimization model in different risk measures and solved it using genetic algorithm. Anagnostopoulos & Mamanis (2010) proposed a multi-objective model with discrete variables. In fuzzy environment, Ammar & Khalifa (2003) implemented fuzzy variables for portfolio optimization problem and Terol et al. (2006) used fuzzy compromise programming for portfolio selection. They used Sharpe's single index model and defined future beta as a fuzzy number. Jana et al. (2009) defined mean value and variance as trapezoidal fuzzy numbers and transformed them to possibilistic form. They proposed a three-objective-model and solved it using fuzzy programming technique. Clustering analysis is one of data mining approaches, which helps control the scale of many problems and it can be used for optimization problem. Tola et al. (2008) used cluster analysis for portfolio optimization. Chen & Huang (2009) used a two-stage method and clustered observations based on four characteristics: rate of return, standard deviation, turnover rate and Treynor index. They implemented fuzzy variables for return and risk variables. For portfolio optimization problem, they used  $\alpha$ -cut method and optimized return and risk separately and in two models.

In this paper, we use a method originally developed by Chen & Huang (2009) with two more characteristics: Sharpe index and semivariance, and cluster mutual funds using Ward method and k-means method separately and compare their results.

## 2. Definition indices

### (1) Rate of return

The net asset value (NAV) is defined as current market value of a fund minus its liabilities divided by outstanding number of shares (Mobiuss, 2007). Rate of return based on net asset value defined as:

$$R_{i,t} = \frac{NAV_{i,t} - NAV_{i,t-1}}{NAV_{i,t-1}} * 100\% , \quad (1)$$

where  $i$  is the number of mutual funds,  $R_{i,t}$  is the rate of return at time  $t$  and  $NAV_{i,t}$  is the net asset value at time  $t$ .

### (2) Variance

$$\sigma_i^2 = \frac{\sum_{t=1}^T (R_{i,t} - \bar{R}_i)^2}{T - 1} \quad (2)$$

$\bar{R}_i$  is the average rate of return of  $T$  months.

### (3) Downside risk measure (semivariance)

$$SV_i = \frac{\sum_{t=1; R_{i,t} < \bar{R}_i}^T (R_{i,t} - \bar{R}_i)}{T - 1} \quad (3)$$

### (4) Turnover rate

Turnover is a measure of fund's transactions, high turnover rate shows that the fund is an active fund and pays more transaction cost.

$$\text{Turnover rate} = \frac{\text{The amount of values in transactions}}{\text{value of total assets}} * 100\% \quad (4)$$

### (5) Treynor index

Treynor (1965) introduced a ratio for performance evaluation of portfolios.

$$TR_i = \frac{\bar{R}_i - R_f}{\beta_i}, \quad (5)$$

where  $TR_i$  is the Treynor index,  $\bar{R}_i$  is the average return,  $R_f$  is the return of risk-free asset and  $\beta_i$  is the measurement of systematic risk and calculated as:

$$\beta_i = \frac{COV(R_i, R_m)}{\sigma_m^2}, \quad (6)$$

where  $COV(R_i, R_m)$  is the covariance between the return of mutual fund  $i$  and the return of the market index and  $\sigma_m^2$  is the variance of  $R_m$ .

### (6) Sharpe index

Sharpe (1966) proposed a measurement for performance evaluation of mutual funds. We use this index because his study was particularly based on mutual funds.

Where  $SR_i$  is the Sharpe index and  $\sigma_i$  is the standard deviation and called total risk.

$$SR_i = \frac{\bar{R}_i - R_f}{\sigma_i} \quad (7)$$

## 3. Clustering

Clustering is a grouping observations or records into classes of similar objects have named and it is widely used in data mining. A cluster is a group of observations, which are similar to one another and are not similar to observations in other clusters. There are two kinds of algorithms in clustering: 1- hierarchical 2- non-hierarchical. Hierarchical clustering has a tree-like cluster structure and creates a dendrogram, where the tree diagram is implemented to display arrangement of clusters. There are various methods to determine distance among clusters, which lead to create different methods in hierarchical clustering (Larose, 2005). In this study, we choose Ward method in hierarchical clustering methods, where dissimilarity among clusters is the Euclidean distance among their centroids. This method minimizes within-clusters sum of squares based on pair wise distances (Decker, & Lenz, 2007).

There are various non-hierarchical clustering methods and k-means is one of the most commonly methods, where it minimizes the sum of distances between each observation and its cluster center (Zio & Bazzo, 2010). In k-means method,  $k$  refers to the number of clusters, which must be determined before clustering process. K-means method follows this algorithm:

Step 1) Determine  $k$  (number of clusters),

Step 2) Assign  $k$  observations randomly to be the initial cluster center locations,

Step 3) For each observation determine the nearest cluster center, which the nearest criterion is usually Euclidean distance, hence, each cluster center includes a subset of observations and our dataset segments to  $k$  clusters,

Step 4) For each cluster, find the new cluster center and substitute them with random center points, new centers are found using weighted average of observations in each cluster. For example the center of points (1,2) and (2,4) is (1.5,3).

Step 5) Repeat Steps 3 to 5 until termination.

Before we precede clustering process, we need to normalize our observations and the proposed model of this paper uses min-max normalization as follows,

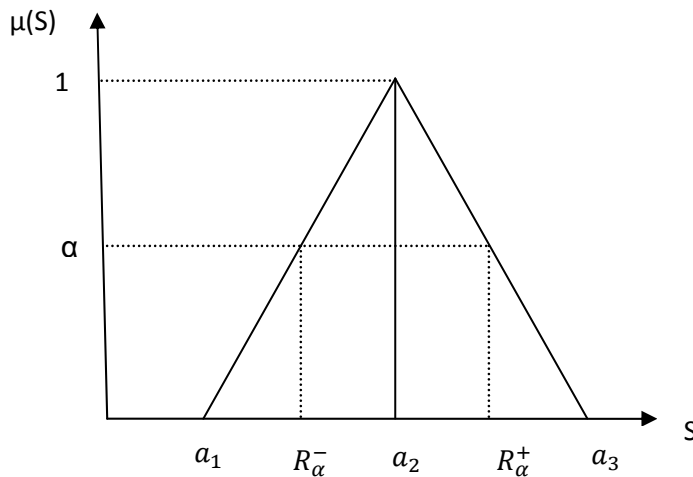
$$x^* = \frac{x - \min(x)}{\text{Range}(x)} = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{8}$$

In Eq. (8),  $x^*$  is a normalized observation,  $x$  is the observation that will be normalized,  $\min(x)$  is the minimum amount of all observations in that index and  $\max(x)$  is the maximum amount of all observations in that index. The advantage of this method is that, after normalization process minimum unit equals to 0, maximum unit equals to 1 and all observations will range zero to one (Larose, 2005).

#### 4. Making fuzzy variables

Note that a crisp number for risk and rate of return obtained from historical data is not usually appropriate for future because there are many uncertainties associated with future and many things may influence on it. Hence, we define risk and rate of return as triangular fuzzy numbers. A triangular fuzzy number could be written as  $[a_1, a_2, a_3]$ . We define fuzzy rate of return membership function as Eq. (9). Fig. 1 shows this function.

$$\mu(S) = \begin{cases} (S - a_1)/(a_2 - a_1) & a_1 < S < a_2 \\ 1 & S = a_2 \\ (a_3 - S)/(a_3 - a_2) & a_2 < S < a_3 \\ 0 & \text{Oth.wise} \end{cases} \tag{9}$$



**Fig. 1.** Membership function of variables

$$\tilde{R}_\alpha = [R_\alpha^-, R_\alpha^+] = [a_1 + \alpha(a_2 - a_1), a_3 - \alpha(a_3 - a_2)] \tag{10}$$

We can make fuzzy risk variable similar to making fuzzy rate of return variable.

#### 5. Fuzzy mean-Semivariance model

Markowitz (1952, 1959) proposed the mean-variance model for portfolio optimization problem. Instead of crisp numbers in the main model we substitute the variables of his model with fuzzy variables as Eq. (11).

$$\begin{aligned}
 &\text{Maximize } E(R_p) = \sum_{i=1}^N \tilde{R}_i x_i && (11) \\
 &\text{Minimize } \sigma_p^2 = \sum_{i=1}^N x_i^2 \tilde{\sigma}_i^2 + 2 \sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j}}^N x_i x_j \tilde{\sigma}_{ij} \\
 &\text{Subject to } \sum_{i=1}^N x_i = 1 \\
 & \quad x_i > 0 \quad i = 1, \dots, N
 \end{aligned}$$

$N$  Number of assets

$x_i$  Proportion invested in asset  $i, i=1, 2, \dots, N$

$\tilde{R}_i$  Fuzzy expected return of asset  $i$

$\tilde{\sigma}_{ij}$  Fuzzy covariance between asset  $i$  and asset  $j$

We assumed short sales are not allowed and there is not any dividend, taxes and transaction costs during the evaluation period. The Markowitz's model is formulated as a convex quadratic programming, where the mean-variance model has an assumption that the rate of return is normally distributed but this assumption does not hold in many cases. Therefore, we propose other risk measures such as semivariance to improve the main model (Chang et al., 2009) and substitute it for the proposed model of this paper. As Chen & huang (2009) discussed we can also eliminate the correlation part in the model because at clustering process the variance among clusters maximized and the correlation between clusters (variables) decreases. Eq. (12) illustrates our new model:

$$\begin{aligned}
 &\text{maximize } E(R_p) = \sum_{i=1}^N \tilde{R}_i x_i && (12) \\
 &\text{minimize } SV_p = \sum_{i=1}^N x_i^2 \tilde{SV}_i \\
 &\text{subject to } \sum_{i=1}^N x_i = 1 \\
 & \quad x_i > 0 \quad i = 1, \dots, N
 \end{aligned}$$

Eq. (12) is one of the most difficult kinds of optimization models to solve. This model is a multi-objective fuzzy non-linear programming. As illustrated in Eq. (10), the risk and return variables are fuzzy and have lower bound and upper bound. Eq. (13) is written using these bounds.

$$\begin{aligned}
 &\text{maximize } \tilde{Z}_R^\alpha(x) = \sum_{i=1}^N [(\tilde{R}_i^-)_\alpha, (\tilde{R}_i^+)_\alpha] x_i && (13) \\
 &\text{minimize } \tilde{Z}_{SV}^\alpha(x) = \sum_{i=1}^N [(\tilde{SV}_i^-)_\alpha, (\tilde{SV}_i^+)_\alpha] x_i^2
 \end{aligned}$$

subject to 
$$\sum_{i=1}^N x_i = 1$$

$$x_i > 0 \quad i = 1, \dots, N$$

The above model could be solved using Kuhn-Tucker conditions considering lower and upper bounds, separately. Hence, we split Eq. (13) into two models that both are multi-objective.

maximize 
$$\tilde{Z}_{\tilde{R}}^\alpha(x) = \sum_{i=1}^N [(\tilde{R}_i^-)_\alpha] x_i \tag{14}$$

minimize 
$$\tilde{Z}_{\tilde{SV}}^\alpha(x) = \sum_{i=1}^N [(\tilde{SV}_i^+)_\alpha] x_i^2$$

subject to 
$$\sum_{i=1}^N x_i = 1$$

$$x_i > 0 \quad i = 1, \dots, N$$

maximize 
$$\tilde{Z}_{\tilde{R}}^\alpha(x) = \sum_{i=1}^N [(\tilde{R}_i^+)_\alpha] x_i \tag{15}$$

minimize 
$$\tilde{Z}_{\tilde{SV}}^\alpha(x) = \sum_{i=1}^N [(\tilde{SV}_i^-)_\alpha] x_i^2$$

subject to 
$$\sum_{i=1}^N x_i = 1$$

$$x_i > 0 \quad i = 1, \dots, N$$

**Definition1 (Complete optimal solution).**  $x^*$  is said to be a complete optimal solution of a r-objective problem if and only if, there exists  $x^* \in X$  such that  $Z_i(x^*) \leq Z_i(X)$  for  $i=1, \dots, r$  and for all  $x \in X$ .

When we solve this problem a complete optimal solution does not always exist, hence Pareto optimal solution is defined as follows

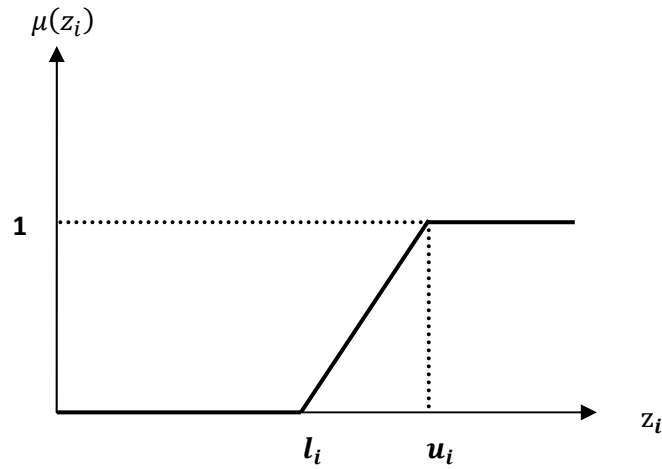
**Definition2 (Pareto optimal solution).**  $x^*$  is said to be a Pareto optimal solution of a r-objective problem if and only if, there does not exist another  $x \in X$ , such that  $Z_i(x^*) \leq Z_i(X)$  for  $i=1, \dots, r$  and  $Z_k(x) \neq Z_k(x^*)$  for at least one  $k, k \in \{1, 2, \dots, r\}$

Eq. (14) considers lower bound of return and upper bound of risk and gives us the worst solution and Eq. (15) considers upper bound of return and lower bound of risk and gives us the best solution. In next section we discuss solution procedure for these two MONLP models using fuzzy theory and obtain Pareto optimal solution.

**6. Fuzzy programming technique**

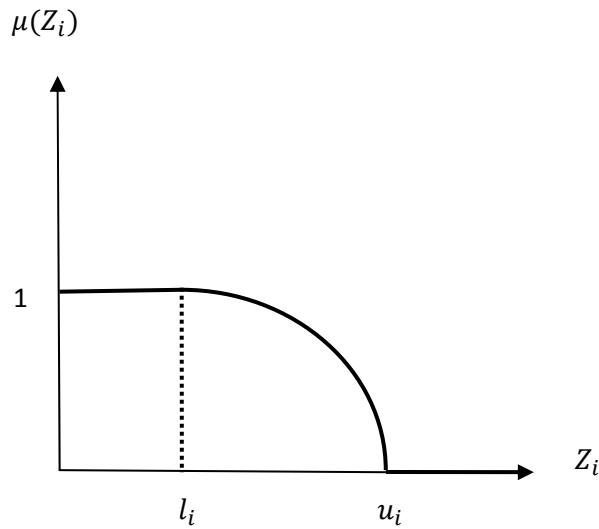
There are different methods to solve MONLP problems and we use the method developed by Zimmermann (1978), which is fuzzy programming technique. In the first stage of fuzzy programming technique, each objective function is optimized with constraints, separately. The objective functions minimizes to obtain lower bound ( $l_i$ ) and maximizes to obtain upper bound ( $u_i$ ). For our problem, we first minimize and maximize the return objective function and then the method minimizes and maximizes the risk objective function considering the constraint. Then we define objective functions as

two fuzzy functions. Fig. 2 and Fig. 3 illustrate the membership function of these fuzzy functions and Eqs. (16-17) define these functions.



**Fig. 2.** Membership function of return objective function

$$\mu(z_i) = \begin{cases} 0 & \text{if } z_i \leq l_i \\ \frac{z_i - l_i}{u_i - l_i} & \text{if } l_i < z_i < u_i \\ 1 & \text{if } z_i \geq u_i \end{cases} \quad (16)$$



**Fig. 3.** Membership function of risk objective function

$$\mu(z_i) = \begin{cases} 0 & \text{if } z_i \geq u_i \\ \frac{u_i - z_i}{u_i - l_i} & \text{if } l_i < z_i < u_i \\ 1 & \text{if } z_i \leq l_i \end{cases} \quad (17)$$

We define  $\beta_i$  as the percentage rate that  $i^{\text{th}}$  objective function nears to its optimized solution or minimum amount of membership function. Eq. (18) shows this definition.

$$\beta_i = \min (\mu(z_i)) \rightarrow \beta_i \leq \mu(z_i) \tag{18}$$

In return objective function we have Eq. (19) as follows,

$$\beta_1 \leq \frac{z_i - l_i}{u_i - l_i} \rightarrow z_i \geq l_i + \beta_1(u_i - l_i), \tag{19}$$

and in risk objective function we have Eq. (20) as follows,

$$\beta_2 \leq \frac{u_i - z_i}{u_i - l_i} \rightarrow z_i \leq u_i - \beta_2(u_i - l_i) \tag{20}$$

Now we can rewrite our model as follows,

$$\text{maximize } P = w_1\beta_1 + w_2\beta_2 \tag{21}$$

$$\text{subject to } \sum_{i=1}^N [(\tilde{R}_i^-)_\alpha] x_i \geq l_i + \beta_1(u_i - l_i)$$

$$\sum_{i=1}^N [(\tilde{S}_i^+)_\alpha] x_i^2 \leq u_i - \beta_2(u_i - l_i)$$

$$\sum_{i=1}^N x_i = 1$$

$$x_i > 0 \quad i = 1, \dots, N$$

$$0 \leq \beta_1, \beta_2 \leq 1$$

$$w_1, w_2 \geq 0$$

$P$  percentage that objective functions additionally near to their optimized condition,

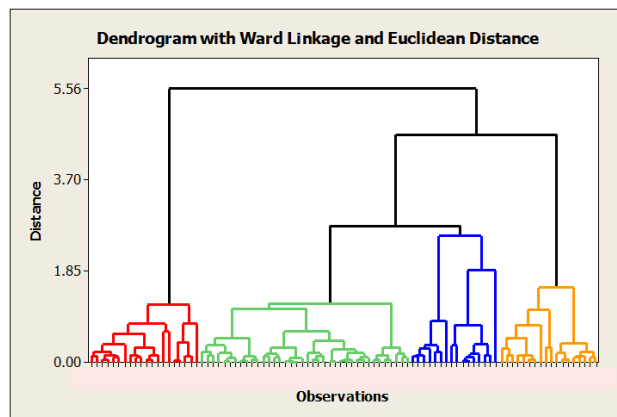
$w_1$  Weight of return objective function,

$w_2$  Weight of risk objective function,

We can assume different  $\alpha$ -levels and LINGO11.0 software can be used to solve this model. The second constraint of model is non-linear and software gives local optimum but because our solution space is convex, the local optimum is also global optimum.

### 7. Numerical example

In order to study the performance of the proposed model of this paper, we have selected 92 equity mutual funds from NASDAQ from 2007 to 2009 (available online at <http://finance.yahoo.com>).



**Fig. 4.** Dendrogram of Ward method



For these cases NAV and market information are extracted and the characteristics were calculated for each fund based on return rate, variance, semivariance, turnover rate, Treynor index and Sharpe index. Finally, all data were normalized leaving us to have 92 rows and 6 columns where all of numbers ranged [0,1]. For clustering analysis Minitab 14 software helped us. Observations clustered using Ward method and  $k$ -means method considering Euclidean distance. Fig. 4. illustrates the dendrogram obtained from Ward method.

In clustering analysis, the number of clusters should be determined and there is not a common method to accomplish this task. In Ward method we can fix distance or similarity level and obtain number of clusters. In our example distance level is fixed to 2.5. But in  $k$ -means method the number of clusters ( $k$ ) should be specified before clustering process. Trial and error procedure can be useful. We can first fix  $k=2$  and then increase it and find the appropriate number. We observed 4 clusters are more appropriate. Number of clusters more or less than 4 separated a few observations with abnormal characteristics. The results obtained from Ward and  $k$ -means method are shown in Tables 1 and Table 2.

**Table 1**

Results of ward clustering method

	Number of observations	Return rate	variance	Semi variance	Turnover rate	Treynor index	Sharpe index
Cluster 1	20	0.574	0.11	0.127	0.096	0.236	0.151
Cluster 2	38	0.749	0.185	0.204	0.089	0.386	0.363
Cluster 3	16	0.777	0.231	0.222	0.307	0.505	0.495
Cluster 4	18	0.81	0.398	0.522	0.071	0.332	0.366

**Table 2**Results of  $k$ -means clustering method

	Number of observations	Return rate	variance	Semi variance	Turnover rate	Treynor index	Sharpe index
Cluster 1	23	0.574	0.112	0.132	0.089	0.254	0.177
Cluster 2	41	0.763	0.233	0.269	0.132	0.351	0.34
Cluster 3	11	0.829	0.518	0.644	0.165	0.341	0.403
Cluster 4	17	0.785	0.133	0.123	0.131	0.557	0.523

In next stage two characteristics remained (return rate and semivariance) for modeling. Four eliminated characteristics (variance, turnover rate, Treynor index and Sharpe index) helped us for more accurate clustering. After clustering process the mean ( $\mu$ ) and standard deviation ( $\sigma$ ) of each cluster are specified. For making fuzzy variables we used original data and assumed  $a_1, a_2, a_3$  as Eq. (22).

$$\begin{cases} a_1 = \mu - 2\sigma \\ a_2 = \mu \\ a_3 = \mu + 2\sigma \end{cases} \quad (22)$$

Tables 3 and Table 4 show the results of fuzzy variables in Ward and  $k$ -means clustering methods.

**Table 3**

Results of making fuzzy variables in Ward method

	Return rate	Semi variance
Cluster 1	[-0.845, 0.021, 0.896]	[3.138, 89.168, 175.199]
Cluster 2	[0.14, 0.443, 0.746]	[39.241, 135.874, 232.506]
Cluster 3	[0.223, 0.511, 0.799]	[45.597, 146.99, 228.241]
Cluster 4	[0.103, 0.589, 1.075]	[102.201, 328.971, 555.741]

**Table 4**  
Results of making fuzzy variables in *k*-means method

	Return rate	Semi variance
Cluster 1	[-0.716, 0.027, 0.77]	[7.81, 92.55, 177.294]
Cluster 2	-----	-----
Cluster 3	[-0.03, 0.634, 1.298]	[187.3, 402.955, 618.681]
Cluster 4	[0.14, 0.53, 0.92]	[10.848, 86.962, 163.076]

As shown in Table 4, Cluster 2 is ignored because its return rate was less than Cluster 4 and its semivariance was more. Each fuzzy number has a membership function. For example the membership function of return rate in Cluster 1 and Ward method is as follows,

$$\mu(R) = \begin{cases} (R + 0.716)/(0.027 + 0.716) & -0.716 < R < 0.027 \\ 1 & R = 0.027 \\ (0.77 - R)/(0.77 - 0.027) & 0.027 < R < 0.77 \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

The membership function of other variables can be written as Eq. (23). Fuzzy return rate and semivariance are our inputs for portfolio optimization model. In our example, there were two multi-objective problems. One of them is as follows,

$$\begin{aligned} \text{maximize } \tilde{Z}_R^\alpha(x) &= [0.871\alpha - 0.845, -0.871\alpha + 0.896]x_1 + & (24) \\ & [0.303\alpha + 0.14, -0.303\alpha + 0.746]x_2 + \\ & [0.288\alpha + 0.223, -0.288\alpha + 0.799]x_3 + [0.486\alpha + 0.103, -0.486\alpha + 1.075]x_4 \\ \text{minimize } \tilde{Z}_{SV}^\alpha(x) &= [86.031\alpha + 3.138, -86.031\alpha + 175.199]x_1^2 + \\ & [96.632\alpha + 39.241, -96.632\alpha + 232.506]x_2^2 + \\ & [101.322\alpha + 45.597, -101.322\alpha + 248.241]x_3^2 + \\ & [226.77\alpha + 102.201, -226.77\alpha + 555.741]x_4^2 \\ \text{subject to } & \sum_{i=1}^4 x_i = 1 \\ & x_i > 0 \quad i = 1, \dots, 4 \end{aligned}$$

The above problem is a multi-objective fuzzy non-linear problem. To solve this problem we assumed 5 levels for  $\alpha$ : [0, 0.25, 0.5, 0.75, 1].

For example if we assume  $\alpha=0.5$  and as explained before split the model into two models we have Eqs. (25-26).

$$\begin{aligned} \text{maximize } \tilde{Z}_R^\alpha(x) &= -0.409x_1 + 0.291x_2 + 0.367x_3 + 0.346x_4 & (25) \\ \text{minimize } \tilde{Z}_{SV}^\alpha(x) &= 132.183x_1^2 + 184.19x_2^2 + 197.58x_3^2 + 442.356x_4^2 \\ \text{subject to } & \sum_{i=1}^4 x_i = 1 \\ & x_i > 0 \quad i = 1, \dots, 4 \end{aligned}$$

For the first objective function lower bound and upper bound are [-0.409, 0.367] and for the second are [49.221, 442.356].

$$\begin{aligned}
 &\text{maximize } \tilde{Z}_R^\alpha(x) = 0.46x_1 + 0.594x_2 + 0.655x_3 + 0.832x_4 \\
 &\text{minimize } \tilde{Z}_{SV}^\alpha(x) = 46.153x_1^2 + 87.557x_2^2 + 96.258x_3^2 + 215.586x_4^2 \\
 &\text{subject to } \sum_{i=1}^4 x_i = 1 \\
 &\quad x_i > 0 \quad i = 1, \dots, 4
 \end{aligned}
 \tag{26}$$

For the first objective function lower bound and upper bound are [0.46, 0.832] and for the second are [20.783, 215.586]. Now we can write the last model that can be solved with optimization software package as follows,

$$\begin{aligned}
 &\text{maximize } P = w_1\beta_1 + w_2\beta_2 \\
 &\text{subject to } -0.409x_1 + 0.291x_2 + 0.367x_3 + 0.346x_4 \geq -0.409 + 0.776\beta_1 \\
 &\quad 132.183x_1^2 + 184.19x_2^2 + 197.58x_3^2 + 442.356x_4^2 \leq 442.36 - 393.139\beta_2 \\
 &\quad \sum_{i=1}^4 x_i = 1 \\
 &\quad x_i > 0 \quad i = 1, \dots, 4 \\
 &\quad 0 \leq \beta_1, \beta_2 \leq 1 \\
 &\quad w_1, w_2 \geq 0
 \end{aligned}
 \tag{27}$$

Eq. 26 can be written the same as Eq. 27. In this example, we assumed 3 weights:

$[w_1 = 0.25, w_2 = 0.75], [w_1 = w_2 = 0.5], [w_1 = 0.75, w_2 = 0.25]$ . Tables 5-14 illustrate results for our numerical example completely. Except  $\alpha=1$ , in other levels of  $\alpha$ , investment proportion variables and return rate and risk variables have interval form.

**Table 5**  
Pareto optimal solution of Ward method with  $\alpha=0$

	$x_1$	$x_2$	$x_3$	$x_4$	Return rate	Risk
$w_1 = 0.25, w_2 = 0.75$	[0.77, 0.885]	[0, 0.383]	[0, 0.385]	[0.115, 0.155]	[0.09, 0.916]	[3.809, 85.292]
$w_1 = 0.5, w_2 = 0.5$	[0, 0.713]	[0, 0.399]	[0, 0.45]	[0.151, 0.287]	[0.172, 0.947]	[10.13, 99.955]
$w_1 = 0.75, w_2 = 0.25$	[0, 0.199]	[0, 0.346]	[0, 0.555]	[0.099, 0.801]	[0.182, 1.039]	[65.696, 109.746]

**Table 6**  
Pareto optimal solution of Ward method with  $\alpha=0.25$

	$x_1$	$x_2$	$x_3$	$x_4$	Return rate	Risk
$w_1 = 0.25, w_2 = 0.75$	[0.076, 0.426]	[0.154, 0.38]	[0.206, 0.384]	[0.16, 0.214]	[0.183, 0.746]	[16.262, 76.618]
$w_1 = 0.5, w_2 = 0.5$	[0, 0.221]	[0.049, 0.387]	[0.251, 0.447]	[0.166, 0.479]	[0.253, 0.822]	[42.281, 89.495]
$w_1 = 0.75, w_2 = 0.25$	[0, 0]	[0, 0.312]	[0, 0.546]	[0.142, 1]	[0.26, 0.953]	[96.797, 158.893]

**Table 7**  
Pareto optimal solution of Ward method with  $\alpha=0.5$

	$x_1$	$x_2$	$x_3$	$x_4$	Return rate	Risk
$w_1 = 0.25, w_2 = 0.75$	[0.076, 0.243]	[0.261, 0.375]	[0.293, 0.382]	[0.167, 0.203]	[0.276, 0.628]	[25.837, 67.833]
$w_1 = 0.5, w_2 = 0.5$	[0, 0]	[0.235, 0.371]	[0.38, 0.443]	[0.186, 0.385]	[0.335, 0.709]	[50.69, 79.431]
$w_1 = 0.75, w_2 = 0.25$	[0, 0]	[0, 0.261]	[0.244, 0.536]	[0.203, 0.756]	[0.343, 0.789]	[87.54, 128.946]

**Table 8**  
Pareto optimal solution of Ward method with  $\alpha = 0.75$

	$x_1$	$x_2$	$x_3$	$x_4$	Return rate	Risk
$w_1 = 0.25, w_2 = 0.75$	[0.088, 0.122]	[0.314, 0.364]	[0.334, 0.374]	[0.174, 0.19]	[0.362, 0.532]	[36.183, 57.83]
$w_1 = 0.5, w_2 = 0.5$	[0, 0]	[0.173, 0.351]	[0.34, 0.435]	[0.214, 0.241]	[0.42, 0.645]	[52.543, 69.972]
$w_1 = 0.75, w_2 = 0.25$	[0, 0]	[0.201, 0.292]	[0.405, 0.513]	[0.286, 0.303]	[0.432, 0.603]	[54.466, 83.342]

**Table 9**  
Pareto optimal solution of Ward method with  $\alpha = 1$

	$x_1$	$x_2$	$x_3$	$x_4$	Return rate	Risk
$w_1 = 0.25, w_2 = 0.75$	0.117	0.344	0.358	0.181	0.445	46.906
$w_1 = 0.5, w_2 = 0.5$	0	0.327	0.423	0.25	0.508	61.377
$w_1 = 0.75, w_2 = 0.25$	0	0.125	0.477	0.398	0.532	87.662

**Table 10**  
Pareto optimal solution of  $k$ -means method with  $\alpha = 0$

	$x_1$	$x_2$	$x_3$	$x_4$	Return rate	Risk
$w_1 = 0.25, w_2 = 0.75$	[0.022, 0.138]	-----	[0.157, 0.164]	[0.705, 0.814]	[-.005, 0.979]	[12.229, 99.679]
$w_1 = 0.5, w_2 = 0.5$	[0, 0]	-----	[0.14, 0.385]	[0.615, 0.86]	[0.115, 1.065]	[31.867, 132.737]
$w_1 = 0.75, w_2 = 0.25$	[0, 0]	-----	[0.002, 1]	[0, 0.998]	[0.139, 1.298]	[162.427, 187.308]

**Table 11**  
Pareto optimal solution of  $k$ -means method with  $\alpha = 0.25$

	$x_1$	$x_2$	$x_3$	$x_4$	Return rate	Risk
$w_1 = 0.25, w_2 = 0.75$	[0.125, 0.14]	-----	[0.163, 0.174]	[0.686, 0.712]	[0.124, 0.843]	[21.931, 90.468]
$w_1 = 0.5, w_2 = 0.5$	[0, 0]	-----	[0.157, 0.347]	[0.653, 0.843]	[0.22, 0.929]	[41.785, 116.288]
$w_1 = 0.75, w_2 = 0.25$	[0, 0]	-----	[0.065, 0.822]	[0.178, 0.935]	[0.229, 1.077]	[128.316, 163.941]

**Table 12**  
Pareto optimal solution of  $k$ -means method with  $\alpha = 0.5$

	$x_1$	$x_2$	$x_3$	$x_4$	Return rate	Risk
$w_1 = 0.25, w_2 = 0.75$	[0.109, 0.144]	-	[0.17, 0.177]	[0.679, 0.721]	[0.255, 0.72]	[32.834, 81.356]
$w_1 = 0.5, w_2 = 0.5$	[0, 0]	-	[0.146, 0.18]	[0.82, 0.854]	[0.328, 0.76]	[41.959, 100.614]
$w_1 = 0.75, w_2 = 0.25$	[0, 0]	-	[0.146, 0.646]	[0.354, 0.854]	[0.329, 0.881]	[102.067, 129.3]

**Table 13**  
Pareto optimal solution of  $k$ -means method with  $\alpha = 0.75$

	$x_1$	$x_2$	$x_3$	$x_4$	Return rate	Risk
$w_1 = 0.25, w_2 = 0.75$	[0.107, 0.133]	-	[0.175, 0.179]	[0.688, 0.718]	[0.374, 0.603]	[44.5968, 69.936]
$w_1 = 0.5, w_2 = 0.5$	[0, 0]	-	[0.21, 0.275]	[0.725, 0.79]	[0.439, 0.674]	[62.106, 86.298]
$w_1 = 0.75, w_2 = 0.25$	[0, 0]	-	[0.252, 0.499]	[0.501, 0.748]	[0.44, 0.713]	[88.318, 103.97]

**Table 14**Pareto optimal solution of  $k$ -means method with  $\alpha=1$ 

	$x_1$	$x_2$	$x_3$	$x_4$	Return rate	Risk
$w_1 = 0.25, w_2 = 0.75$	0.12	-	0.178	0.702	0.487	56.956
$w_1 = 0.5, w_2 = 0.5$	0	-	0.242	0.758	0.554	73.566
$w_1 = 0.75, w_2 = 0.25$	0	-	0.37	0.63	0.568	89.685

Results show that return rate is still triangular fuzzy number. For example in  $k$ -means method with weights [ $w_1 = 0.75, w_2 = 0.25$ ] we can say that return rate is (0.139, 0.568, 1.298). However, membership function of risk variable is not known because its objective function is a quadratic function. However, we make fuzzy variables and solve the problem using  $\alpha$ -cut method to have fuzzy variables in results. In  $\alpha$ -cut method when  $\alpha=1$  means that the problem is solved in certain condition. In other words, we solve our problem in fuzzy and certain condition simultaneously. The results show that in Ward method investment proportion of variable  $x_3$  is higher than others and in  $k$ -means method variable  $x_4$  is higher. These variables have more influence in forming portfolio and we can say that these clusters are better than others. Referring to data before clustering, number of funds in Cluster 3 in Ward method and Cluster 4 in  $k$ -means method is illustrated in Table 15.

**Table 15**

Comparison of number of clusters

	Number of clusters																
Ward	5	10	11	12	32	41	49	51	63	73	75	78	79	85	91	92	
$k$ -means	5	8	12	20	32	34	41	51	52	56	58	73	75	83	90	91	92

As shown in Table 15 although more than half of funds are common in two methods but we should specify that clustering method is effective on our methodology. On the other hand, our methodology suggests 16 or 17 mutual funds out of 92 mutual funds and also determines the investment proportion in each cluster.

## 8. Conclusions

Since mutual funds developed rapidly in recent years, evaluating their performance has been an important subject. Several multi-criteria decision making methods were explored and in this paper multi-objective portfolio optimization is chosen. In current study, performance indices such as rate of return, variance, semivariance, turnover rate and Treynor ratio and Sharpe ratio are calculated and then clustering based on these indices is done. We used Ward method as a hierarchical and  $k$ -means as a non-hierarchical method separately. We made fuzzy variables from return rate and semivariance. Risk measurement is very important in portfolio optimization. Semivariance is used as a downside risk measure. A multi-objective fuzzy mean-semivariance model is made. This model is solved with fuzzy technique programming.

As shown in results, solutions with confidence level less than one are interval. These interval numbers helps investors to decide better. They know the range of return and risk and choose a number in investment proportion interval. It means that our methodology considers investors preferences. Results show that we cannot prefer Ward method rather than  $k$ -means and vice versa because the Pareto optimal solutions of clustering methods do not dominate each other. Because the clustering method was effective in our methodology, future studies can follow this methodology without clustering and compare the results and as another work can complete the optimization model and solve it using advanced techniques such as metaheuristics.

## References

- Ammar, E., & Khalifa, H. A. (2003). Fuzzy portfolio optimization: A quadratic programming approach. *Chaos, Solitons and Fractals*, 18, 1045–1054.
- Anagnostopoulos, K. P., & Mamanis, G. (2010). A portfolio optimization model with three objectives and discrete variables. *Computers & Operations Research*, 37, 1285–1297.
- Basso, A., & Funari, S. (2001). A data envelopment analysis approach to measure the mutual fund performance. *European Journal of Operational Research*, 135, 477–492.
- Chang, C. H., Lin, J. J., Lin, J. H., & Chiang, M. C. (2010). Domestic open-end equity mutual fund performance evaluation using extended TOPSIS method with different distance approaches. *Expert Systems with Applications*, 37, 4642–4649.
- Chang, T. J., Yang, S. C., & Chang, K. J. (2009). Portfolio optimization problem in different risk measures using genetic algorithm. *Expert Systems with Applications*, 36, 10529–10537.
- Chen, L. H., & Huang, L. (2009). Portfolio optimization of equity mutual funds with fuzzy return rates and risks. *Expert Systems with Application*, 36, 3720–3727.
- Deb, G. S., & Banerjee, A. (2009). Downside risk analysis of Indian equity mutual funds: a value at risk approach. *International Research Journal of Finance and Economics*, 23, 216–230.
- Decker, R., & Lenz, H. J. (2007). *Advances in Data Mining*. Springer, New York.
- Jana, P., Roy, T. K., & Mazumder, S. K. (2009). Multi-objective possibilistic model for portfolio selection with transaction cost. *Journal of Computational and Applied Mathematics*, 228, 188–196.
- Larose, D. T. (2005). *Discovering Knowledge in Data*. New Jersey: John Wiley & Sons.
- Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance*, 7, 77–91.
- Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. New York: Wiley
- Mobius, M. (2007). *Mutual Funds: An Introduction to the Core Concepts*. Singapore: John Wiley & Sons (Asia) Pte Ltd.
- Murthi, B. P. S., Choi, Y. K., & Desai, P. (1997). Efficiency of mutual funds and portfolio performance measurement: A non-parametric approach. *European Journal of Operational Research*, 98, 408–418.
- Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business*, 39, 119–138.
- Terol, A. B., Gladish, B. P., Parra, M. A., & Uria, M. V. R. (2006). Fuzzy compromise programming for portfolio selection. *Applied Mathematics and Computation*, 173, 251–264.
- Tola, V., Lillo, F., Gallegati, M., & Mantegna, R. N. (2008). Cluster analysis for portfolio optimization. *Journal of Economic Dynamics and Control*, 32, 235–258.
- Treynor, J. (1965). How to rate management of investment funds. *Harvard Business Review*, 43, 63–75
- Vercher, E., Bermudez, J. D. & Segura, J. V. (2007). Fuzzy portfolio optimization under downside risk measures. *Fuzzy Sets and Systems*, 158, 769–782.
- Zimmermann, H. J. (1978). Fuzzy programming and linear programming with several Objective Functions. *Fuzzy Sets and Systems*, 1, 45–55.
- Zio, E., & Bazzo, R. (2010). Multi-objective optimization of the inspection intervals of a nuclear safety system: A clustering-based framework for reducing the Pareto Front. *Annals of Nuclear Energy*, 37, 798–812.