Yugoslav Journal of Operations Research
28 (2018), Number 3, 385–414
DOI: https://doi.org/10.2298/YJOR170217023J

A MULTI-CRITERIA DECISION-MAKING APPROACH TO PERFORMANCE EVALUATION OF MUTUAL FUNDS: A CASE STUDY IN SERBIA

Milena JAKŠIĆ
University of Kragujevac, Faculty of Economics,Serbia
milenaj@.ac.rs

Predrag MIMOVIĆ
University of Kragujevac, Faculty of Economics, Serbia
mimovicp@kg.ac.rs

Miljan LEKOVIĆ
University of Kragujevac, Faculty of Hotel Management and Tourism in Vrnjačka Banja, Serbia
m.lekovic@kg.ac.rs

Received: February 2017 / Accepted: November 2017

Abstract: In this paper we evaluate the performance of nine mutual funds in the Republic of Serbia in the period 2011-2013 by integrating traditional approaches for measuring absolute efficiency and the selected multi-criteria decision-making methods for measuring relative efficiency. The aim of our research is to test selection abilities of Serbian portfolio managers. Performance evaluation of mutual funds, being by its nature a complex problem of multi-criteria decision-making, must be solved by the methods that have, at least, the same level of complexity. Research results indicate that mutual funds have inferior performance, which on the other hand, confirms that the national portfolio managers lack selection abilities.

Keywords: Sharpe Index, Treynor Index, Jensen's alpha Index, Ivaluation, Ierformance, Efficiency, Mutual Funds, Multi-criteria Decision-Making, AHP, DEA, DEAHP.

MSC: 90B50, 90B90, 97M40.

1. INTRODUCTION

In modern business, with strong dynamic changes in its environment, the task of mutual funds management is to continuously monitor users' preferences of financial services, competition activities, performance of internal processes, and overall financial situation. Therefore, performance measurement and evaluation is the basis for reviewing the current situation in business behavior and the possibilities for its change. Performance measurement allows mutual funds not only to measure the degree of realization of the defined goals, but also to observe the key factors that lead to the improvement or deterioration of business results.

Financial environment constantly imposes the need for finding and defining new concepts and models of performance measurement in order to improve operations quality of mutual funds, and the efficiency of financial system in any market-oriented economy. Crisis on financial markets, followed by the problems, faced by financial institutions, caused primarily by lower prices of securities, requires a review of the established models and management approaches in the management of financial assets. So, managers face the issue on whether active management of mutual fund portfolio brings better results than those that would be realized by investing in assets faithfully reflected in some leading stock index. The intention is to determine whether active management of mutual fund portfolio helps managers to achieve higher return than the market return.

Modern management aims at developing a wide range of models that will allow portfolio of securities construction, enabling the achievement of stable return in the medium and long term. In the past, investors were almost solely interested in high-return funds investment, but the bankruptcy of many such investments forced them to pay particular attention to another dimension of fund performance, i.e., to risk. Experience shows that high-return mutual funds owe their score more to the high level of the undertaken risk, and the overall market trends, than to the ability of portfolio managers.

Bearing this in mind, this study will focus on measuring and evaluating the performance of mutual funds. The main objective is to observe selection abilities of Serbian portfolio managers, based on the analysis of performance of mutual funds in Serbia from 2011 to 2013. Our focus is on mutual funds, given their dominance, not only in Serbia but in the world, and also on their number and the value of assets they manage. So, our general research objective points to two specific objectives.

The first specific objective is to compare risk-weighted return of mutual funds in Serbia with the risk-weighted return of the leading Belgrade Stock Exchange index, Belex15.

The second is to obtain a more comprehensive and a more objective performance measure of the observed mutual funds than the one would be in the case if only traditional performance measurement indices were observed. Hence, we apply multi-criteria decision-making methods.

Based on the defined research subject and objectives, we test the following hypotheses:

H_1 : Mutual fund portfolio has superior/inferior performance compared to the market portfolio.

H_2 : Integrated application of traditional and multi-criteria methods gives a more objective and a more comprehensive performance measure of mutual funds than the one got in the case with their partial implementation.

According to the defined research subject, the objective, and the established research hypotheses, the paper is structured as follows: the first part focuses on reviewing the literature in the field of performance measurement of mutual funds. This is followed by a brief overview of the traditional performance measures of mutual funds and the description of DEA (Data Envelopment Analysis), AHP (Analytic Hierarchy Process), and DEAHP (Data Envelopment Analytic Hierarchy Process) methods, in section 3. What follows, in section 4, is the problem description and an overview of the traditional approach to performance measurement in the case of nine mutual funds in the Republic of Serbia. Risk-weighted returns of mutual funds are compared with risk-weighted returns of the leading Belgrade Stock Exchange index, Belex15, using the following performance measures: Sharpe index (S_i), Treynor index (T_i), and Jensen's alpha index (α_i). Section 5 deals with individual and combined application of AHP and DEAHP methods in evaluating performance of the selected funds and the analysis of the results. Due to the volume of the article, less attention is dedicated to the method of forming DEA model, which resulted in a mere presentation of the results.

2. LITERATURE REVIEW

From early 60s, measuring the performance of mutual funds has become an integral part of financial literature in developed countries. Scientific and professional literature abounds in works that directly or indirectly deal with the issue of measuring and evaluating performance of mutual funds. Though, we will mention some of the most important in terms of theoretical, methodological, and practical significance for the context of multi-criteria methodology used in this paper. Tangen [44] classifies all models for measuring organizational performance into three categories: 1) the first class – fully integrated models, 2) the second class – balanced, multidimensional models, and 3) the third class – financial, one-dimensional models. According to Tangen, the most advanced performance measurement models belong to the first class, regarding their meeting high standards both in terms of available information and the measures that explain causal relationships throughout the organization. The third class consists of models that mostly use traditional performance measures. Even though their objectives are lower, it is important to respect the basic principles of performance measurement. Finally, the second class consists of models that take a more balanced approach to performance measurement than the third-class models, using non-financial measures, different horizons and organizational levels of observation. Although each class has its own specifics, according to Tangen, it is difficult to draw a strict dividing border between them, and therefore, he recommends the lower classes methods to be used in situations where the existing performance measurement

system moves between two classes. There is a wide range of performance measures, and the choice depends on what should be measured and how, as well as on the complexity of the observed organization.

Measuring the performance of mutual funds has become an integral part of financial literature in developed countries in early 1960s. The first empirical analysis of the performance of mutual funds was conducted by Friend, Brown, Herman, and Vickers in their work " *A Study of Mutual Funds*" , published in 1962 [34]. A few years later, Jack Treynor[45], William Sharpe [39], andMichael Jensen[25], independently of each other, introduced standard performance measures, later known as Sharpe, Treynor, and Jensen's alpha index. Starting from Jensen's study, conducted in 1967, most academic studies conclude that net performance of mutual funds is inferior in comparison with market performance, i.e. the majority of papers suggest that actively managed mutual funds are not able to outperform market index returns. Analyzing the performance of 115 mutual funds in the period 1945-1964, Jensen (1967) concludes that their managers failed to achieve returns higher than the expected, considering the level of risk taken. Chang and Lewellen [6],Bogle [5],Droms and Walker[16],Harlow and Brown[20] reach similar conclusion. However, in the late 1980s and early 1990s, conflicting studies appeared, like the one presented by Ippolito [21], with the conclusion that mutual funds own enough private information to outweigh the created costs [30].

Financial literature is especially famous for performance evaluation of European mutual funds, carried out by Otten and Bams [30], based on the sample of 506 funds in five countries: France (99 funds), Germany (57 funds), Italy (37 funds), the Netherlands (9 funds), and Great Britain (304 funds). The conclusion of their study is that the average European mutual fund is able to add value, i.e. exceed the relevant market indices, as indicated by positive net alphas. Unfortunately, the obtained results lack statistical significance, which has, in truth, been achieved by the addition of management fees, when mutual funds, in the case of four out of five countries analyzed, achieved positive and statistically significant gross alphas.

On the other hand, literature on mutual funds and their performance measurement in less developed countries, such as the countries of Central and Eastern Europe, is relatively scarce, despite the fact that these countries have, with the fall of socialism and the transition to market-oriented economic system, attracted considerable investors' attention. The issue of performance evaluation of mutual funds in Central and Eastern Europe attracted researchers attention at the beginning of 2000s. Analyzing the performance of mutual funds in Poland in the period 2000-2008, based on the sample of 140 funds, Bialkowski and Otten [4] conclude that Polish mutual funds, on average, are not able to add value, i.e. outperform the relevant market indicies, as indicated by negative net alphas. The abovementioned authors, however, acknowledge that the addition of management fees leads to positive and significant alphas for domestic funds and to negative alphas, without statistical significance for international funds. These results suggest that domestic mutual funds in Poland are more successful than the international funds due to information superiority of the domestic investors over the foreign, as well

as to their managers selection abilities, but who charge excessively high fees.

Swinkels and Rzezniczak [42] evaluated performance of Polish mutual funds over one year shorter period, 2000-2007, based on the sample of 38 Polish mutual funds. In measuring performance, these authors got positive alphas, but not statistically significant, which implies that mutual fund portfolio has the same performance as the market portfolio. In other words, they failed to prove either superiority or inferiority of fund performance compared to market performance. Markovic-Hribernik and Vek[27] got similar results, analyzing performance of mutual funds in Slovenia, belonging to the Energy policy sector, in the period from January 2005 to August 2009. Seven out of nine surveyed funds had positive alpha indices of small nominal value, but none of them had the necessary statistical significance, so the authors could not confirm selection superiority of managers of mutual funds.

Jagric et al. [23] measured the performance of mutual funds in Slovenia as well, but the results of their research were somewhat different. The authors limited their study to the period 1 July 2000 – 31 December 2003, and the funds older than three years. All nine of the analyzed funds achieved positive alpha index values, six of which were statistically significant. This suggests that, based on the present research, the managers of Slovenian mutual funds in the reporting period were able to outperform the market by showing remarkable selection abilities. This is further confirmed by the results obtained by Podobnik et al. [32], analyzing the performance of Slovenian mutual funds on the sample of fourteen funds in the period from 31 December 1999 to 31 August 2006. All the observed funds realized positive alpha indices, while 50% were statistically significant. In the same work, they evaluated performance of Croatian and Bosnian mutual funds. Out of fourteen surveyed mutual funds in Croatia in the period from 1 January 2004 to 31 December 2005, eleven funds achieved positive alpha indices, but only one was statistically significant. In Bosnia, eight out of nine analyzed funds, in the three-year period from 1 April 2003 to 1 April 2006, reached positive alpha indices, reflecting the potential selection superiority of their managers. However, as in the case of Croatia, only one alpha index had the necessary statistical significance. These authors reached the conclusion about obvious dominance of Slovenian mutual funds as compared to Croatian and Bosnian funds when performance and selection ability of their managers are concerned.

3. METHODOLOGY

Corporate or organizational performance is multi-dimensional, influenced by numerous and diverse factors such as: 1) financial factors that affect financial position of a company or organization, 2) strategic factors of a qualitative nature that define the company's internal activities and their relationship with the market (organization, management, market trends, etc.), and 3) economic factors that define the economic and business environment. The synthesis of these factors into the overall evaluation index is a subjective process that depends on decision makers' system of values, their preferences, and subjective assessment. An overview of

the previous research shows limited efficiency of traditional methods for measuring performance. Referring to the multi-dimensional nature of performance measurement, researchers are expected to have a good theoretical understanding of the nature of performance in terms of the ability to identify measures appropriate to the research context, and to rely on a strong theoretical background in terms of the nature of measures, i.e. what performance are measured and, implicitly, which performance measurement methods to combine in a particular situation and in what way. These findings are consistent with the multi-criteria analysis paradigm, so scientific and professional literature abounds in papers dealing with the issue of evaluating corporate performance.

Thus, Pendaraki and Zopounidis [31] and Verheyden and De Moor [49] developed PROMETHEE II model to evaluate performance of mutual funds. Alptekin [1] evaluated investment and pension funds in Turkey by using TOPSIS method; Chang et al. [7] also apply TOPSIS method for evaluation of performance of mutual funds. Murti et al. [28], [29], measured efficiency of 731 mutual funds, grouped into 7 categories, using at that time still unrecognized DEA approach. They found a significant, positive correlation between their index of efficiency and Jensen alpha index for all categories of assets. Wang et al. [48] identify the evaluation of mutual funds as a sort of fuzzy multi-criteria problem and combine the AHP method with fuzzy methods in the process of determining the relative importance of the criteria. Basso and Funari [3] evaluated performance of 47 mutual funds by using DEA method, showing that DEA method can be more than a useful supplement to traditional approaches to performance measurement. Following the example of 30 private mutual funds, Eling [18] also applied DEA method, indicating its comparative advantages compared to traditional performance measures. In their work, Wu et al. [49] demonstrated the use of the modified DELPHI method, combined with AHP method, in evaluating performance of mutual funds. Wang et al. [47] considered the evaluation of mutual funds as a kind of fuzzy multi-criteria problem, and combined AHP method with fuzzy methods in the process of determining relative importance of the criteria. The efficiency of American mutual funds using the DEA method criteria was measured by Anderson et al. [2] and Daraio and Simar [13]. Galagedera and Silvapulle [19] used DEA methods to assess the relative efficiency of 257 mutual funds in Australia, while Lozano and Gutierrez [26] anlysed relative efficiency of a Spanish mutual fund using six different DEA linear programming models. Murthi & Choi [29] used the same inputs and outputs in the application of DEA method, and performed associated performance measurement based on DEA method with traditional Sarp index. Sengupta [38] finds that 70% of respondent portfolio was relatively efficient, but with significant variations depending on the category of funds. Chen & Li [10] first applied DEA in the evaluation of performance of mutual funds in China, and after them, Ding [15], Deng & Yuan [14], who developed the dynamic DEA model, while Xu & Zhang [51] applied the input-oriented BCC DEA model. Sebastian & Ester [37] in their study also assessed that DEA can be used to evaluate performance of mutual funds, etc. Despite the fact that non-parametric techniques, such as DEA, obviously can be

a useful instrument for measuring performance of mutual funds, the problem is that they only measure relative efficiency and do not allow mutual comparison, which would allow ranking. Still, it could be very useful for investors in the process of optimizing their investment portfolios. Therefore, it is desirable and useful to combine multiple techniques and methods, in order to obtain a comprehensive, objectified, and complete score, which takes into account multi-dimensional nature of mutual fund performance, without neglecting traditional ratio numbers, but on the contrary, relying on them.

3.1. Traditional performance measures of mutual funds – Sharpe index (S_i), Treynor index (T_i), and Jensen's alpha index (α_i)

The base line in the performance measurement of mutual funds is *Capital Asset Pricing Model (CAPM)*, developed, independently from each other by Jack L. Treynor (1961-1962), John Lintner (1965a-1965b), William F. Sharpe (1964), and John Mossin (1966), based on the previous work of Harry Markowitz. According to CAPM, return of mutual fund is a linear function of systemic risk (ß) and selection ability (α), i.e. equals the sum of risk-free return, market premium, and selection ability of managers [27]:

$$R_{i,t} = \alpha_i + R_{f,t} + \beta_i \left(R_{m,t} - R_{f,t} \right) + \varepsilon_{i,t} \tag{1}$$

where:

$R_{i,t}$ – average return of mutual fund i in time t,

α_i – Jensen's alpha index,

$R_{f,t}$– average risk-free return in time t,

β_i– beta coefficient of mutual fund portfolio i,

$R_{m,t}$– average market return in time t,

$\varepsilon_{i,t}$– stochastic specific return of fund i in time t (residual return).

Capital Asset Pricing Model requires that the expected returns of mutual funds are linearly dependent on their covariance with the market [42]. From *CAMP*, basic performance measures are derived: Sharpe index (S_i), Treynor index (T_i), and Jensen's alpha index (α_i). The higher these indices, the more efficient their mutual funds are, i.e. their portfolios, indicating their better performance.

Sharpe index (S_i) is calculated by dividing risk premium, i.e. excess return, by standard deviation of return as a measure of total risk (σ_i :

$$S_i = \frac{R_i - R_f}{\sigma_i} \tag{2}$$

The advantage of using Sharpe ratio in evaluating fund performance is that its calculation does not require benchmark as a substitute for the market. So, the choice of benchmarks does not affect the ranking of funds according to this index, whereas the major drawback of Sharpe ratio lies in the fact that it is a reliable performance indicator only of non-diversified, or poorly diversified portfolio.

On the other hand, Treynor ratio (T_i) is similar to Sharpe ratio (S_i), except that, instead of standard deviation as a measure of volatility of fund returns around their mean values, beta coefficient is used (ß):

$$T_i = \frac{R_i - R_f}{\beta_i} \tag{3}$$

So, Treynor ratio is calculated by dividing the rate of return above the risk-free rate return by beta coefficient as a measure of systemic risk. Beta coefficient measures market exposure of mutual fund, i.e. sensitivity of fund return to the market index [42], and can be represented by the following formula:

$$\beta_i = \frac{\sigma_i \times \rho_{i,t}}{\sigma_m} \tag{4}$$

where:

β_i– beta coefficient of mutual fund portfolio i,

σ_i– standard deviation of mutual fund i,

σ_m– standard deviation of market index,

$\rho_{i,m}$– correlation coefficient of mutual fund i and the market.

Positive beta coefficient means that return of mutual fund is moving in the direction of market return, while negative beta coefficient indicates the contrary. The value of beta coefficient between 0 and 1 indicates the movement weaker than the market, while beta coefficient greater than one testifies to fluctuations more powerful than the market. In calculating beta coefficient, inter alia, a correlation coefficient is used as a measure of the degree to which two series of numbers tend to move together upward or downward. Value of the correlation coefficient ranges from -1 (perfect negative correlation) to +1 (perfect positive correlation), and is determined as follows:

$$\rho_{i,m} = \frac{Cov_{i,t}}{\sigma_i \times \sigma_m} \tag{5}$$

where:

$Cov_{i,t}-$ covariance between mutual fund return and market return.

The conclusion is that higher Sharpe index means higher excess return per unit of total risk as measured by standard deviation, while higher Treynor index means higher excess return per unit of systemic risk as measured by beta coefficient. If portfolio is perfectly diversified, both performance measures, Sharpe and Treynor index, will give the same result because the total risk is equal to the systemic risk. If Treynor index is higher than the Sharpe ratio, it indicates insufficient diversification and the presence of non-systemic risk.

However, although they stand for useful instruments of performance measurement of mutual funds, neither Sharpe nor Treynor ratio show extra return as a result of active portfolio management. That is why Jensen derived alpha index (α_i) from *CAMP* regression equation, which eliminates the aforementioned disadvantages:

$$\alpha_i = R_i - \left[R_{f,t} + \beta_i \left(R_m - R_f\right)\right] \tag{6}$$

Alpha index represents the difference between the fund's actual return and the expected return for a given level of risk. If the actual return of the mutual fund is higher than the expected, alpha index is positive, the fund performance superior, and the manager has managed to achieve extra return and outperform the market, showing their selection ability. However, if the realized rate of return is lower than the expected on the basis of portfolio risk, alpha index is negative, the fund performance inferior, and the mutual fund manager lacks the necessary selection skills. Finally, equality of actual and expected return indicates the average performance of the mutual fund, which is considered the market, and alpha index in this case is equal to zero. It should be added as important that Jensen's alpha must be statistically significant in order to be even considered. If alpha is not statistically significant, mutual fund portfolio has the same performance as the market portfolio. The process of determining the statistical significance (*t-statistic*) is as follows: 1) first, the corresponding hypotheses are formulated, Ho: α =0 and $H_1 : \alpha \neq 0$; 2) then, alpha's standard error is calculated ($S_e(\alpha)$) [43]:

$$S_e(\alpha) = \sqrt{\frac{\frac{\sum \varepsilon_i^2 \times \sum x_i^2}{n-2}}{n \times \sum (x_i - x)^2}} \tag{7}$$

where:

$\varepsilon_i = \left(R_i - R_f\right) - \left(\alpha + \beta_i \left(R_m - R_f\right)\right)$, $x_i = R_m - R_f$;

Finally, alpha index is divided by the calculated standard error, and the resulting value is compared with the corresponding critical value:

$$t(\alpha) = \frac{\alpha}{S_e(\alpha)} \tag{8}$$

The conclusion is that Jensen's alpha in the regression equation measures securities selection skills of mutual fund managers, pointing to their inferiority (α <0) or superiority (α >0). Inferior manager has a significantly negative Jensen's alpha, while the superior manager has a positive and statistically significant alpha index value [23].

3.2. DEAHP approach

Ramanathan [33] proposes a hybrid DEAHP ([25], [21], [42],[22], etc.) method as a way to overcome the shortcomings of the partial application of DEA and AHP methods. AHP (Saaty, [35], [36]) is an intuitive method for formulating and analyzing decisions, where a problem is hierarchicaly structured and pair-wise comparisons are made, based on a 1-9 comparison scale [36]. As a method that can be successfully used to measure relative impact of a number of relevant factors on possible outcomes, as well as for prediction, i.e. distribution of relative probability of outcomes, it has been used for solving a number of complex decision-making problems. A good overview of AHP application was given by Vaidya and Kumar [46], Sipahi and Timor [40]), Ishizaka and Labib [22], and Subramanian and Ramanathan [41].

DEA ([8], [9]; [11]; [12]; [17]) is a mathematical, non-parametric approach for calculating efficiency, based on linear programming, which does not require a specific functional form. It is used to measure performance of decision-making units (DMU) by reducing multiple inputs to a single " virtual" input, and multiple outputs to a single " virtual" output, using weight coefficients, whereby for each organizational unit, the corresponding linear programming model is formed and solved. DEA method has proven to be successful, especially when evaluating performance of non-profit organizations that operate outside the market, because, in their case, financial performance indicators, such as revenue and profit, do not measure efficiency in a satisfactory manner. All data on inputs and outputs for each decision-making unit are entered into a certain linear program, which is actually one of the DEA models. In this way, performance of the observed decision-making units is evaluated, which is the ratio of weighted output sum and weighted input sum. DEA points to relative efficiency because decision-making units are observed and measured in relation to other units. Efficiency ranges from 0 to 1, so any deviation from 1 is attributed to excess inputs or to the lack of outputs.

DEA model is formulated in the form of the following equation:

$$maxh_{j0} = \frac{\sum_{r=1}^{s} u_{rj0} y_{rj0}}{\sum_{i=1}^{m} v_{ij0} x_{ij0}} \tag{9}$$

where:

y_{rj} – Output value

x_{ij} – Input value

u_{rj} - Weight coefficient of output y_{rj}

v_{ij} - Weight coefficient of input x_{ij}

$r = 1, 2, \ldots, s$ - Number of recorded products

$i = 1, 2, \ldots, m$ - Number of used resources

$j = 1, 2, \ldots, n$ – Number of DMU

In DEAHP problem model, DEA method is used for obtaining local decision-making priorities from the comparison matrix in respect of the observed elements in AHP model. Tables 1 and 2 show typical AHP method and DEAHP method comparison matrices, respectively. As Ramanathan suggests, elements a_{ij}, $a_{ij}>0$, $a_{ij}=1/a_{ji}$, $a_{ii}=1$ for each i in AHP comparison matrix become elements of DEAHP comparison matrix, adjusted to DEA method, in order to calculate local priorities. Each matrix row is viewed as a typical DMU, and each column as an output. In addition, matrix contains column with the so-called *dummy*, i.e. fictitious input, which takes a value of 1 for each DMU, to implement DEA method (Tables 3 and 4).

Table 1: Traditional AHP pairwise comparison matrix

	Element 1	Element 2	....	Element n
Element 1	1	a_{12}	...	a_{1N}
Element 2	1/ a_{12}	1		a_{2N}
....	...	...	....	...
Element N	1/ a_{1N}	1/ a_{2N}	...	1

Source: Ramanathan, R. (2006). Data envelopment analysis for weight derivation and aggregation in the analytic hierarchy process, Computers & Operations Research, 33, p. 1296.

Table 2: DEAHP pairwise comparison matrix and assessment of their effectiveness

	Output 1	Output 2	...	Output *n*	Fictitious input
DMU_1	1	a_{12}	...	a_{1N}	1
DMU_2	$1/ a_{12}$	1	...	a_{2N}	1
...	...	...	...	...	...
DMU N	$1/ a_{1N}$	$1/ a_{2N}$	...	1	1

Source: Ramanathan, R. (2006). Data envelopment analysis for weight derivation and aggregation in the analytic hierarchy process, Computers & Operations Research, 33, p. 1296.

Ramanathan proves that DEA method application with AHP comparison matrices provides objectified values of decision-making priority elements, thus reducing subjectivity of assessment using AHP method, and eliminating rank inversion, which occurs by adding or excluding an irrelevant alternative, a characteristic problem when applying AHP. The calculated DEA efficiencies can be interpreted as local priorities of decision-making units. Finally, DEA is used for aggregation of finite decision-making priority elements. When DEA approach is used in this sense, alternatives are seen as decision-making units, DMU, and their local priorities, calculated in relation to each criterion, as outputs, using dummy inputs column. On the other hand, unlike classic DEA approach that measures relative efficiency only, DEAHP method, which implicitly includes the ability of AHP to contain both quantitative and qualitative decision-making factors, results in more complete performance assessment of the observed decision-making units.

Table 3: AHP comparison matrix of alternatives and criteria

	Criterion 1	Criterion 2	....	Criterion *J*
Alternative 1	y_{11}	y_{12}	...	y_{1J}
Alternative 2	Y_{21}	Y_{22}		y_{2J}
....	...	...	....	...
Alternative *N*	Y_{N1}	Y_{N2}	...	y_{NJ}

Source: Ramanathan, R. (2006). Data envelopment analysis for weight derivation and aggregation in the analytic hierarchy process, Computers & Operations Research, 33, p. 1298.

Table 4: DEA approach to evaluating the efficiency of alternatives in relation to the defined criteria

	Criterion 1	Criterion 2	....	Criterion J	Fictitious input
DMU 1	y_{11}	y_{12}	...	y_{1J}	1
DMU 2	Y_{21}	Y_{22}		y_{2J}	1
....	...	...	....	...	1
DMU N	Y_{N1}	Y_{N2}	...	y_{NJ}	1

Source: Ramanathan, R. (2006). Data envelopment analysis for weight derivation and aggregation in the analytic hierarchy process, Computers & Operations Research, 33, p. 1298.

4. DESCRIPTION OF PROBLEMS AND MODELS STRUCTURING

4.1. Traditional approach

We use Belgrade Stock Exchange index, *Belex15* as a benchmark, while average annual rate of return on treasury bills of the National Bank of Serbia is used as risk-free rate of return [24]. Data on return of mutual funds in Serbia, as well as data on average annual rates of return on treasury bills of the National Bank of Serbia was collected from Securities Commission annual reports, while data on Belex15 index trends was taken from the Belgrade Stock Exchange website. It is important to emphasize that management fees were not taken into account in the research since mutual funds publish information on their returns on a gross basis, so that the selection ability of mutual fund portfolio managers will be measured by gross Jensen's alpha.

Table 5: Performance of Mutual Funds in Serbia in the Period 2011-2013

Name of the fund	Sharpe index	ß coefficient	Treynor index	Jensen's alpha	t-statistic
Fima ProActive	-1.672	-0.396	0.472	-0.272	-3.660
Ilirika Cash Dinar	-2.711	0.131	-2.101	-0.231	-3.219
Ilirika Cash Euro	-1.297	-0.251	0.365	-0.133	-2.544
Ilirika Balanced	-3.032	-0.334	0.751	-0.271*	-6.979
Ilirika Dynamic	-3.916	-0.324	0.801	-0.308*	-11.900
KomBank In-Fond	-2.331	0.422	-0.491	-0.153*	-4.919
Triumph	-0.457	-0.787	0.163	-0.278	-1.381
Raiffeisen Cash	-1.394	0.437	-0.241	-0.039	-2.523
Raiffeisen World	-1.352	0.252	-0.122	-0.131	-2.971
Belex15	**-0.620**	**1.000**	**-0.023**	**0.000**	

(*) Indicates statistical significance at the 5% level

Source: PhD Milena Jakšić, Associate Professor. *Scientific area: Financial markets and financial instruments*

Sharpe index, as the first performance measure introduced, which summarizes both benefits and costs of investing, i.e. both return and risk, usually has some value between 0.5 and 3. Based on the " rule of thumb" , if annual Sharpe index is higher than 1.0, fund has a fairly good year, while extraordinary funds have Sharpe index greater than 2.0. In the conducted research, Sharpe ratio is negative for all the observed mutual funds in Serbia, which is to be expected in periods of severe crisis when the goal of active management is not to get more, but to lose less, i.e. to achieve lower negative return. The interpretation of the negative Sharpe index is the same as that of the positive one. In other words, the rule, the higher the index, the better the fund performance, is still valid.

Much more important information than the absolute value of Sharpe index is that this index is for all funds, except for mutual fund *Triumph*, lower than Sharpe index for benchmark *Belex15*, which is -0.620 (Table 5). Therefore, according to Sharpe ratio, eight out of nine analyzed funds have inferior performance compared to the benchmark. However, considering that Sharpe index (S_i is a reliable performance indicator of only non-diversified or poorly diversified portfolio, research must include the calculation of indicators such as Treynor (T_i) and Jensen's alpha index (α_i).

The calculated Treynor ratio is for most funds positive and greater than Treynor ratio for benchmark *Belex15*, which is $T_i = -0.023$. Fund with the highest Treynor ratio – *Ilirika Dynamic* ($T_i = 0.801$) is the fund with the highest excess return per unit of systemic risk, while the largest negative excess return per unit of systemic risk is realized by mutual fund *Ilirika Cash Dinar* ($T_i = -2.101$).

Accordingly, Sharpe index indicates inferior, and Treynor index superior performance of Serbian mutual funds, and Treynor index is for each mutual fund higher than Sharpe index, which is explained by the presence of high non-systemic risk, caused by insufficient portfolio diversification. Furthermore, it should be noted that every possible ranking of funds according to Sharpe and Treynor index would be different, which confirms the conclusion that mutual fund portfolios in Serbia are not well diversified [20]. Regardless of their undeniable usefulness, Sharpe and Treynor indices do not show whether active management helped managers outperform the market, i.e. Belgrade Stock Exchange index, *Belex15*. The answer to this question is given by Jensen's alpha, which must be statistically significant to be taken into account. In the conducted research, alpha indices are negative for all the observed mutual funds in Serbia in the period 2011-2013, while *Ilirika Balanced, Ilirika Dynamic* and *KomBank InFond* funds have a negative and statistically significant value of alpha index. Since the result of the said funds is statistically significant, the research hypothesis H_1 is accepted. Therefore, about 30% of the analyzed mutual funds have inferior performance relative to market portfolio.

In the analyzed period, mutual funds in Serbia lost more value than the market index, which means that active management achieved results worse than the expected. Inferiority of fund performance would be even greater if the management fees were included in the analysis and if net Jensen's alpha was calculated, or, if the analysis included transaction costs. Serbian mutual funds managers lack selection abilities, i.e. the needed action selection skills [20].

4.2. AHP evaluation model of mutual funds

Multi-criteria decision-making techniques, such as Analytic Hierarchy Process, Analytic Network Process, DEMATEL (*DEcision MAking Trial and Evaluation Laboratory*) have extensively been used in evaluating organizational performance both independently and in combination with other multi-criteria or traditional approaches.

The main assumptions underlying the application of AHP evaluation model of mutual funds relate to the following:

The main purpose of the model is performance evaluation of nine selected mutual funds in the Republic of Serbia;

Time period, in which the problem is solved, is exactly limited (three years, i.e. 2011-2013);

The criteria by which a solution to the problem is sought are: 1) Value of mutual funds' assets; 2) Value of investment units of mutual funds; 3) Rate of return per investment unit, and 4) Rate of return on average net assets of the fund (Table 6; *Figure 1*).

Table 6: Data statistics for 2013

Mutual fund	Value of the assets of mutual funds	Value of investment units of mutual funds	Rate of return per investment unit of mutual funds	Rate of return on average net assets of the funds
Fima ProActive	134.836.745	413,23	9,00%	8,84
Ilirika Cash Dinar	230.917.008	1.505,77	9,76%	9,76
Ilirika Cash Euro	24.629.078	1.211,01	2,55%	2,55
Ilirika Balanced	140.034.367	1.449,27	-0,47%	-0,05
Ilirika Dynamic	35.587.397	309,69	8,60%	8,60
KomBank InFond	62.112.576	653,96	5,64%	5,54
Raiffeisen Cash	2.450.654.497	1.544,12	11,50%	11
Raiffeisen World	575.360.438	1.290,24	8,02%	8
Triumph	26.978.215	390,94	-19,60%	-20,73

Having established the appropriate evaluation matrices for the selected criteria in relation to the defined model objective, as well as for alternatives in relation to each individual criterion, using quantitative data from the Belgrade Stock Exchange, the National Bank of Serbia, and data from the annual reports of the Securities Commission of the Republic of Serbia, and using the *Superdecisions* software package, the calculation of the local, and then the synthesis of the final priority criteria and alternatives have been done.

Two questions were asked: First, which criteria, and how much graded, according to the 1-9 scale [35], is considered relatively more important in relation to the purpose of the model? And the second question is: in relation to each individual criterion (observed value of the index), which mutual fund is preferred, or considered better on the 1-9 scale of comparison?

In the case that the process of evaluation and comparison involves more decision makers, it is possible to use geometric mean to combine and objectify evaluation:

$$w_i = \sqrt[k]{\pi_{k=1}^{k=K} w_{ik}}, \ \forall i \tag{10}$$

Where w_i is the final weight of factor i, and the relative weight of element i, calculated on the basis of evaluation of the evaluator k.

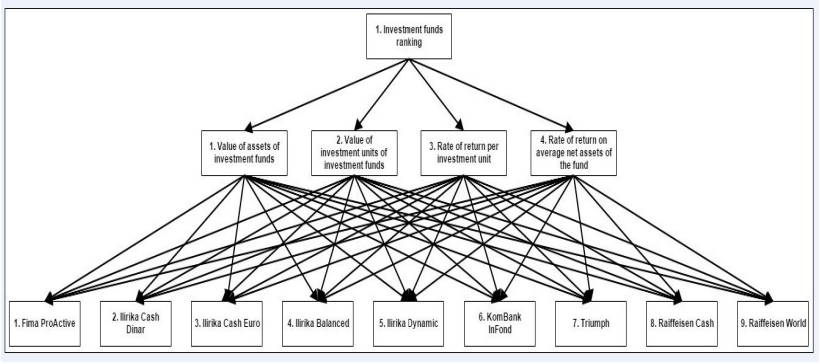

Figure 1 The hierarchical structure of the AHP model for evaluation and ranking of mutual funds in RS. Source: Authors

According to estimates of experts in this field, and based on the 1-9 comparison scale, calculations were made, and the following results were obtained:

Table 7: Relative Importance of the Criteria of the AHP evaluation model of mutual fund performance in the Republic of Serbia, in 2013, Calculated Using the Superdecisions Software Package

Criterion	Relative importance of criteria
Value of assets of mutual funds	0.164523
Value of investment units of mutual funds	0.093934
Rate of return per investment unit	0.433048
Rate of return on average net assets of the fund	0.308946

$CI = 0.0806$

Source: Own calculation

Table 7 shows the calculated relative importance of the selected criteria. The values obtained are the result of expert evaluation and comparison, according to the 1-9 scale. Therefore, the criterion *Rate of return per investment unit* has the highest relative importance (0.433048), followed by the criterion *Rate of return on average net assets of the fund* (0.308946), etc.

Table 8 gives the final priorities of alternatives, i.e. mutual funds in 2013. Tables show that there has been a change in the ranking of individual mutual funds, except for *KomBank Infond*. Furthermore, it can be seen that, in the reporting period, there was significant deterioration in performance and, consequently, the ranking of the mutual fund *Triumph*, while the funds *Raiffeisen Cash* and *Ilirika Cash Dinar* retained their good positions.

Figure 2 presents a comparative overview of benchmark values (*column ideal*) of priorities of mutual funds per year, because it is known that AHP priorities can be

interpreted differently depending on the context of the problem. These values are also the result of authors' and experts' evaluation and comparison based on the 1-9 scale in relation to the defined criteria, i.e. their values during the period. Tables show that there has been a change in the ranking of individual mutual funds, except for *KomBank Infond*. Furthermore, it can be seen that, in the reporting period, there was significant deterioration in performance and, consequently, the ranking of the mutual fund *Triumph*, while *Raiffeisen Cash* and *Ilirika Cash Dinar* funds retained their good positions.

Table 8: Final Priorities and Ranking of mutual Funds in 2013, using the Superdecisions Software Package

Alternatives	Total	Normal	Ideal	Ranking
31 Fima ProActive	0.0603	0.1205	0.3922	3
32 Ilirika Cash Dinar	0.0925	0.1850	0.6018	2
33 Ilirika Cash Euro	0.0232	0.0464	0.1509	8
34 Ilirika Balanced	0.0233	0.0466	0.1515	7
35 Ilirika Dynamic	0.0501	0.1002	0.3259	5
36 KomBank InFond	0.0340	0.0680	0.2213	6
37 Raiffeisen World	0.0544	0.1088	0.3541	4
38 Raiffeisen Cash	0.1537	0.3073	1.0000	1
39 Triumph	0.0086	0.0172	0.0561	9

Source: Own calculation

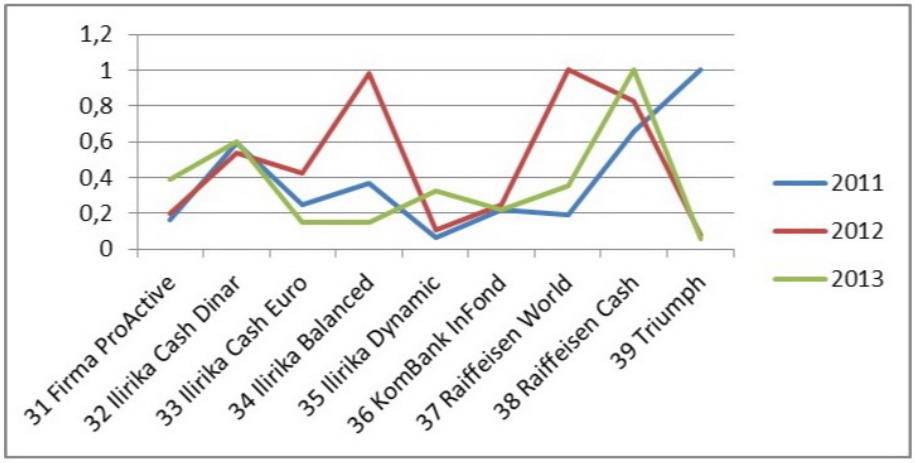

Figure 2 Priorities of mutual funds in the period 2011-2013, calculated using the AHP method (ideal values)

5. DEAHP APPROACH TO EVALUATING PERFORMANCE OF MUTUAL FUNDS IN THE PERIOD 2011-2013

In accordance with Ramanathan's suggestions [33], Table 9 shows DEAHP criteria comparison matrix for the calculation of their relative efficiency. Input values presented in the table are taken from AHP criteria comparison matrix, thus creating conditions for the application of standard input-oriented DEA model. Furthermore, column with the values of fictitious inputs was added as a condition for DEA model application. As a result of the applied DEAHP approach, relative efficiencies of the selected criteria are obtained, and shown in Table 10. Due to the fact that the number of DMU is smaller than the sum of the numbers of outputs and inputs, which is contrary to the condition for a successful application of DEA method, the result shows that three out of four criteria are relatively efficient, while the criterion *value of investment unit* is relatively inefficient. For the purposes of applying the integrated DEAHP method, as already said in the part referring to DEAHP method, it is not necessary to calculate relative efficiency of criteria, so that the combined use of AHP and DEA method continues for alternatives, i.e. mutual funds, viewed as DMU, but now for each criterion separately. Table 11 shows DEAHP comparison matrix of mutual funds in relation to the criterion *value of assets of mutual funds*, and Table 12 relative efficiency of mutual funds compared to the same criteria, for 2013. The values of output in Table 10 stand for the pairwise comparison of alternatives (mutual funds) using AHP method, and typically, column with the values of fictitious input is added as a condition for the establishment and application of DEA method. Relative efficiencies shown in Table 12, calculated using DEAHP method, show that mutual fund *Raiffeisen Cash* is relatively efficient, while other mutual funds are relatively inefficient.

Table 9: Pairwise criteria comparison matrix – DEAHP Method

DMU	Output 1	Output 2	Output 3	Output 4	Dummy input
Value of assets *(C_1)*	1	3	0.333	0.333	1
Value of investment unit *(C_2)*	0.333	1	0.333	0.333	1
Rate of return of investment unit *(C_3)*	3	3	1	2	1
Rate of return on average net assets of the fund *(C_4)*	3	3	0.5	1	1

Source: Own calculation

Table 10: Efficiency of criteria – DEAHP Method

DMU	Efficiency
C_1	1.00000
C_2	0.33333
C_3	1.00000
C_4	1.00000

Source: Own calculation

Table 11: DEAHP alternative comparison matrix in relation to the criterion value of assets of mutual Funds in 2013

DMU	Dummy input	O1	O2	O3	O4	O5	O6	O7	O8	O9
Fima ProActive	1	1	0.25	4	0.5	4	3	0.2	0.143	4
Ilirika Cash Dinar	1	4	1	5	4	5	4	0.25	0.167	5
Ilirika Cash Euro	1	0.25	0.2	1	0.25	0.5	0.25	0.167	0.111	0.5
Ilirika Balanced	1	2	0.25	4	1	4	3	0.25	0.125	4
Ilirika Dynamic	1	0.25	0.2	2	0.25	1	0.333	0.167	0.111	2
KomBankInvFond	1	0.333	0.25	4	0.333	3	1	0.2	0.125	3
Raiffeisen World	1	5	4	6	4	6	5	1	0.167	6
Raiffeisen Cash	1	7	6	9	8	9	8	6	1	9
Triumph	1	0.25	0.2	2	0.25	0.5	0.333	0.167	0.111	1

Source: Own calculation

Table 12: Efficiency of alternatives in terms of the criteria value of assets – DEAHP Method, 2013

DMU	Efficiency
Fima ProActive	0.44444
Ilirika Cash Dinar	0.57143
Ilirika Cash Euro	0.11111
Ilirika Balanced	0.44444
Ilirika Dynamic	0.22222
KomBankInvFond	0.44444
Raiffeisen World	0.71429
Raiffeisen Cash	1.00000
Triumph	0.22222

Source: Own calculation

In the same way, DEAHP comparison matrices of alternatives (mutual funds) were formed in relation to other criteria, and the corresponding relative efficiencies calculated, as shown in Table 13:

Table 13: Local Priorities of DMU (mutual Funds), calculated in relation to the outputs (Criteria)

DMU	Efficiency in relation to the criterion C_1	Efficiency in relation to the criterion C_2	Efficiency in relation to the criterion C_3	Efficiency in relation to the criterion C_4
Fima ProActive	0.44444	0.42857	0.88889	0.88889
Ilirika Cash Dinar	0.57143	1	1	1
Ilirika Cash Euro	0.11111	1	0.66667	0.44444
Ilirika Balanced	0.44444	1	0.55556	0.44444
Ilirika Dynamic	0.22222	0.18182	0.8	0.77778
KomBankInvFond	0.44444	0.375	0.57143	1
Raiffeisen World	0.71429	0.85366	0.77778	0.55556
Raiffeisen Cash	1	1	1	1
Triumph	0.22222	0.28571	0.11111	0.11111

Source: Own calculation

Table 14 shows DEAHP comparison matrix of DMU, i.e. mutual funds, for 2013. Input values are now local priorities of DMU, calculated in relation to each criterion individually using DEAHP method, and column with the values of fictitious inputs is included as well. Final priorities of alternatives, calculated using a hybrid DEAHP method, are shown in Table 15.

Table 14: Comparison matrix of alternatives in relation to the formulated criteria – DEAHP Method for 2013

DMU	I1	O1	O2	O3	O4
Fima ProActive	1	0.44444	0.42857	0.88889	0.88889
Ilirika Cash Dinar	1	0.57143	1	1	1
Ilirika Cash Euro	1	0.11111	1	0.66667	0.44444
Ilirika Balanced	1	0.44444	1	0.55556	0.44444
Ilirika Dynamic	1	0.22222	0.18182	0.8	0.77778
KomBankInvFond	1	0.44444	0.375	0.57143	1
Raiffeisen World	1	0.71429	0.85366	0.77778	0.55556
Raiffeisen Cash	1	1	1	1	1
Triumph	1	0.22222	0.28571	0.11111	0.11111

Source: Own calculation

Table 15: Final priorities of alternatives calculated by using DEAHP method for 2013

DMU	Efficiency
Fima ProActive	0.88889
Ilirika Cash Dinar	1.00000
Ilirika Cash Euro	1.00000
Ilirika Balanced	1.00000
Ilirika Dynamic	0.80000
KomBankInvFond	1.00000
Raiffeisen World	0.85366
Raiffeisen Cash	1.00000
Triumph	0.28571

Source: Own calculation

In the same way, DEAHP method is used to calculate final pirorities of the observed mutual funds in 2011 and 2012. The calculated priorities are shown in Table 15 and then compared with the values obtained by AHP method (Figure 3), as well as through the whole observed period. In doing so, to compare priorities obtained using AHP method and DEAHP method, priorities in the column *Ideal* were used from the AHP, which, as already mentioned, represent a sort of benchmark, and are obtained by dividing all the priorities in the column *normal* by the highest priority in that column. Analysis of the results shows that, depending on the chosen approach, the ranking inversion of individual mutual

funds occurred, as regards 2013, and the same conclusion could be drawn for the remaining period. This is not unexpected, given the different conceptual, logical, and theoretical base underlying some of the observed methods. On the other hand, Table 16 is more illustrative and important from the standpoint of research objectives, as they show DEAHP-based trends in final priorities as regards the observed mutual funds for the reference period 2011-2013. It may be noted that for the majority of mutual funds and their overall performance, 2012 was the least favorable year, 2011 was slightly better, while 2013 was relatively the best, in which, with the exception of the mutual fund *Triumph*, all funds had better overall performance compared to 2012, which is partly consistent with the conclusions reached through partial application of AHP method. Table 17 summarizes the ranking of all mutual funds, from a traditional approach standpoint, which includes three representative indices, as well as an overview of ranking determined by applying AHP and DEAHP methods. It can be concluded that the best overall performance (or the least bad) in the observed period was in the mutual fund *Raiffeisen Cash*, ranked 1 according to the value of the Jensen Alpha Index, AHP, and DEAHP method, while the worst performance was in *Ilirika Dynamic* and *Triumph* mutual funds.

Table 16: Final priorities of mutual funds in the period 2011-2013, calculated using DEAHP method

Mutual fund	2011	2012	2013
Fima ProActive	1	0.72917	0.88889
Ilirika Cash Dinar	1	1	1
Ilirika Cash Euro	0.85714	1	1
Ilirika Balanced	1	1	1
Ilirika Dynamic	0.55556	0.55556	0.8
KomBankInvFond	0.85714	0.77778	1
Raiffeisen World	0.85714	1	0.85366
Raiffeisen Cash	1	1	1
Triumph	1	0.375	0.28571

Source: Own calculation

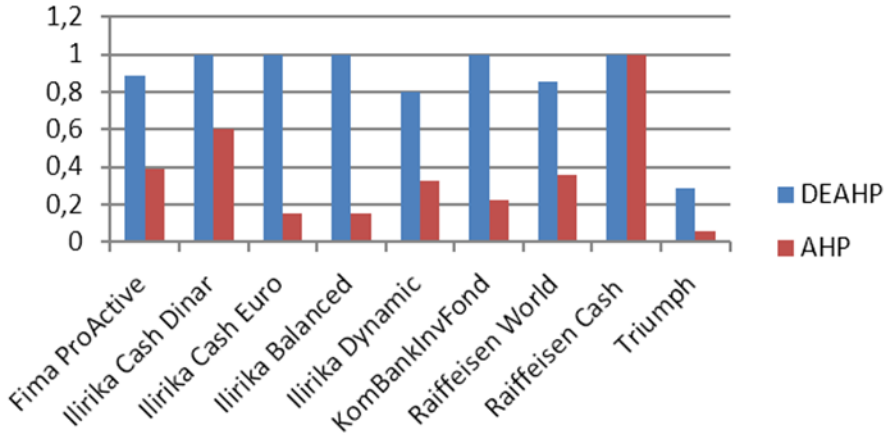

Figure 3 Comparison of the results obtained using AHP and DEAHP methods in 2013

Table 17: Ranking of mutual funds in the period 2011-2013 by approached applied

Mutual fund	Sharpe index rank	Treynor index rank	Jensen's alpha index rank	AHP rank	DEAHP rank
Fima ProActive	5	3	7	3	6
Ilirika Cash Dinar	7	9	5	2	1
Ilirika Cash Euro	2	4	3	8	1
Ilirika Balanced	8	2	6	7	1
Ilirika Dynamic	9	1	9	5	7
KomBank In-Fond	6	8	4	6	1
Triumph	1	6	8	9	9
Raiffeisen Cash	4	7	1	1	1
Raiffeisen World	3	5	2	4	8

6. CONCLUSION

Measurement and evaluation of mutual fund performance has became an integral part of the global financial literature at the beginning of the 1960s, and it has constantly attracted considerable attention of economists. However, in the countries of South East Europe, including Serbia, this issue has become popular forty years later, as evidenced by the small number of papers dealing with this subject.

The existing international studies and the present research have complemented literature in this research area, and allowed for the first research step when domestic literature is taken into consideration. The motive for the conducted research lies in the intention to partially alleviate the chronic shortage of domestic literature on performance measurement and evaluation of mutual funds. So, regarding theory, the work is expected to contribute to the enrichment of literature in this field in Serbia, while, in terms of application, potential contribution of the research is reflected in: a) the application of the presented methods and models in the measurement and evaluation of mutual funds performance by potential investors on the Serbian financial market; b) understanding the significance of portfolio management when considering the interaction between and within the defined clusters; c) the definition of specific recommendations for the evaluation of the overall performance of mutual funds, and d) critical examination of the selected performance measures of mutual funds.

The motive behind the research lies in the intention to make the academic community more familiar with the possibilities of application of non-traditional methods and models in evaluating and ranking performance not only of mutual funds, but of other financial and non-financial institutions, too.

We measured performance of nine mutual funds in Serbia in the period 2011-2013, first by applying traditional performance measures, Sharpe index (S_i), Treynor index (T_i), and Jensen's alpha index (α_i). According to Sharpe index, eight out of nine analyzed funds have inferior performance compared to the benchmark, Treynor ratio is for most funds positive and greater than Treynor ratio for benchmark *Belex15*, while negative alpha indices are recorded in all the examined mutual funds in Serbia. The research results clearly indicate that Serbian mutual fund portfolio has inferior performance relative to the market portfolio, which means that the task of active management is not achieved since active management achieved results worse than the expected. Causes of poor results should be sought in the lack of experience and poor selection skills of Serbian portfolio managers, but also in the fact that, in times of crisis, correlation coefficients tend to one, which makes the benefits of diversification disappear. The obtained results are in line with the conclusion obtained by [25], [7], [20] and others.

After that, AHP model for evaluating the performance of mutual funds was formed, with three levels, four criteria, and nine alternatives. By comparing alternatives, i.e. mutual funds in relation to the selected criteria, based on the 1-9 scale, and using the *Superdecisions* software package, their priorities and rankings

were determined. Finally, DEAHP approach was applied in the calculation of the final priorities of mutual funds, their ranking performed, and the results compared, by years in the reporting period, so as with the results obtained using AHP method. Some mutual funds, such as *Ilirika Cash Dinar, Ilirika Cash Euro* and *Ilirika Balanced* demonstrated superior performance, in the reporting period, compared to other mutual funds while maintaining high level of relative efficiency, but, in general, bearing in mind the traditional indicators and the identified benchmark, showed inferior performance (except *Raiffeisen Cash*).

The multi-criteria analysis has proved in practice as a convenient theoretical and methodological instrument for covering and solving numerous financial decision-making issues, both in companies and in financial institutions. The diverse nature of the factors influencing financial decision-making process, complexity of financial business and economic environment, subjective nature of many financial decisions, present some of the characteristics of financial decisions that enable the application of a multi-criteria methodological framework. The need for simultaneous observation of several criteria, including personal preferences of investors, is an important component of the management function, especially in institutions that professionally deal with money management, such as banks, pension funds, and mutual funds. The reasons for this should be sought in the fact that the structure of their clients' preferences often does not coincide with the preferences of investors, i.e. their financial goals, risk inclination and aversion, investment horizon, etc. The application of multi-criteria decision-making allows decision-makers (manager) to participate actively in the process of making financial decisions, and helps them to understand the complexity and uncertainty of the business environment, and to deal with them.

This means that managers' role is not reduced to passive implementation of the optimal solution (if any!) obtained from the applied multi-criteria model, but they can actively participate in the process of structuring and modeling the problems, as well as in analysing, interpreting, and implementing the obtained results. It can be said that the multi-criteria analysis provides a wide array of techniques for the synthesis of several criteria in the problems of evaluating performance in order to select, rank, classify, and describe a set of alternative options, as evidenced by numerous scientific and professional references.

The proposed approach of combining traditional measures based on ratio numbers, multi-criteria methods, and robust non-parametric methods provides a flexible, systematic, and objective framework for comprehensive performance measurement of mutual funds, and, implicitly, a reliable basis for making quality investment decisions. In theoretical and methodological terms, some doubts remain related to the functioning of DEAHP method in the case of inconsistent evaluation matrices [33], which can, however, be verified or denied in future empirical research. Some efforts have been devoted to overcome this [47]. It would be also useful to carry out solution sensitivity analysis and to check whether and how a change in the relative importance of the selected criteria in AHP method affects the ranking of alternatives. It would also be interesting to compare the results with those obtained by other multi-criteria methods.

In essence, there are certain restrictions related to the conducted study of performance of mutual funds, and the most important one relates to the time dimension of data. Empirical research was, due to the impossibility of collecting sufficiently long time series, conducted on the basis of a relatively small amount of annual data. Therefore, the conclusion of the analysis should be interpreted with caution since the available time series are not long enough to allow for a high degree of reliability of the obtained evaluation. The results of this analysis can be characterized as preliminary research aiming at presenting methodological aspects of a future work with a larger database. Future research will focus on increasing the amount of data by disaggregation, i.e. collection of data for periods of less than one year.

REFERENCES

[1] Alptekin, N., "Performance Evaluation of Turkish Type a Mutual Funds and Pension Stock Funds by Using TOPSIS Method", *International Journal of Economics and Finance*, 1 (2) (2009) 11-22.
[2] Anderson, R., Brockman, C., Giannikos, C., and McLeod, R., "A Non-Parametric Examination of Real Estate Mutual Fund Efficiency", *International Journal of Business and Economics*, 3 (3) (2004) 225-238.
[3] Basso, A., and Funari, S., "Data Envelopment Analysis Approach to Measure the Mutual Fund Performance", *Europen Journal of Operational Research*, 135 (2001) 477-492.
[4] Bialkowski, J., and Otten, R., "Emerging Market Mutual Fund Performance: Evidence for Poland", *The North American Journal of Economics and Finance*, 22 (2) (2011) 118-130.
[5] Bogle, J., "Investing in the 1990s", *Journal of Portfolio Management*, 17 (3) (1991) 5-14.
[6] Chang, E. C., and Lewellen, W. C., "An Arbitrage Pricing Approach to Evaluating Mutual Fund Performance", *Journal of Financial Research*, 8 (1985) 15-30.
[7] Chang, C.H., Lin, J.J., Lin, J.H., and Chiang, M.C., "Domestic open-end equity mutual fund performance evaluation using extended TOPSIS method with different distance approaches", *Expert Systems with Applications*, 3 (2010) 4642–4649.
[8] Charnes A., Cooper, W.W., and Rhodes, E., Measuring the efficiency of decision making units. *European Journal of Operational Research* 2 (1978) 429–444.
[9] Charnes, A., Cooper, W. W., Lewin A. Y., and Seiford L. M., *Data envelopment analysis: theory, methodology, and applications*, (eds.) Kluwer Academic Publisher, Norwell, MA, 1994.
[10] Chen, G., and Li, G. J., "Relative Appraisals of Investment Fund Performance", *Journal ofSichuan University Social Science*, 6 (2001) 32-37.
[11] Cook, W. D., Cooper, W. W., Seiford, L. M., and Tone, K., *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-solver Software*, Kluwer Academic Publishers, MA, USA, 2001.
[12] Cook, W.D., Kaoru, T., and Zhu, J., "Data envelopment analysis: Prior to choosing a model", *Omega* 44 (2014) 1-4.
[13] Daraio, C., and Simar, L., "A robust nonparametric approach to evaluate and explain the performance of mutual funds", *European Journal of Operations Research*, 175 (1) (2006) 516-542.
[14] Deng, C., and Yuan, Q., "Performance of Mutual Funds on Dynamic DEA", *Systems Engineering*, 1 (2007) 111-117.
[15] Ding, W., "Evaluation of Mutual Funds Performance Based on Data Envelopment Analysis Model", *Quantitative and Technical Economics Research*, 3 (2002) 98-101.
[16] Droms, W. G., and Walker, D. A., "Investment Performance of International Mutual Funds", *Journal of Financial Research*, 17 (1994) 1-14.
[17] Dyson, R. G., Allen, R. Camanho, A. S., Podinovski, V. V., Sarrico, C. S., and Shale, E. A., "Pitfalls and protocols in DEA", *European Journal of Operational Research*, 132 (2) (2001) 245-259.
[18] Eling, M., "Performance Measurement of Hedge Funds Using Data Envelopment Analysis", *University of Sent Gallen, Institute of Insurance Economics, Working Papers on Risk Management and Insurance*, 25 (2006) 1-35.

[19] Galagedera, D., and Silvapulle, P., "Australian Mutual Fund Performance Appraisal Using Data Envelopment Analysis", *Managerial Finance*, 28 (9) (2002) 60–73.
[20] Harlow, W., and Brown, K., "The Right Answer to the Wrong Question: Identifying Superior Active Portfolio Management", *Journal of Investment Management*, 4 (4) (2006) 15-40.
[21] Ippolito, R., "Efficiency with costly information: a study of mutual fund performance", *Quarterly Journal of Economics* 104 (1989) 1-23.
[22] Ishizaka, A., and Labib, A., "Review of the Main Developments in the Analytic Hierarchy Process", *Expert Systems with Applications*, 38 (11) (2011) 14336–14345.
[23] Jagric, T., Podobnik, B., Strasek, S., and Jagric, V., "Risk-adjusted Performance of Mutual Funds: Some Tests", *South-Eastern Europe Journal of Economics* 2 (2007) 233-244.
[24] Jakšić, M., Leković, M., and Milanović, M., "Measuring the Performance of Mutual Funds – A Case Study", *Industry*, 43 (1) (2015) 37-51.
[25] Jensen, M. C., "The Performance of Mutual Funds in the Period 1945-1964", *Journal of Finance*, 23 (1967) 389-416.
[26] Lozano, S., and Gutierez, E., "Data envelopment analysis of mutual funds based on second-order stochastic dominance", *European Journal of Operational Research*, 189 (1) (2008) 230-244.
[27] Markovic-Hribernik, T., and Vek, U., "Do Mutual Fund Performance and the Abilities of Fund Managers in Slovenia Deviate from those in Developed Markets? ", *Romanian Economic and Business Review*, 8 (1) (2013) 130-139.
[28] Murti, B. P. S., Choi, J. K., and Desai, P., "Efficiency of Mutual Funds and Portfolio Performance Measurement: A Non-Parametric Approach", *European Journal of Operational Research*, 98 (1997) 408-414.
[29] Murti, B.P.S., and Choi, J. K., "Relative Performance Evaluation of Mutual Funds: A Non-Parametric Approach", *Journal of Business Finance & Accounting*, 28 (7-8) (2003) 853-876.
[30] Otten, R., and Bams, D., "European Mutual Fund Performance", *European Financial Management*, 8 (1) (2002) 75-101.
[31] Pendaraki, K., and Zopounidis, C., "Evaluation of Equity Mutual Funds Performance Using a Multicriteria Methodology", *Operational Research*, 3 (1) (2003) 69-90.
[32] Podobnik, B., Balen, V., Jagric, T., and Kolanovic, M., "Croatian and Slovenian Mutual Funds and Bosnian Investments Funds", *Czech Journal of Economics and Finance*, 57 (3-4) (2007) 159-177.
[33] Ramanathan, R., "Data Envelopment Analysis for Weight Derivation and Aggregation in the Analytic Hierarchy Process", *Computers & Operations Research*, 33 (2006) 1289-1307.
[34] Redman, L. A., Gullet, N. S., and Manakyan, H., "The Performance of Global and International Mutual Funds", *Journal of Financial and Strategic Decisions* 13 (1) (2000) 75-85.
[35] Saaty, Thomas L., *The Analytic Hierarchy Process*, New York, McGraw-Hill, New York, 1980.
[36] Saaty, Thomas L., and Kearns, Kevin P., *Analytical Planning - The Organization of Systems*, Pergamon Press, Pittsburgh, 1985.
[37] Sebastian, L., and Ester, G., "Data envelopment analysis of mutual funds based on second-order stochastic dominance", *European Journal of Operational Research*, 189 (1) (2008) 230-244.
[38] Sengupta, J., "Efficiency tests for mutual fund portfolios", *Applied Financial Economics*, 13 (12) (2003) 869-876.
[39] Sharpe, W. F., "Mutual Fund Performance". *Journal of Business*, 39 (1966) 119-138.
[40] Sipahi, S., and Timor, M., "The Analytic Hierarchy Process and Analytic Network Process: Anoverview of Applications", *Management Decision*, 48 (5) (2010) 775 – 808.
[41] Subramanian, N., and Ramanathan, R., "A Review of Applications of Analytic Hierarchy Process in Operations Management", *International Journal of Production Economics*, 138 (2) (2012) 215–241.
[42] Swinkels, L., and Rzezniczak, P., "Performance evaluation of Polish mutulal fund managers", *International Journal of Emerging Markets* 4 (2009) 26-42.
[43] Teall, L. J., *Financial Market Analytics*, CT: Quorum Books , 1999, ISBN: 1-56720-198-9, 316 pages, Hard Cover HG4515.T43 1999, 332.6'0151 - dc21.
[44] Tangen, S., "Analysing the requirements of performance measurement systems", *Measuring Business Excellence*, 9 (4) (2005) 46-54.
[45] Treynor, J. L., "How to Rate Management of Investment Funds", *Harvard Business Review*, 43 (1965) 63-73.
[46] Vaidya, O. S., and Kumar, S., "Analytic Hierarchy Process: An Overview of Applications", *European Journal of Operational Research*, 169 (1) (2006) 1-29.

[47] Wang, Y.M., and Chin, K.S., "A new data envelopment analysis method for priority determination and group decision making in the analytic hierarchy process", *European Journal of Operational Research*, 195 (2009) 239-250.
[48] Wang, S.-Y., Lee, C.-F., "Fuzzy multi-criteria decision-making for evaluating mutual fund strategies", *Applied Economics*, 43 (24)(2011) 3405-3414.
[49] Verheyden, T. and De Moor, L., "Process-Oriented Social Responsibility Indicator for Mutual Funds: A Multi-Criteria Decision Analysis Approach", *International Journal of Multi-Criteria Decision Making*, 6 (1) (2016) 66-69.
[50] Wu, C-R., Chang, H-Y. and Wu, L-S., "A Framework of Assessable Mutual Fund Performance", *Journal of Modelling in Management*, 3 (2) (2008) 125-139.
[51] Xu, M. and Zhang, B., "Fund Performance Appraisal Using DEA", *Mathematics in Practice and Theory*, 6 (2009) 27-32.