The Quantitative Diversity Index in Multi-Objective Portfolio Model

Seyed Babak Ebrahimi*

*Corresponding Author, Associate Prof., Faculty of Industrial Engineering, K.N. Toosi University of Technology, Tehran, Iran. (Email: B_ebrahimi@kntu.ac.ir)

Mostafa Abdolahi Moghadam

Ph.D. Candidate, Department of Financial Engineering, Faculty of Industrial Engineering, K.N. Toosi University of Technology, Tehran, Iran. (Email: mostafamoghadam68@gmail.com)

Nasser Safaie

Assistant Prof., Faculty of Industrial Engineering, K.N. Toosi University of Technology, Tehran, Iran. (Email: safaie@kntu.ac.ir)

Abstract

The primary purpose of investors is maximizing the utility that is characterized by two essential criteria include risk and return. Regarding investors' uncertainty about the future, one of the main ways to reduce risk is to diversify the investment portfolio. In this research, we proposed an index conducted by Euclidean distance for assessing portfolio diversity. Besides, we designed a multi-objective model to select optimal stock portfolios with considering value at risk (VaR), which is one of the critical indicators of unacceptable risk, portfolio Beta as systematic risk, and portfolio variance as unsystematic risk simultaneously. The model presented in this paper aims to maximize diversification while minimizing value at risk and stock risks. Furthermore, maximizing returns are considered as a limitation of this model. Since the proposed model is nonlinear and concerning computational complexity, it is NP-hard; therefore, we utilized the PSO and the GE metaheuristic algorithms that are improved for solving multi-objective problems to solve the model. The results of the model implementation in multiple iterations showed that the average yield of selected portfolios by the model is higher than the desirable condition. The evaluation of stock performance indicators also shows the satisfactory performance of the multi-objective model.

Keywords: Systematic/Unsystematic risks, VaR, Diversity index, Portfolio optimization.

JEL: C31, C52, D46, D81, G11

Introduction

The current literature highlights the importance of choosing the optimal set of investments in the capital market to maximize the expected wealth of investors. In doing so, an investor needs proper methods or criteria to identify and measure the potential value of each investment opportunity. These criteria should be sufficiently reliable and accurate so that investors can decide with high confidence and low risk. Risk and returns are two main critical factors in capital market decisions. The selection of a set of stocks, called portfolio, is usually driven by the interaction between risk and return. The higher the risk of an investment, demands a higher return (Jones, 2010).

Since the early 1950s of the inauguration of modern portfolio theory, the rate of return and the risk of a portfolio has been recognized as the most important factors for investors in every capital market. Markowitz theory (1952) proved that the risk and returns could be at an optimal point by investing in a diversified portfolio of financial assets. The degree of risk-taking among individuals is a tradeoff between risk and expected return, and asset returns are unpredictable or risky. Diversification means choosing different financial assets to reduce the risk of one specific asset. A diversity index is a mathematical measure that can show how much the initial wealth is distributed between different assets. In other words, the diversity index contributes investors to choose the appropriate number of assets for investing based on their initial wealth. Diversity has many advantages including aggregate competition in the capital market, maximizing investor wealth and mitigates portfolio risk (Chan, Peter, and William, 1989). According to the stock portfolio theory, portfolio risk is not only affected by the average standard deviation, but also by the diversity of investment. In other words, the larger the variety in an investment portfolio, the lower would be the risk (Reilly and Keith, 2002). This goal requires that the variability of the return on a particular asset be adjusted to the variability of the return on other assets in the portfolio, which would reduce the unsystematic risk (Platanakis, Athanasios, and Charles, 2018; Jackwerth, and Anna, 2016)

This research seeks to provide an alternative model to select the optimal stock portfolio and a useful tool to estimate the degree of diversification by adjusting risk win returns. VaR, which is one of the significant indicators of undesirable risk, is integrated with systematic and unsystematic risks when forming a portfolio. We also present the Euclidean distance measure for stock portfolio diversification and formulate a multi-objective model to choose optimal stock portfolios. The Euclidean distance is a novel measure to index

the level of diversity in the portfolio, this index calculates the distance between different assets and optimizes the diversification. The results of the model are compared with Sterling diversity index which is a well-known index that integrated variety and balance into a dual concept that can explicit the condition of different parts of society. The model attempts to maximize diversification by minimizing the VaR and stock risk.

Moreover, maximizing returns are considered to be a constraint of this model. A genetic algorithm is used to optimize the nonlinear multi-objective model and compare the results of different dimensions to validate the framework. The results display that the average yield of selected portfolios by the model is higher than the desirable condition and confirm the positive performance of the multi-objective model.

According to investment theory, financial asset's risks can be classified into systematic risk and unsystematic risk. An unsystematic risk, referred to a controllable risk, is exclusive to an asset because the risk is related to a portion of the return on an asset. This amount of risk is specific to a company or an industry, and it is due to several factors such as worker strikes, management practices, advertising competition, and changes in consumer tastes. Systematic risk, an uncontrollable risk, is related to the general market conditions such as interest rates, the national currency rate fluctuations, inflation rates, monetary and financial policies, and political conditions (Gagliardini, and Christian, 2013) Systematic risks cannot be eliminated at all (Kim et al. 2018).

Markowitz (1952) suggests that the risk and returns could be at an optimal point by investing in a diversified portfolio of financial assets. He created two-directional reforms in the management of investment, with the idea that a financial decision to be taken from the swap between risk and the return of the stock market. First, he assumed that the investor performs a quantitative evaluation of the risk and return of the stock portfolio at the same time pays attention to integrate the portfolio return and the motion of the portfolio returns, which it is the main idea of diversification. Second, the financial decision-making process assumes to be an optimization problem; the investor chooses a portfolio among the various types of available combinations which has the least variance (Georgalos, Ivan, and David, 2018). Markowitz approach is a diversification method that is used in the analysis of the portfolio of investments. This kind of diversification involves the inclusion of covariance between the securities and the combination of less correlated capital assets to reduce the risk in the portfolio without jeopardizing returns. In other words, the less correlation in an investment portfolio will reduce the risk of the portfolio (Francis, and Dongcheol, 2013).

Other methods introduced include the use of the Hirschman-Herfindahl Index (HHI) and the Shannon entropy index (Chen, Yong, Xianhua, and Lingling, 2014). Oh, et al. (2005) used a genetic algorithm to optimize the index-based portfolio. Their goal was to build a portfolio that had the same performance as the stock index. The proposed algorithm was applied to one of the Korean stock market indices from 1999 to 2001 and was compared with traditional methods of constructing the index-based basket. The results show that the genetic algorithm has many advantages over traditional methods when the fluctuations of the market are increasing. It is fully effective and shows the average performance when the market trend is constant.

Research background

This study contributes to the literature as follows: First, we develop an innovative multi-objective mathematical model by combining VaR and portfolio diversification as a task to optimize portfolio, stock returns, and risks. Second, we consider both systematic and unsystematic risks in the model formulation using the Euclidean distance criterion as a tool for the quantitative assessment of the diversity index. Third, we apply a real-world case study to verify the proposed model and solve the proposed model using two robust meta-heuristic algorithms. Fourth, the model can be tested in any other capital market in the world.

Applying the diversification strategy to portfolio optimization is considered by many scholars and organizations (Steinberg, 2018; Pola, 2016; Briere, Kim, and Ariane, 2015; Dang, 2019; Kara, Ayşe and Gerhard-Wilhelm, 2019; Paut, Rodolphe, and Marc, 2019; Aluko, Oladapo, and Bolanle, 2018; Beaudreau, Maggie, and Philip, 2018). For instance, Liu (2018) used experimental data of several key cryptocurrencies to study the role of diversity and investment in the digital asset market. Similarly, Kajtazi et al. (2018) searched the effects of considering bitcoin to the ideal portfolio using the mean-CVaR method. They reported that this consideration could play a significant role in portfolio diversification.

Diyarbakırlıoğlu and Satman (2013) offered a new approach to evaluate the diversification risk of an investment portfolio by the covariance matrix of returns. They solved their problem using Maximum Diversification Index (MDI) through the genetic algorithm. Their problem is verified by existed stock returns data, and results show the MDI can be powerfully applied to define a large set of investable assets. Oyenubi (2016) reported an acceptable description for the elusiveness of the optimum amount of stocks in a portfolio. He used the Portfolio Diversification Index (PDI) to quantify diversification.

Having used a novel quantification approach, Oyenubi attempted to quantify the level of the interdependency of stocks in a set, and the problem is solved by using Pareto based algorithm to find optimal portfolios. Jadhao and Chandra (2017) utilized sample entropy and approximate entropy indicators for diversifying portfolio in the rotation strategy based on size and style. Oloko (2018) investigated the benefits of diversification in Nigeria's stock market. Kalashnikov et al. (2017) suggested a new integrated approach to address the Lean Six Sigma project portfolio and solved the problem using a binary mixed-integer bi-objective quadratic model. Their model is solved via branch-and-bound solver of CPLEX software, and it is verified using numerical examples.

Chang et al. (2009) examined the optimization of the portfolio considering different scales for risk measurement using a genetic algorithm. In their research, the genetic algorithm was used because of its ability to solve complex problems in various risk assessments. The results showed that most of the optimization problems, including cardinal limitations, can be solved through a genetic algorithm within a reasonable time; if the mean-variance, semi-variance, and variance associated with skewness are used as a risk measurement criterion. He also found that the smaller portfolio has better performance than a larger one. The stock returns derived from the genetic algorithm are less than other models, but risk reduction and adjusted risk-based criteria offset the reduction in returns implying the superiority of the response from the genetic algorithm.

Investors should consider the following rules when launching their diversification strategy. First, the whole investment portfolio in one stock does not have a significant impact on the overall strategy. Second, targeting a maximum of 20 stocks in the portfolio would result in better control of the managerial costs. Third, investing more than 10% in one stock is not recommended because it can be against diversification. Forth, to achieve optimal diversity, the focus on investing in companies that significantly are affected by others should be avoided; because they are affiliated companies or large suppliers or customers. Investing in a large company that is affected by others, may prevent diversification.

Mathematical Model

Assumptions

First, investors are generally risk-averse and more interested in higher expected returns per less risk, yet they seek to balance between risk and returns. As a result, the stock portfolio is selected based on minimizing the amount of fitness

function. Second, investors do not tend to invest in a small stock that has low liquidity, and there are no restrictions on transaction costs and taxes. Third, there are no restrictions on the market and the short sale. A mathematical model to consider portfolio diversity, stock returns, and risks simultaneously are developed from a portfolio analysis viewpoint. To this end, at first, we provide the assumptions, indices, parameters, and decision variables of the proposed model and then using these relevant components, we formulate the mathematical model.

Indices

$j=1,2,...,n$　Index of stocks of active investor companies
$t=1,2,...,T$　Time period index

Parameters

r_{jt}	The returns of stock j in the period t
$\hat{r}_j$	The average return on stock j
σ_j^2	The variance of stock j
σ_{ij}	Covariance between stock i and j
d_{ij}	Euclidean distance between two selected stocks i and j
n	Total number of stocks in the portfolio
T	Number of time periods past
k	Number of related characteristics of each stock
β_j	Systematic risk of stock j
S_{jk}	The value of characteristic k from the stock j
M	A big positive number
E_Ω^{max}	Maximum possible expected returns of the stock portfolio
α	Desired confidence level
U	The maximum investment in stocks
λ	Percentage of the minimum expected return of the portfolio
$MaxNp$	The maximum number of stocks to be selected for the investment portfolio

Decision variables

x_j	Percentage of stock j in the investment portfolio
q_j	Equal to 1, if the stock j is in the investment portfolio; otherwise 0
y_t	Equal to 1, if the portfolio returns in time t are negative, otherwise 0
VaR	VaR
SR	The systematic risk value of the investment portfolio
USR	The unsystematic risk value of the investment portfolio

Multi-objective model

After describing the relevant components, the proposed multi-objective model can be formulated as follow. This model has three main objective functions that attempt to maximize stock diversification and minimize VaR and stock risks. Three mentioned objective functions along with its relevant constraints shows as follow:

Objective functions

$$Max: Z_1 = \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} \times x_i \times x_j \qquad i \neq j \qquad (1)$$

$$Min: Z_2 = USR + SR \qquad (2)$$

$$Min: Z_3 = VaR \qquad (3)$$

Subject to

$$USR = \sum_{j=1}^{n} x_j^2 \times \sigma_j^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} \times x_i \times x_j \qquad i \neq j \qquad (4)$$

$$SR = \sum_{j=1}^{n} x_j \times \beta_j \qquad (5)$$

$$\sum_{j=1}^{n} x_j = 1 \qquad (6)$$

$$\sum_{j=1}^{n} q_j = MaxNp \qquad (7)$$

$$x_j \leq U \times q_j \qquad \forall j \in n \qquad (8)$$

$$VaR + \sum_{j=1}^{n} r_{jt} \times x_j \geq -M \times (1 - y_t) \qquad \forall t \in T \qquad (9)$$

$$\sum_{t=1}^{T} y_t = [\alpha T] \qquad (10)$$

$$\sum_{j=1}^{n} \overline{r}_j \times x_j \geq (1 - \lambda) \times E_{\Omega}^{max} \qquad (11)$$

$$E_{\Omega}^{max} = \underset{j}{Max} \{ \overline{r}_j \times x_j \} \qquad (12)$$

$$y_t \in \{0,1\} \qquad \forall t \in T \qquad (13)$$

$$x_j \geq 0 \qquad \forall j \in n \qquad (14)$$

$$q_j \in \{0,1\} \qquad \forall j \in n \tag{15}$$

where,

$$d_{ij} = \sqrt{\sum_{k=1}^{K} \left(s_{ik} - s_{jk} \right)^2} \quad \forall j, i \in n \tag{16}$$

$$d_{ij} = d_{ji} \qquad \forall j, i \in n \tag{17}$$

$$\begin{cases} d_{ij} = 0, & \forall j = i \\ d_{ij} > 0, & \forall j \neq i \end{cases} \tag{18}$$

$$\beta_j = \frac{\operatorname{cov}\left(\overline{r}_j \times \overline{r}_m \right)}{\sigma_n^2} \tag{19}$$

The objective function (1) is to maximize the diversity of stock portfolio. Since the distance is an essential component to measure the difference between elements, then this criterion should play a fundamental role in the classification of investments. Therefore, variety can be calculated using the difference between the elements. Accordingly, a portfolio of stocks as N={S1, S2, …, Sn} is considered, and the Euclidean distance between the two selected stocks *i* and *j* is defined using equations (16) to (18). Finally, the integrated multi-criteria diversity index based on the Euclidean distance between each pair of elements can be introduced as D=∑ni=1 ∑nj=1 dij xixj. In this regard, increasing the value of *D* means increasing the variation between items. Then, the variety of elective elements is estimated as total Euclidean distances between each pair of them.

On the other hands, the objective function (2) is to minimize the systematic risk and unsystematic risk. Different indices have been proposed for risk measurement of stock portfolios, including the portfolio standard deviation introduced by Markowitz (1952). In order to calculate the variance of the stock portfolio, the weight of each share must be determined (x_j) in the stock portfolio. The stock variance is calculated as equation (4) which represents the unsystematic risk. Also, equation (5) measure the systematic risk of the stock portfolio to calculate parameter β_j the equation (19) is used. The objective function (3) is to minimize VaR which it represents the risk of loss for investments and estimates how much of the investment may be lost within a given time period when the market condition is stable.

Equation (6) implies that the purchased stocks must be precisely the same as all available resources. Constraints (7) shows that the selected stocks in an investment portfolio should be less than or equal to the appropriate maximum

amount. Constraints (8) specified an upper bound for the weight of each share (x_j) in the stock portfolio if this stock is selected, it can increase the diversification of the stock portfolio. In determining the upper bound for the decision variables, the investor's opinion is decisive, and it is determined by the minimum number of shares which the investor is willing to invest in them. Constraints (9) shows that the portfolio returns plus the VaR is related to the y_t. Equation (10) shows that in α percentage of time the variable y_t is equal to one. When this variable is equal to one, the corresponding constraint in the first constraints category is redundant, in other words, the portfolio returns in α percent of the time or in [αT] of the time period can be negative. For the rest of the time, y_t is equal to zero; it means the portfolio returns plus the VaR in periods that portfolio return is negative should be greater than zero. In constraints (11) $\hat{r}_j$ represents the average return of each stock and E_Ω^{max} is the maximum expected return of the stock's portfolio and the expression (1-λ) E_Ω^{max} indicates the minimum expected return of the portfolio. Given that the portfolio return is assumed to be the weighted sum of the return, the value of E_Ω^{max} can be calculated using the equation (12). Finally, constraints (13)-(15) show the binary or non-negative decision variables.

Solution approach

To solve the proposed multi-objective model, two well-known meta-heuristics algorithms including a genetic algorithm (GA), and particle swarm optimization (PSO) are used. Moreover, Lingo software is used to obtain the ideal values of each objective functions.

Genetic Algorithm (GA)

GA is a highly effective and efficient random and meta-heuristic optimization method that has been used to solve many complex problems developed by Holland (1974). In this algorithm, first, the problem variables are chosen randomly; then they are combined to draw other points. GA as one of these algorithms is basically a computer search method composed of the gene and chromosome structures.

This algorithm initially begins with a set of random solutions (chromosomes) which is known as the population base, and then the value of each chromosome is determined according to the fitness function. Therefore, higher qualities of chromosomes have a greater chance of producing offspring; on this basis, the choice of parents is taken, and then the offspring are created by crossover operator on the parents. Finally, some of the genes of the offspring change with the mutation process, and then the new offspring are

replaced with the weakest chromosomes in the initial population. The main steps to solve an optimization problem via GA can be illustrated as Figure. 1 Also, the mutation and crossover operators are presented in (Holland, 1974), and the reader advised to see the research.

Step 1. Create the initial population;

a) Generate random chromosomes.
Step 2. Perform the main loop;

b) Calculate the fitness functions of each chromosome.
c) Select two chromosomes from the initial population using the roulette-wheel operator.
d) Do a crossover operator.
e) Do mutation operator.
f) Repeat steps *b* to *e* until enough members to form the next generation be created.
Step 3. Repeat step 2 while stopping criteria is not satisfied.
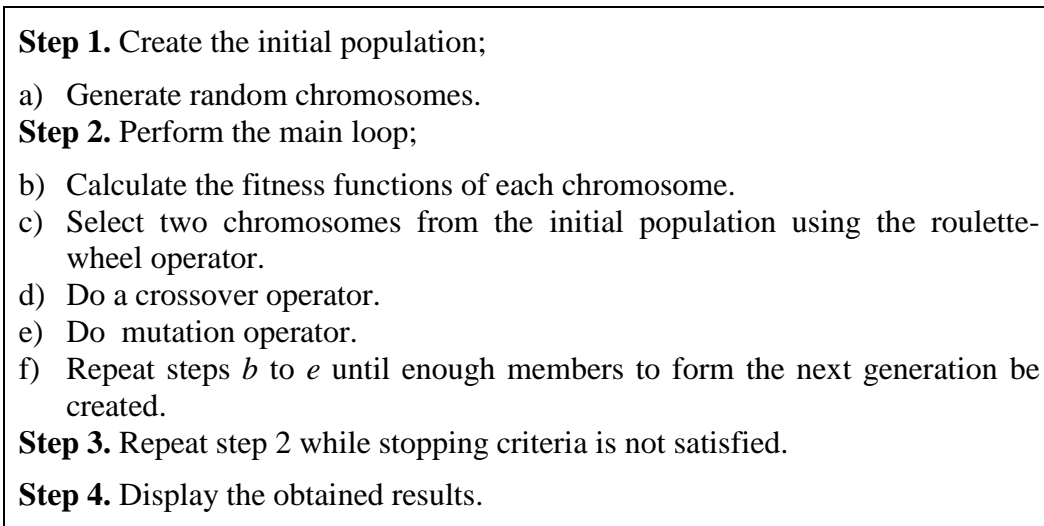
Step 4. Display the obtained results.

Figure 1. The pseudo-code of the genetic algorithm

In the face of multi-objective optimization problems, one of the approaches to solving is to use an LP-metric approach which was introduced by Zelany (1974). This method is one of the compromise programming methods, and it works without achieving knowledge from the decision-maker, and it attempts to minimize the distance between some of the reference points and the probable solution (deviation). In this method, the choice of the reference point and the criterion for measuring the distance is an important topic.

Based on this model, each objective of the problem (*k* objectives) is solved separately, and after normalizing, the answers are combined to find the optimal solution (the optimal answer that is the closest answer to the ideal answer). The mathematical form of this method can be displayed as follows.

$$Minimize \quad Z = \sum_{l=1}^{k} \left(\frac{Z_l^* - Z_l}{Z_l^*} \right)^p \tag{20}$$

p=1 refers to the same weight of all deviations, and an increase of *p* means more weight of larger deviations. Now, to form a fitness function (goal) for using in the genetic algorithm, the LP-metric approach is used according to

the following form:

$$Minimize \quad Z = \frac{Z_1^* - Z_1}{Z_1^*} + \frac{Z_2^* - Z_2}{Z_2^*} + \frac{Z_3^* - Z_3}{Z_3^*} \tag{21}$$

Where Z_1 is the objective function of the greatest diversity (diversification), Z_2 is the objective function of total risks includes systematic and unsystematic risks, and Z_3 is the objective of the VaR. Also, Z_1^*, Z_2^*, and Z_3^* are the optimal amount of each of the objectives in terms of the problem constraints.

Particle Swarm Optimization

Eberhart and Kennedy (1995) proposed a particle swarm optimization algorithm which is inspired by the social behavior of animals, such as the collective migration of birds and fishes. Initially, this algorithm was used to explore the effective patterns on the simultaneous flying of birds, the sudden change of direction, and the optimal deformation of their groups. The change in the location of birds in the search space is influenced by the experience and knowledge of themselves and their neighbors. Therefore, the position of other birds affects the search of a bird. The result of the modeling of this social behavior is the process of searching for birds in the direction of successful areas. Birds learn from each other and move on to their best neighbors based on their knowledge. The basis of this algorithm is on the principle that at any given moment, each bird adjusts its location in the search space, according to the best place ever located and the best place in its entire neighborhood. The following relationships are also used to update the velocity and location of each of the particles.

$$V_i(t) = w \times V_i(t-1) + c_1 rand_1 \times (P_{i.best} - x_i(t-1)) + c_2 rand_2 \times (P_{g.best} - x_i(t-1)) \tag{22}$$

$$x_i(t) = x_i(t-1) + V_i(t) \tag{23}$$

Where w is the inertial weighting factor or moving in its path, which indicates the effect of the velocity vector of the previous iteration ($V_i(t-1)$) on the velocity vector in the current iteration ($V_i(t)$). Also, c_1 and c_2 represent the constant coefficient of training or motion in the direction of the best value of the examined particle and the best value among all population, respectively. Moreover, $rand_1$ and $rand_2$ are two random numbers with uniform distribution in (0, 1). The $x_i(t)$ and $x_i(t-1)$ represent the position vector of particles in the current iteration and previous iteration, respectively.

The best position found for particle i is defined by $P_{i\ .best}$ while $P_{g\ .best}$

represents the best position found by the best particle among all particles.

To prevent the excessive movement speed of a particle when they move from one location to another or velocity vector divergence, the velocity variations are limited to ($V_{min} \leq V_i(t) \leq V_{max}$). The pseudo-code of the particle swarm optimization algorithm is presented in Figure. 2.

```
Initialize particle
For each particle
Calculate fitness value of the particle fᵢ
/*updating particle's best fitness value so far*/
If fᵢ is better than Pᵢ.best
Set current value as the new Pᵢ.best
End For
/*updating population's best fitness value so far*/
Set Pg.best to the best fitness value of all particles
For each particle
Calculate particle velocity according to equation (22)
Update particle position according to equation (23)
Calculate the fitness value of the particle fᵢ
/*Updating population's best fitness value so far*/
Set Pg.best to the best fitness value of all particles
End For
End
```

Figure 2. The pseudo code of particle swarm optimization algorithm

Also, the chromosome used for these two algorithms is shown in Figure. 3 By way of example, we suppose that eight stocks ($j=5$) exist and *MaxNS*=5. So, activate genes of the chromosome is less than or equal to *MaxNS* (see constraint (7). For instance, it supposed with 5 includes gens 1, 3, 4, 6, and 7. The initial chromosome with random data between (0, 1) are generated which sum of them be equal to 1 (based on the constraint (6) of the model). So, the proposed chromosome represents the percentage of stock j in the investment portfolio (x_j).

0.23	0	0.17	0.11	0	0.14	0.35	0

Figure 3. The example of proposed chromosome used by algorithms

Parameter setting

Here, the parameters of algorithms are tuned to achieve the best performance, to this end Taguchi method are used (Taguchi, 1986). This method as one of

the designs of experiments approaches seeks to tune parameters using a set of the orthogonal array instead of full factorial experiments. To perform this method, Minitab software is utilized, and since the LP-metric approach is used to combine objective functions, so following response as "smaller is the better" for Signal-to-Noise ratio is used.

$$S/N = -10 \times \log(\Sigma(Y^2)/n)) \qquad (24)$$

To do this at first, the levels of the parameters of each algorithm are provided in Table 1. Then, by performing the Taguchi method using Minitab software, the orthogonal arrays L^9 and L^{27} are chosen for GA and PSO respectively. Finally, these two algorithms are ran using these sets of experiments and the results are presented in Tables 2-3. Furthermore, to decide on these results the Signal-to-Noise plots are illustrated in Figure. 5-6.

Table 1. The parameters levels of each algorithm

Algorithms	Parameters	Parameters level		
		Level 1	Level 2	Level 3
PSO	C_1	0.5	1	2
	C_2	0.5	1	2
	W	0.5	0.75	1
	$N\text{-}pop$	100	200	300
	$Max\ iteration$	200	400	600
GA	Pc	0.7	0.8	0.9
	Pm	0.05	0.10	0.15
	$N\text{-}pop$	100	200	300
	$Max\ iteration$	200	400	600

Table 2. The orthogonal array L^9 and results for GA

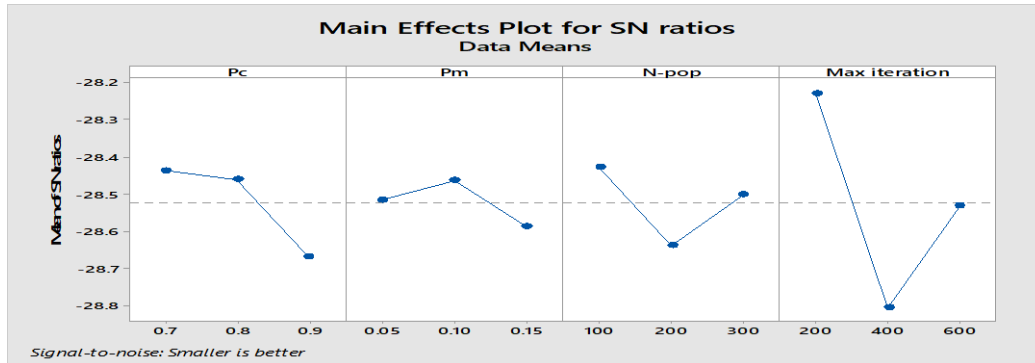| Pc	Pm	$N\text{-}pop$	$Max\ iteration$	Response
0.7	0.05	100	200	25.23688
0.7	0.1	200	400	27.47174
0.7	0.15	300	600	26.57779
0.8	0.05	200	600	26.85483
0.8	0.1	300	200	25.36971
0.8	0.15	100	400	27.27282
0.9	0.05	300	400	27.93733
0.9	0.1	100	600	26.68067
0.9	0.15	200	200	26.77781

Figure 4. The Signal-to-Noise plots for GA

Table 3. The orthogonal array L^{27} and results for PSO

| C_1	C_2	W	N-pop	Max iteration	Response
0.5	0.5	0.5	100	200	26.51833
0.5	0.5	0.5	100	400	25.59528
0.5	0.5	0.5	100	600	28.93421
0.5	1	0.75	200	200	27.42218
0.5	1	0.75	200	400	28.92652
0.5	1	0.75	200	600	25.59709
0.5	2	1	300	200	26.7662
0.5	2	1	300	400	26.97564
0.5	2	1	300	600	25.57436
1	0.5	0.75	300	200	26.71327
1	0.5	0.75	300	400	26.30941
1	0.5	0.75	300	600	25.34503
1	1	1	100	200	27.99651
1	1	1	100	400	26.18041
1	1	1	100	600	25.10245
1	2	0.5	200	200	26.97777
1	2	0.5	200	400	26.78379
1	2	0.5	200	600	26.02241
2	0.5	1	200	200	26.20638
2	0.5	1	200	400	25.45439
2	0.5	1	200	600	27.79816
2	1	0.5	300	200	26.10074
2	1	0.5	300	400	27.69621
2	1	0.5	300	600	27.77765
2	2	0.75	100	200	25.41882
2	2	0.75	100	400	28.62868
2	2	0.75	100	600	28.57616
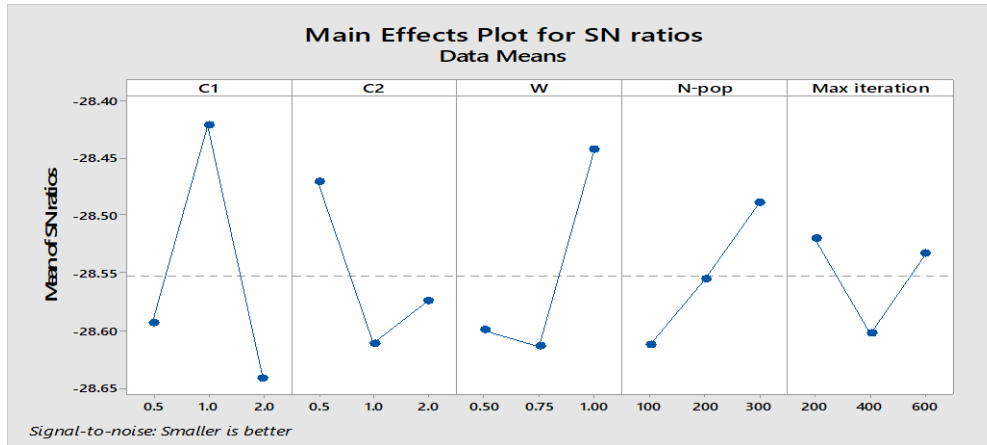
Figure 5. The Signal-to-Noise plots for PSO

Finally, based on Figure. 4 it can be found that the values of 0.7, 0.1, 100, and 200 are selected for *Pc*, *Pm*, *N-pop*, and *Max iteration*, respectively. Likewise, Figure. 5 shows that the values of 1, 0.5, 1, 300, and 200 are selected for C_1, C_2, *W*, *N-pop*, and *Max iteration*, respectively. Therefore, these values set for final running via two algorithms.

The Model Test

To implement the model and measure its efficiency, we performed a test of data from the top 30 companies in Tehran Stock Exchange (TSE) by considering several features such as volume, value, no. of trades, closing price, and market capitalization are reported in Table 4. These mentioned values are used as S_{jk} (characteristic k from the stock j) to calculate d_{ij} (Euclidean distance between two selected stocks i and j).

Table 4. The data of several characteristics of TSE

No. of stock	Stock	Volume	Value	No. of trades	Closing price	Market Cap.
1	PSER1	85129	1346570522	30	15818	949080000000
2	SROD1	3419185	6250270180	137	1828	1480680000000
3	GHND1	213094	2458456478	105	11537	2322400000000
4	LAMI1	85005	822253365	4	9673	580380000000
5	STEH1	703972	1865525800	64	2650	4637500000000
6	SKHS1	53487	576054990	10	10770	1346250000000
7	LTOS1	6935364	16499230956	3926	2379	11895000000000
8	KFAN1	107918	597541966	12	5537	1107400000000
9	KPRS1	134324	741468480	25	5520	687820000000

10	TAMI1	13607192	75180532959	949	5525	895050000000
11	KSKA1	255239	866791644	69	3396	849000000000
12	SDAB1	16242679	30337442919	1240	1868	1681200000000
13	SGRB1	8707993	35504401399	662	4077	2038500000000
14	SHGN1	1098484	3337877588	173	3039	1528310000000
15	SSEP1	13316758	16381365367	680	1230	3013500000000
16	BSTE1	5100963	5629783797	345	1104	6624000000000
17	SRMA1	1221949	4292722164	143	3513	519924000000
18	DMVN1	207080	4753109980	144	22953	688590000000
19	CIDC1	2972619	4214050974	249	1418	6806400000000
20	SPKH1	3589819	25670837173	457	7151	1430200000000
21	SHSI1	183910	3213792791	124	17475	536701000000
22	SMAZ1	2698947	5347058509	210	1981	2442570000000
23	SURO1	5125563	18111440764	598	3534	2473800000000
24	SKOR1	3397425	4126797445	217	1215	2673000000000
25	KHOC1	5552238	24675514218	698	4444	2888600000000
26	SIMS1	14952728	24632102295	892	1647	1770640000000
27	ABAD1	7113800	15744651573	855	2213	531120000000
28	SFKZ1	41145366	74593472749	1523	1813	10198100000000
29	SSOF1	983403	2694829420	153	2740	1507000000000
30	NSTH1	11588115	31572081897	847	2725	2043750000000

Also, the returns values of stocks in each time period are provided in Table 5 and the values of average returns, β coefficients, and variance, per 12 months are given in Table 6. Furthermore, the values of d_{ij} and σ_{ij} can be obtained using their related formulas and the values of desired confidence level (α), percentage of the minimum expected return of portfolio (λ), maximum investment in stocks (U), and the maximum number of stocks in the investment portfolio (MaxNp) are assumed as 0.2, 0.6, 0.1, and 20, respectively.

Moreover, in order to find the ideal values of each objective function, the model is solved using Lingo software and applying these values of parameters. Therefore, values of 1.637845E+15, 0.04760250, and 0.1019152 are obtained as goal values for Z_1^*, Z_2^*, and Z_3^*, respectively.

Table 5. The values of return per time period

Stock	Time period											
	1	2	3	4	5	6	7	8	9	10	11	12
1	0.074	0.009	-0.021	-0.007	0.087	0.063	0.008	0.047	-0.05	0.043	0.04	0.094
2	0.018	-0.01	0.002	0.031	0.025	0.084	0.076	0.023	-0.00	0.067	0.075	0.067

									2			
3	0.083	0.099	0.022	0.056	0.019	0.042	-0.015	-0.011	0.06	0.021	0.085	0.054
4	0.095	-0.041	0.074	0.076	0.054	0.083	0.095	0.031	0.056	0.059	0.016	-0.07
5	0.049	0.066	0.008	-0.076	0.039	0.024	0.095	0.043	-0.01	0.068	0.029	0.045
6	0.08	0.006	0.027	0.03	0.042	0.005	0.052	0.031	0.021	0.097	0.092	0.037
7	0.077	0.017	0.003	0.093	0.001	0.059	0.045	0.094	0.088	0.044	0.052	0.037
8	0.02	0.008	0.098	0.037	0.021	0.049	0.03	0.003	0.048	-0.018	0.078	0.092
9	0.035	0.038	0.019	-0.003	0.064	0.056	0.022	0.087	0.055	0.006	0.08	0.046
10	0.033	0.078	0.053	0.016	0.052	0.089	0.05	0.083	0.043	0.025	0.056	0.036
11	0.007	0.043	0.027	0.067	0.01	0.032	0.014	0.09	0.066	0.049	0.044	0.084
12	0.047	0.023	0.069	0.083	0.026	0.076	0.026	0.067	0.049	-0.026	0.084	0.056
13	0.084	0.022	0.008	0.045	0.082	0.064	-0.002	0.037	0.053	0.084	0.035	0.009
14	0.052	0.013	0.005	0.053	0.082	0.043	0.036	0.032	-0.05	0.023	0.097	-0.001
15	0.037	-0.02	0.016	-0.018	0.001	0.003	0.068	0.088	0.091	0.033	0.076	0.052
16	0.037	0.051	0.051	0.087	0.062	0.096	0.075	0.062	0.009	-0.01	0.043	0.029
17	0.05	0.043	0.054	0.083	-0.02	0.072	0.059	0.084	0.042	0.043	0.044	0.031
18	0.026	0.097	0.083	0.092	0.08	0.034	0.024	0.079	0.049	0.041	0.065	0.037
19	0.014	0.063	0.058	0.032	0.032	0.082	0.016	0.056	0.021	0.075	0.081	-0.013
20	0.083	0.004	0.088	-0.011	0.077	0.057	-0.01	0.021	0.073	0.088	0.094	0.067
21	0.032	0.032	-0.002	0.058	0.081	0.062	0.043	0.07	0.022	0.004	0.029	0.034
22	0.00	0.06	0.06	0.09	0.03	0.03	-	0.00	0.07	-	-	0.03

	8	4	8	9	4	6	0.029	4	5	0.012	0.017	7
23	0.048	0.042	0.015	0.028	0.061	0.026	0.061	0.035	0.062	0.075	0.026	0.086
24	0.096	0.041	0.036	-0.06	0.02	0.089	0.068	0.05	-0.02	0.045	-0.092	0.099
25	0.027	0.075	0.035	0.031	0.093	0.046	0.003	-0.012	0.017	0.06	0.042	0.033
26	0.028	0.065	0.018	0.011	0.075	-0.006	0.044	0.054	0.025	0.073	0.008	0.09
27	0.055	0.031	0.052	0.038	0.03	-0.004	-0.015	-0.044	0.016	0.039	0.048	0.049
28	0.099	0.042	0.009	0.002	0.098	0.067	-0.011	0.022	0.059	0.078	0.005	0.094
29	0.021	0.062	-0.01	0.014	0.056	0.091	0.076	0.04	0.002	0.075	-0.013	0.069
30	0.097	0.027	0.065	-0.01	0.049	0.078	0.023	0.016	0.074	0.024	0.09	0.024

Table 6. The values average return, risk, and variance

No. of stock	Stock	$\hat{r}_i$	β_i	σ_i^2
1	PSER1	0.0323	0.87093	0.0020148
2	SROD1	0.0380	0.06868	0.0011504
3	GHND1	0.0429	2.01574	0.001351
4	LAMI1	0.0440	0.22713	0.0027536
5	STEH1	0.0317	0.47747	0.0019204
6	SKHS1	0.0433	0.60596	0.0009699
7	LTOS1	0.0508	1.85937	0.0010876
8	KFAN1	0.0388	0.04512	0.0012971
9	KPRS1	0.0421	0.27834	0.0007863
10	TAMI1	0.0512	0.83632	0.000522
11	KSKA1	0.0444	0.77006	0.0007828
12	SDAB1	0.0483	0.06571	0.0010219
13	SGRB1	0.0434	0.75252	0.000943
14	SHGN1	0.0321	2.26474	0.001517
15	SSEP1	0.0356	1.22938	0.0015803
16	BSTE1	0.0493	2.46817	0.0009395
17	SRMA1	0.0488	0.86654	0.0007515
18	DMVN1	0.0589	0.70185	0.0007121
19	CIDC1	0.0431	0.95924	0.0009286
20	SPKH1	0.0526	2.53406	0.0016097

21	SHSI1	0.0388	0.60239	0.0006408
22	SMAZ1	0.0306	0.62865	0.0016452
23	SURO1	0.0471	0.26890	0.0004872
24	SKOR1	0.0310	0.89911	0.0036724
25	KHOC1	0.0375	0.41613	0.0008477
26	SIMS1	0.0404	0.85472	0.000953
27	ABAD1	0.0246	0.67493	0.0009492
28	SFKZ1	0.0470	0.49921	0.0016605
29	SSOF1	0.0403	0.52709	0.0013175
30	NSTH1	0.0464	0.21421	0.0011461

Results

The results of testing the models using various approaches are reported. They are representing the proportions of the budgets to be invested in each company's stock. Several criteria are used to compare three approaches such as the return rate, the real diversity index, systematic risk, unsystematic risk, amount of stock in the portfolio, VaR, CVaR, Treynor ratio, and Sharpe ratio. The comparison of these approaches is illustrated in Table 7.

Also, it should be noted that all calculations are based on using the branch and bound solver of the Lingo 9 software and MATLAB software. After solving the problem with the data of the previous section and the mentioned goals, Tables 7-9 have been obtained.

Table 7. Comparison of related criteria in three methods

Criteria	LP-metric based methods		
	Lingo (local optimum)	GA	PSO
No. of stock in the basket	15	20	18
Portfolio return	7.2601982	7.3625147	6.8695723
Sterling diversity index	5.75207E+12	5.849442E+12	5.7125365E+12
Unsystematic risk	0.0001763	0.00014698	0.00018759
Systematic risk	1.003981	1.009777	1.022365
VaR	0.1365164	0.1275286	0.1478549
Treynor ratio	2.848356	2.736524	2.7475632
Sharpe ratio	0.268086	0.245869	0.2763576
CPU time	2700.53	58.36	112.42

According to Table 7, the GA has the least CPU time and VaR, and the highest portfolio return and sterling diversity index. So, it can be selected as the best approach and the related decision variables of it are provided in Table 8.

Table 8. Obtained values of decision variables using GA

| Variable	Value	Variable	Value	Variable	Value	Variable	Value
E_Ω^{max}	0.00508	$x(19)$	0.10000	$q(7)$	1.00	$q(25)$	0.00
$x(1)$	0.00000	$x(20)$	0.00000	$q(8)$	1.00	$q(26)$	1.00
$x(2)$	0.01848	$x(21)$	0.03430	$q(9)$	1.00	$q(27)$	0.00
$x(3)$	0.02140	$x(22)$	0.02262	$q(10)$	1.00	$q(28)$	1.00
$x(4)$	0.03222	$x(23)$	0.00000	$q(11)$	0.00	$q(29)$	0.00
$x(5)$	0.10000	$x(24)$	0.03918	$q(12)$	1.00	$q(30)$	0.00
$x(6)$	0.02104	$x(25)$	0.00000	$q(13)$	1.00	$y(1)$	1.00
$x(7)$	0.10000	$x(26)$	0.01795	$q(14)$	0.00	$y(2)$	0.00
$x(8)$	0.02733	$x(27)$	0.00000	$q(15)$	1.00	$y(3)$	0.00
$x(9)$	0.03438	$x(28)$	0.10000	$q(16)$	1.00	$y(4)$	0.00
$x(10)$	0.03156	$x(29)$	0.00	$q(17)$	1.00	$y(5)$	0.00
$x(11)$	0.00000	$x(30)$	0.00	$q(18)$	0.00	$y(6)$	0.00
$x(12)$	0.01618	$q(1)$	0.00	$q(19)$	1.00	$y(7)$	0.00
$x(13)$	0.02490	$q(2)$	1.00	$q(20)$	0.00	$y(8)$	1.00
$x(14)$	0.00000	$q(3)$	1.00	$q(21)$	1.00	$y(9)$	0.00
$x(15)$	0.09204	$q(4)$	1.00	$q(22)$	1.00	$y(10)$	0.00
$x(16)$	0.10000	$q(5)$	1.00	$q(23)$	0.00	$y(11)$	0.00
$x(17)$	0.06643	$q(6)$	1.00	$q(24)$	1.00	$y(12)$	0.00
$x(18)$	0.00000						

According to Table 8, the following 20 companies can be selected as an optimal portfolio: SROD1, GHND1, LAMI1, STEH1, SKHS1, KFAN1, KPRS1, TAMI1, KSKA1, SGRB1, SHGN1, BSTE1, SRMA1, DMVN1, SPKH1, SMAZ1, SURO1, KHOC1, ABAD1, and SSOF1. And the values of Z_1, Z_2, and Z_3 result in 5.849442E+12, 1.004157 and 0.1275286, respectively.

Furthermore, to show the efficiency of the proposed multi-objective model, it is compared with Markowitz and diversification-risks models. These values are presented in Table 9.

Table 9. Compression of the proposed model with Markowitz and diversification-risks models using LP-metric based GA

Criteria	Models			
	Markowitz	Diversification-risks	Diversification-risks-VaR	Diversification-risks-CVaR
Portfolio return	7.2324586	7.1854796	7.3625147	7.4101649
Sterling diversity index	5.72326	5.94315	5.849442	5.713564
Unsystematic risk	0.0001632	0.000149687	0.00014698	0.00014911
Systematic risk	1.0037235	1.0056985	1.009777	1.008813
Treynor ratio	2.475325	2.811276	2.736524	2.684121
Sharpe ratio	0.2138545	0.2643252	0.245869	0.240647
CPU time	51.12	54.75	58.36	61.05

To validate the framework, their model includes Markowitz, diversification-risks, diversification-risks-VaR, and diversification-risks-CVaR models are solved 15 times with different population sizes to display the average returns of stock portfolios, as shown in Figure 6. In this investigation, we have used the conventional method of calculation VaR (VaR= $-\sigma^2 Z_\alpha -\mu$) with the assumption that the distribution of returns is normal with a mean of μ and σ^2 variance at the α confidence level, based on these assumptions the conditional value at risk (CVaR) model is calculated which is one of the developed versions of VaR model. This risk measure quantifies the amount of tail risk of the portfolio that proposed in 2000 by Rockafellar and Uryasev.
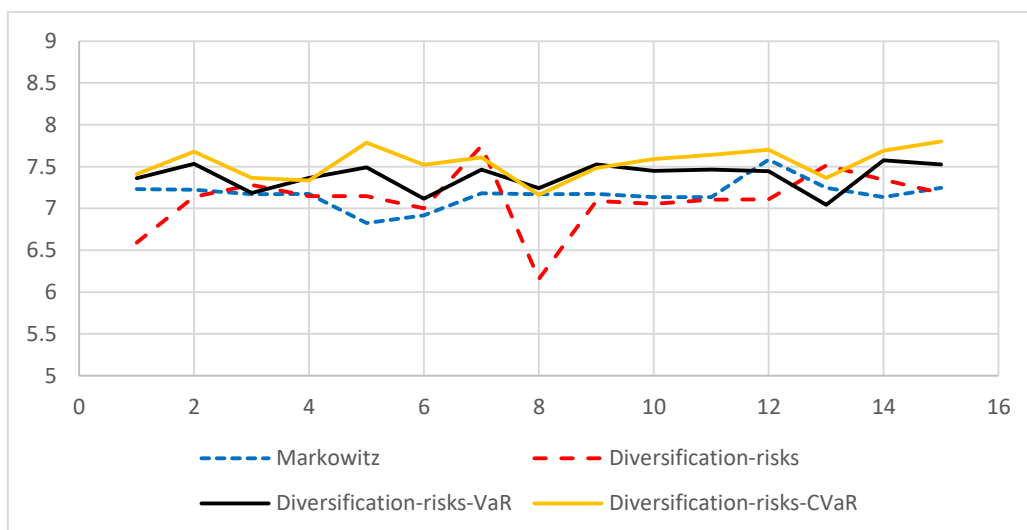
Figure 6. Average of stock returns in different iteration using four model

Since the optimal return rate for investment based on a market index is defined as an interval between 6.5 and 8.5, so based on Fig. 6, only in the diversification-risks model the average return is lower than the desired value. In other models, the average portfolio returns are higher than the mentioned values. Therefore, it can be concluded that the proposed model can be chosen as a suitable model.

Conclusion

This study introduces the Euclidean distance criterion as a measure of stock portfolio diversification and uses a multi-objective model attempts to select optimal stock portfolios. The model aims to maximize diversification and minimize VaR, and stock risks including systematic and unsystematic risks. Also, maximizing returns are considered as a constraint of this model. Since the proposed model is nonlinear (and regarding computational complexity, it is NP-hard) the study utilizes two meta-heuristic algorithms to solve the model. It further validates portfolio selection by using the data of the top 30 active companies in the TSE for the 12 months. The findings show that the GA has the least CPU time and VaR, and the highest portfolio return and sterling diversity index. So, it can be selected as the best approach, and the related decision variables can be reported based on the results of 20 active companies in the TSE that were selected as the optimal portfolio. Moreover, to show the effectiveness of the proposed multi-objective model, we compared our proposed model with Markowitz and diversification-risks models. After testing our model for 15 times with different population sizes, we found that the average portfolio returns are higher than the desirable values resulted from the market index.

Various studies have been conducted to investigate the relationship between returns and risk in the selection of stock portfolios. However, in the field of quantification of diversity index, no particular mathematical formulation has not been introduced so far. In this research, the Euclidean distance criterion has been tested as a tool for the quantitative assessment of the diversity index. Also, the research model is designed based on deterministic parameters, the unfeasibility of short selling, and without considering the transaction cost. For the future study, it is suggested that in addition to the uncertainties of parameters of a model, the effect of short selling and transaction costs on the index of diversity be studied. In addition, the different methods in VaR computations can be utilized in the model and the results should be compared.

References

Aluko, O. A., Fapetu, O., & Azeez, B. A. (2018). International portfolio diversification in the Nigerian stock market: A global financial crisis perspective. *Future Business Journal*, *4*(2), 189-194.

Beaudreau, A. H., Chan, M. N., & Loring, P. A. (2018). Harvest portfolio diversification and emergent conservation challenges in an Alaskan recreational fishery. *Biological Conservation*, *222*, 268-277.

Briere, M., Oosterlinck, K., & Szafarz, A. (2015). Virtual currency, tangible return: Portfolio diversification with bitcoin. *Journal of Asset Management*, *16*(6), 365-373.

Chan Kim, W., Hwang, P., & Burgers, W. P. (1989). Global diversification strategy and corporate profit performance. *Strategic management journal*, *10*(1), 45-57.

Chang, T. J., Yang, S. C., & Chang, K. J. (2009). Portfolio optimization problems in different risk measures using a genetic algorithm. *Expert Systems with Applications*, *36*(7), 10529-10537.

Chen, Y., Shi, Y., Wei, X., & Zhang, L. (2014). How does credit portfolio diversification affect banks' return and risk? Evidence from Chinese listed commercial banks. *Technological and Economic Development of Economy*, *20*(2), 332-352.

Dang, V. D. (2019). The effects of loan portfolio diversification on Vietnamese banks' return. In *International Econometric Conference of Vietnam* (pp. 928-939). Springer, Cham.

Diyarbakırlıoğlu, E., & Satman, M. H. (2013). The maximum diversification index. *Journal of Asset Management*, *14*(6), 400-409.

Eberhart, R., & Kennedy, J. (1995, October). A new optimizer using particle swarm theory. In *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science* (pp. 39-43). Ieee.

Francis, J. C., & Kim, D. (2013). *Modern portfolio theory: Foundations, analysis, and new developments* (Vol. 795). John Wiley & Sons.

Gagliardini, P., & Gouriéroux, C. (2013). Granularity adjustment for risk measures: Systematic vs unsystematic risks. *International Journal of Approximate Reasoning*, *54*(6), 717-747.

Georgalos, K., Paya, I., & Peel, D. A. (2018). On the contribution of the Markowitz model of utility to explain risky choice in experimental research. *Journal of Economic Behavior & Organization*.

Holland, J. H. (1974). Erratum: Genetic algorithms and the optimal allocation of trials. *SIAM Journal on Computing*, *3*(4), 326-326.

Jackwerth, J. C., & Slavutskaya, A. (2016). The total benefit of alternative assets to pension fund portfolios. *Journal of Financial Markets*, *31*, 25-42.

Jadhao, G. and Chandra, A., (2017). Application of VIX and entropy indicators for portfolio rotation strategies. *Research in International Business and Finance*, *42*, pp.1367-1371.

Jones, C. P. (2007). *Investments: analysis and management*. John Wiley & Sons.

Kajtazi, A., & Moro, A. (2019). The role of bitcoin in well diversified portfolios: A comparative global study. *International Review of Financial Analysis*, *61*, 143-157.

Kalashnikov, V., Benita, F., López-Ramos, F., & Hernández-Luna, A. (2017). Bi-objective project portfolio selection in Lean Six Sigma. *International Journal of Production Economics*, *186*, 81-88.

Kara, G., Özmen, A., & Weber, G. W. (2019). Stability advances in robust portfolio optimization under parallelepiped uncertainty. *Central European Journal of Operations Research*, *27*(1), 241-261.

Kim, W., Kim, Y. M., Kim, T. H., & Bang, S. (2018). Multi-dimensional portfolio risk and its diversification: A note. *Global Finance Journal*, *35*, 147-156.

Liu, W. (2018). Portfolio diversification across cryptocurrencies. *Finance Research Letters*.

Markowitz, H. M., & Todd, G. P. (2000). *Mean-variance analysis in portfolio choice and capital markets* (Vol. 66). John Wiley & Sons.

Oh, K. J., Kim, T. Y., & Min, S. (2005). Using genetic algorithm to support portfolio optimization for index fund management. *Expert Systems with Applications*, *28*(2), 371-379.

Oloko, T. F. (2018). Portfolio diversification between developed and developing stock markets: The case of US and UK investors in Nigeria. *Research in International Business and Finance*, *45*, 219-232.

Oyenubi, A. (2016). Diversification measures and the optimal number of stocks in a portfolio: An information theoretic explanation. *Computational Economics*, 1-29.

Paut, R., Sabatier, R., & Tchamitchian, M. (2019). Reducing risk through crop diversification: An application of portfolio theory to diversified horticultural systems. *Agricultural Systems*, *168*, 123-130.

Platanakis, E., Sakkas, A., & Sutcliffe, C. (2019). Harmful diversification: Evidence from alternative investments. *The British Accounting Review*, *51*(1), 1-23.

Pola, G. (2016). On entropy and portfolio diversification. *Journal of Asset Management*, *17*(4), 218-228.

Reilly, F. K., & Brown, K. C. (2011). *Investment analysis and portfolio*

management. Cengage Learning.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of risk*, *2*, 21-42.

Steinberg, J. B. (2018). International portfolio diversification and the structure of global production. *Review of Economic Dynamics*, *29*, 195-215.

Taguchi, G. (1986). *Introduction to quality engineering: designing quality into products and processes* (No. 658.562 T3).

Zelany, M. (1974). A concept of compromise solutions and the method of the displaced ideal. *Computers & Operations Research*, *1*(3-4), 479-496.

Bibliographic information of this paper for citing:

Ebrahimi, Seyed Babak; Abdolahi Moghadam, Mostafa & Safaie, Nasser (2021). The Quantitative Diversity Index in Multi-Objective Portfolio Model. *Iranian Journal of Finance*, 5(1), 122-146.