Int. J. Nonlinear Anal. Appl. 15 (2024) 7, 135–152
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2023.30955.4526

Designing and explaining the portfolio optimization model using censored models and meta-heuristic algorithm

Zeinab Kazemi[a], Gholamreza Zomorodian[a,*], Mirfeiz Fallah Shams Layalestani[a], Farhad Hanifi[b]

[a] *Department of Financial Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran*
[b] *Department of Business Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran*

(Communicated by Sirous Moradi)

Abstract

One of the important topics discussed in the stock market, which should be considered by both natural and legal investors, is choosing an optimal investment portfolio. In this regard, investors are studied in order to select the best portfolio based on risk and return. However, traditional investment methods do not focus on portfolio optimization and only consider the highest return and lowest risk. This research addresses the gap in solving the problem of wide portfolio optimization by comparing answers using more effective and efficient metaheuristic optimization algorithms, thus reducing the probability of error. During this research, metaheuristic optimization methods are well-designed and studied, and then used to optimize the portfolio despite real market limitations. The developed algorithms are all implemented to solve the extended portfolio optimization problem. In this research, more effective and efficient metaheuristic optimization algorithms are used to solve the problem of wide portfolio optimization and by comparing the answers, the probability of error can be almost zero. The stock portfolios formed by the model based on censoring models have more returns and less risk (variance) than the invasive weed algorithm, showing the superiority of the proposed model in comparison to the invasive weed algorithm. The findings of the research have filled the research gap in investment portfolio valuation and demonstrate that the proposed model has effectively considered investment portfolio selection conditions and determined an optimal investment portfolio.

Keywords: investment portfolio, portfolio optimization, censorship models, invasive weed algorithm
2020 MSC: 49M25, 00A71

1 Introduction

Choosing the right portfolio is one of the most important goals for investors. Two factors that can play an important role in the selection process are return and portfolio risk [14]. This problem can be written as a mathematical programming problem and can be easily solved. For example, Meghwani and Thakur [11] state that some investors choose their desired stocks based on the past performance of existing stocks, while others use other factors such as liquidity in selecting stocks. Additionally, fundamental and technical analysis are also used in stock and portfolio selection [15]. However, Zhou et al. [21] indicate that depending on the current conditions and the level of knowledge

*Corresponding author
Email addresses: zeinabkazemi1987@gmail.com (Zeinab Kazemi), gh.zomorodian@gmail.com (Gholamreza Zomorodian), fallahshams@gmail.com (Mirfeiz Fallah Shams Layalestani), hanifi_farhad@yahoo.com (Farhad Hanifi)

Received: March 2023 *Accepted:* June 2023

of the investors, different strategies and methods are used to choose stocks and portfolios. The methods for solving the portfolio optimization problem can be divided into two categories: precise and search-oriented methods. 1) Linear programming method: The linear programming method is one of the exact methods. This method is very simple, and its ultimate goal is to reduce portfolio risk as much as possible while achieving a certain return [8]. However, it is important to note that this method only solves the problem in a linear way, and generally a better result can be obtained by using non-linear methods [20]. 2) Non-linear programming method (quadratic): The non-linear or quadratic method is the original Markowitz method [10], which can be used to find the optimal portfolio [9]. One of the drawbacks of these precise methods is their time-consuming nature, which has been largely addressed today with high-speed computers. By solving the Markowitz problem, you can optimize a single portfolio, and by solving this problem for several portfolios (with different risk and return), you can obtain the frontier efficiency [13].

In this article, the issue of designing and explaining the portfolio optimization model using censored models and meta-heuristic algorithms is discussed. Meta-heuristic algorithms are algorithms that are generally inspired by nature and can be used to solve nonlinear problems with constraints. These algorithms have three different types: 1) single-member search algorithms, 2) group search algorithms, 3) combined algorithms.

Since deciding on the types of investments and choosing the most suitable investment method is one of the most fundamental issues in investment management, the present research aims to design and explain an appropriate model for optimizing the portfolio using censored models and meta-heuristic algorithms. The first step is to accurately define the problem and identify the decision parameters and variables, so that it can be more compatible with real-world problems. After validating the model, and taking into account that the desired final result of the research is the optimization of the stock investment portfolio with the approach of minimizing financial risks and maximizing stock returns, efforts are made to provide suitable solutions to solve the mathematical model of the problem using meta-heuristic algorithms. Considering the multi-objective nature of the investigated model (maximization of yield and minimization of risk), there is no dominant answer, and the results will be examined by using more effective and efficient meta-heuristic algorithms.

2 Literature review

In recent years, studies have been conducted to optimize stock portfolios, and the use of various algorithms to optimize investment portfolios has increased. Researchers such as Wang et al. [19] have used mathematical programming to solve the stock portfolio problem. Mokhin proposed an innovative method using the genetic algorithm for different stock portfolios, with risk calculated in different ways. Their goal was to investigate the efficiency of the genetic algorithm for solving the problem of stock portfolio optimization with different risk models, particularly for portfolios that consider integer constraints. By using different risk calculation models in this genetic algorithm method, investors can obtain the efficiency limit for a fixed amount of capital. They found that a smaller portfolio is more efficient than a larger one [19].

In another study, Aghamohammadi et al. [1] in a research entitled "Research on multi-objective evaluation algorithms in order to solve the problem of portfolio optimization and other practical cases in economic and financial issues" found that the development and improvement of multi-objective evaluation algorithms and the creation of complex formulations in the financial and economic fields have led to a two-way desire for both research societies. The classification chosen for this research provides a distinction between the portfolio optimization problem and the use cases in this field.

Mercangöz and Ergolu [12] conducted a study entitled "Prediction of the mean variance model for the selection of limited portfolio assets" which investigated the limited selection of portfolios using the multi-objective evolutionary algorithm. They proposed self-adjusting multi-objective particle mass optimization as an effective method for portfolio optimization. By using predictive multi-objective evaluation, the mean-variance model was assumed to be another solution to the common Markowitz mean-variance model, which is used to solve the limited portfolio optimization problem. The research investigated the limited portfolio selection using the multi-objective evolutionary algorithm. They used self-adjusting multi-objective particle mass optimization algorithm as an effective method for portfolio optimization. In the proposed model, future capital returns were predicted using the artificial neural network model, and then optimization was done using multi-objective evaluation algorithms. The research results showed that the Pareto solutions approach includes maintaining sufficient diversity and also a complete comparison with the Markowitz model. Additionally, the superiority of the self-adjusting multi-objective particle mass optimization algorithm and its responsiveness in the optimization of stock portfolio selection were demonstrated.

The use of genetic algorithms and financial coefficients has also been studied by Akbay et al. [2]. In their study, they investigated the series-based group stock portfolio (GSP) optimization approach using the genetic grouping algorithm (GGA) with symbolic summation approximations. The purpose of the study was to derive GSPs that not only have similar stock price series in groups but also have high returns. To achieve this goal, the study considered stock price series, which is a type of time series. Because the number of time series data points is large, extracting data from it can be time-consuming, so dimensionality reduction techniques were used to speed up the exploration process. Two dimensionality reduction techniques for time series, namely symbolic summation approximation (SAX) and extended symbolic summation approximation (ESAX), were used in the proposed approach. The results showed that the return on investment (ROI) of the proposed approach using the fit functions with SAX is approximately 16% to 18% and better than the ROI obtained with ESAX. However, the proposed approach with ESAX achieved better group similarity than SAX.

Finally, in a study by Salehpoor and Molla-Alizadeh-Zavardehi [16] entitled "A flexible trading system: optimizing the investment portfolio with the aim of risk and return using a learning model in order to maximize profit", they found that the use of a risk-based balanced return verification system with transaction cost and retraining mechanism leads to stop loss in the market, and in other words, the proposed trading system reacts to the effects of transaction cost and continuously improves the benchmarks of yield funds.

3 Research method

Research modelling based on stock optimization with financial risk-return objectives is presented as follows: To solve the problem of choosing an investment portfolio, the symbols used in this study are based on the Markowitz model. Markowitz was the first researcher to use variance or standard deviation as a measure of risk. He postulated that the classical form can be formulated as the following equation:

3.1 Mathematical modelling

Markowitz [10] provides a definition of risk measure using variance or standard deviation. His model with risk minimization is presented as follows:

$$\min \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \delta.$$

The above equation minimizes the total variance (risk) associated with the investment portfolio. Markowitz's model is based on the following assumptions: investors are risk-averse and have increasing expected utility, and their final utility curve for wealth is decreasing. In the real world, investors face many constraints, such as trading limits and portfolio size. If the number of variables increases or if there are additional restrictions, such as restrictions on investment weight, ceiling and floor restrictions, the problem of optimizing the investment portfolio becomes more complex. Despite these limitations, researchers strive to develop the Markowitz model by incorporating these limitations into the original model. Therefore, to apply the above equation model, the investment portfolio optimization problem is examined as a problem with a multi-objective function, where the expected return is maximized (the first objective function) and the risk is minimized (the second objective function). Therefore, the standard form of the objective and constraints for optimization in this study will be as described in the following equations:

$$\max : z_1 = \sum_{i=1}^{n} \mu_i x_i$$

$$\min : z_2 = \sum_{i=1}^{n} \sigma_i^2 x_i^2 + \sum_{i} \sum_{j} \sigma_{ij} x_i x_j$$

$$i \neq j$$

$$z_3 = \sum_{i=1}^{n} B_i x_i \to 1$$

$$St: \quad \sum_{i=1}^{n} x_i = 1$$

$$l_i y_i \leq x_i \leq u_i y_i$$

$$\sum_{i=1}^{n} y_i = k$$

$$y_i = 0 \text{ or } 1$$

$$x_i \geq 0$$

3.2 Definition of parameters

Z_i: optimization of stocks with the goals of maximization of return - minimization of risk, u_i: average return of the i-th share, σ_i: standard deviation of the return of the i-th share, σ_{ij}: covariance of the i-th share and the j-th share, β_i: systematic risk of the i-th share, n: total number of shares, K: the number of selected shares, u_i: the maximum investment amount in the i-th share and l_i: the minimum investment amount in the i-th share.

3.3 Definition of decision variables

0 or 1: $y_i = 1$ to invest in the i-th share or not: X_i percentage of investment in the i-th share, w_i: weight (importance) of the i-th share in terms of financial risks that is obtained from a MADM.

Two important components in making investment decisions are the level of risk and the return on capital assets. Choosing the optimal asset portfolio often involves a trade-off between risk and return, with higher risk assets offering the potential for higher returns. Identifying the efficiency frontier related to the asset portfolio allows investors to obtain the highest expected return from their investment, based on their utility function and degree of risk aversion and risk tolerance. Each investor, based on their risk tolerance and risk aversion, chooses a point on the efficient frontier and determines the composition of their portfolio with the goal of maximizing returns and minimizing risk.

Optimizing the portfolio means choosing the best combination of financial assets in a way that maximizes the return of the investment portfolio while minimizing portfolio risk. Multi-objective optimization problems (MOPs) in meta-heuristic models have a standard format, which is defined as optimization problems that must satisfy more than one objective, provide multiple optimal solutions, and include more than one objective function that are optimized simultaneously. This is referred to as MOP, which optimizes the objective function vector as follows [7]:

$$F(X) = [F_1(X), F_2(X), ..., F_m(X)]^T$$

while, $X = [X_1, X_2, ..., X_d]$ a vector of variables and "d" and "m" are the number of variables and objectives. A simple approach to solve the multi-objective optimization problem is to use various weights to transform a multi-objective function into a single-objective optimization problem. This problem can be formulated based on the following equation [4]:

$$F = \sum_{i=1}^{m} w_i f_i$$

while m is the number of objective functions, and w_i and f_i are weight factors and objective functions, respectively. However, this method can be time-consuming and is considered a major limitation. A common solution to multi-objective optimization problems is to maintain a set of the best solutions in an archive and updating the archive in each iteration. In this method, the best solutions are defined as non-dominated solutions or Pareto optimal solutions. A solution is considered a non-dominated solution if and only if it satisfies the following conditions.

Pareto dominance [16]:

$$U = (u_1, u_2, u_3, ..., U_N) < (v_1, v_2, v_3, ..., V_n)$$

If and only if U is fractionally less than v in the target space, which means [3, 6]:

$$\begin{bmatrix} F_i(U) \leq F_i(V) \\ F_i(U) \prec F_i(V) \end{bmatrix} \begin{matrix} v_i \\ \exists_i \end{matrix} \quad i - 1, 2, ..., m$$

Pareto optimal solution: vector U is an optimal solution if and only if none of the other solutions can dominate it. A set of Pareto optimal solutions is called Pareto optimal front (PF OPTIMAL). figure below; It shows that among the three solutions A, B, C, c has the highest value for f1 and f2 and as a result it is considered the dominant solution; On the contrary, both solutions A, B can be considered as non-dominant solutions [18].

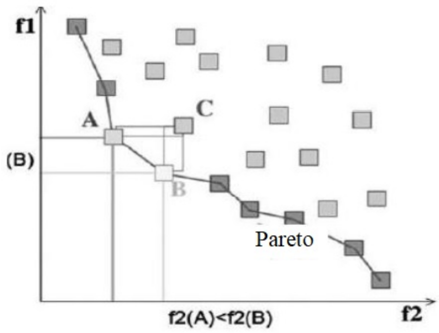

Figure 1: A set of Pareto optimal solutions

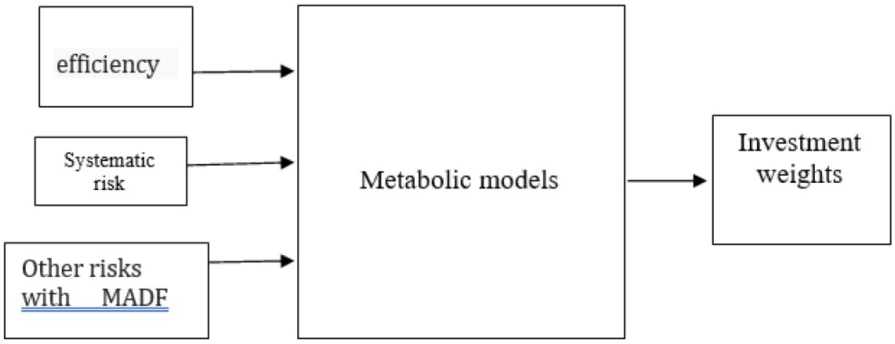

Figure 2: Metabolic models for investment weights

Stock price index: The stock price index of companies shows the general trend of market movement, in fact, this trend is the degree of success of the country's capital market. In order to calculate the rate of return of shares of different industries, the following formula is used [17]:

$$R_{it} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} \times 100$$

where $P_{i,t}$ and $P_{i,t-1}$ show the stock price of the i-th industry at the moment "t" and $t-1$, respectively, and $R_{i,t}$ shows the rate of return of the i-th industry at the moment of time t.

The process of minimizing the variance of the risky stock investment portfolio is expressed as follows:

$$\min\ \sigma_p^2 = x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + x_3^2\sigma_3^2 + x_4^2\sigma_4^2 + 2x_1x_2\ cov(1,2) + ... + 2x_3x_4\ cov(3,4) \tag{3.1}$$

st:

$$r_p = x_1r_1 + x_2r_2 + x_3r_3 + x_4r_4 \tag{3.2}$$

$$x_1 + x_2 + x_3 + x_4 = 1 \tag{3.3}$$

equation (3.1) shows how to calculate the return variance of an investment portfolio consisting of stocks of selected industries in each portfolio. In equation (3.1), represents the portfolio return variance, represents the variance of the return of industries' shares in the portfolio, and represents the share of the industry in the formed basket. Equations (3.2) and (3.3) state that the return of a portfolio is equal to the weighted average return of stocks in the portfolio. Solving the above-constrained quadratic programming problem leads to the calculation of the optimal weights of risky stocks in the portfolio. If we form the Lagrange function to minimize the variance of the investment portfolio as follows, then we have:

$$\begin{aligned} L =& x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + x_3^2\sigma_3^2 + x_4^2\sigma_4^2 + 2x_1x_2\ cov(1,2) + ... + 2x_3x_4\ cov(3,4) + \lambda(\bar{r}_p - x_1r_1 + x_2r_2 + x_3r_3 + x_4r_4) \\ &+ \gamma(1 - x_1 - x_2 - x_3 - x_4) \end{aligned} \tag{3.4}$$

in equation (3.4), it represents the constant level of portfolio efficiency. The formation of the first-order optimization conditions is completed in the form of the following matrix:

$$\begin{bmatrix} 2\sigma_1^2 & 2cov(1,2) & 2cov(1,2) & cov(1,2) & -r_1 & -1 \\ 2cov(2,1) & 2\sigma_2^2 & 2cov(2,3) & 2cov(2,4) & -r_2 & -1 \\ 2cov(3,1) & 2cov(3,2) & 2\sigma_3^2 & 2cov(3,4) & -r_3 & -1 \\ 2cov(4,1) & 2cov(4,2) & 2cov(4,3) & 2\sigma_4^2 & -r_4 & -1 \\ -r_1 & -r_2 & -r_3 & -r_4 & 0 & 0 \\ -1 & -1 & -1 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_1 \\ x_1 \\ x_1 \\ \lambda \\ \gamma \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \bar{r}_p \\ 1 \end{bmatrix} \tag{3.5}$$

On the other hand, if the total return on investment consists of return on holding risky assets and risk-free assets (risk-free interest rate) as follows:

$$r_k = yr_p + (1-y)r_f \tag{3.6}$$

Assuming that the utility function of a risk-averse person is as follows:

$$U_k = r_k - \frac{15}{1000}\sigma_k^2 \tag{3.7}$$

Then, by combining relations (3.6) and (3.7), we can write:

$$U_k = yr_p + (1-y)r_f - \frac{15}{1000}\sigma_k^2 \tag{3.8}$$

By deriving the high utility function with respect to "y", the optimal share of the set of risky assets (y^*) can be obtained as follows:

$$y^* = \frac{r_p - r_f}{0.03sigma_p^2} \tag{3.9}$$

in relation (3.9), the value is obtained from the optimization process above and is taken into account exogenously. At the same time, the variance matrix of the covariance of risky assets is estimated from the estimation of heterogeneous multivariate models of fixed conditional variance and dynamic conditional variance.

4 Findings

4.1 Ranking of industries using censorship models

In this section, five industries (according to the available data) are studied: (1. Pharmaceutical materials and products, 2. Automobiles and parts manufacturing, 3. Machinery and equipment, 4. Food and beverage products except sugar and sugar, and 5. Chemical products) are ranked by censorship models and based on five indicators (1. Stock liquidity, 2. Stock yield, 3. Earnings per share, 4. Price-to-earnings ratio, and 5. Systematic risk).

The first step in censorship models is assigning relative weights to indicators. As you can see in table 1, four experts (in accordance with the standard of the investigated method) assigned one of the linguistic descriptions (very high, high, medium, low and/very low) to each of the indicators.

Table 1: The importance of indicators from the point of view of four experts

Indicators	**First expert**	**Second expert**	**Third expert**	**Fourth expert**
Stock liquidity	Very high	High	Very high	High
Stock returns	Very low	Very low	Average	Normal
Earnings per share	Low	Low	Very low	Very low
Price-to-earnings ratio per share	High	Very high	High	Low
Systematic risk	Normal	Average	Low	Very high

In figure 3 and in table 2, we see the transformation of linguistic features into trapezoidal numbers of censorship models. (in the form of a trapezoidal distribution with four minimum point values, two middle points and a maximum point).

The weights presented in the table below are presented according to the fuzzy standard.

Table 2: Linguistic features and trapezoid numerical weights [5]

Features	Weights (Numberic)
Very low	(0,0,0,0,3)
Low	(0,0.3,0.3,0.5)
Average	(0.2,0.5,0.5,0.8)
High	(0.5,0.7,0.7,1)
Very high	(0.7,1,1,1)

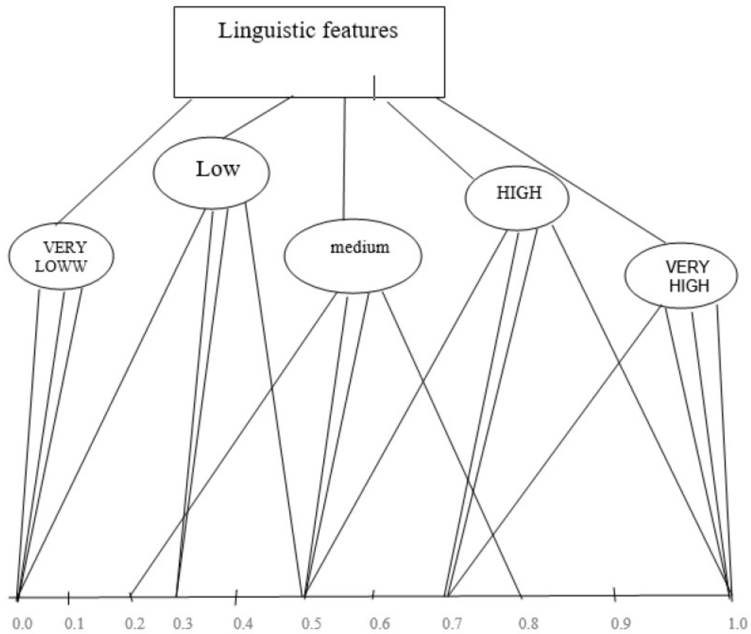

Figure 3: The transformation of linguistic features into trapezoidal numbers of censorship models

Now, using tables 1 and 2, we calculate the trapezoidal numerical weights of each index. For example, to calculate the trapezoidal numerical weight for the stock liquidity index, we must calculate the numerical equivalent (very high + high + very high + high) at four trapezoidal points (minimum point, two middle points and maximum point), that is:

$$\text{Lowerbound} = (0.7 + 0.5 + 0.7 + 0.5)/4 = 0.6$$
$$\text{First middle weight} = (1 + 0.7 + 1 + 0.7)/4 = 0.8$$
$$\text{Second middle weight} = (1 + 0.7 + 1 + 0.7)/4 = 0.85$$
$$\text{Upperbound} = (1 + 1 + 1 + 1)/4 = 1$$

Therefore, the trapezoidal numerical weight for the stock liquidity index (W_1) is equal to:

$$W_1 = (0.6, 0.85, 0.85, 1)$$

in the same way, trapezoidal numerical weights can be calculated for stock return indices (W_2), earnings per share (W_3), price-to-earnings ratio (W_4), and systematic risk (W_5).

$$W_2 = (0.1, 0.25, 0.25, 0.55)$$
$$W_3 = (0, 0.15, 0.15, 0.4)$$
$$W_4 = (0.425, 0.675, 0.675, 0.875)$$
$$W_5 = (0.275, 0.575, 0.575, 0.775)$$

In the second step of the censorship models, to evaluate each of the five industries (1- pharmaceutical materials and products, 2- automobile and parts manufacturing, 3- machinery and equipment, 4- food and beverage products except sugar and sugar, and 5- products chemistry) we discuss the basics of each of the five indices (1- stock liquidity, 2- stock yield, 3- earnings per share, 4- price-earnings ratio, and 5- systematic risk). For this purpose, linguistic descriptions (very good, good, average, bad and very bad) are used in tables 3 to 7.

Table 3: Evaluation of industries according to the stock liquidity index (according to the researcher's findings)

Stock liquidity	**First expert**	**Second expert**	**Third expert**	**Fourth expert**
Pharmaceutical materials and products	Very good	Good	Good	Very bad
Automobile and parts manufacturing	Bad	Very good	Average	Bad
Equipment and machinery	Very bad	Average	Very bad	Good
Food and beverage products	Average	Very bad	Very good	Very good
Chemical products	Good	Bad	Bad	Average

Table 4: Evaluation of industries according to the stock return index (according to the researcher's findings)

Stock returns	**First expert**	**Second expert**	**Third expert**	**Fourth expert**
Pharmaceutical materials and products	Good	Bad	Good	Average
Automobile and parts manufacturing	Very Bad	Very good	Average	Bad
Equipment and machinery	Bad	Good	Very good	Good
Food and beverage products	Very good	Very bad	Very bad	Very good
Chemical products	Average	Average	Bad	Very bad

Table 5: Evaluation of industries in relation to profit per share index (according to the researcher's findings)

Earnings per share	**First expert**	**Second expert**	**Third expert**	**Fourth expert**
Pharmaceutical materials and products	Good	Very good	Bad	Good
Automobile and parts manufacturing	Very bad	Bad	Very bad	Very bad
Equipment and machinery	Bad	Very bad	Very good	Bad
Food and beverage products	Average	Average	Good	Very good
Chemical products	Very good	Good	Average	Average

Table 6: Evaluation of industries according to the index of the ratio of price to profit per share (according to the researcher's findings)

Price-to-earnings ratio per share	**First expert**	**Second expert**	**Third expert**	**Fourth expert**
Pharmaceutical materials and products	Bad	Bad	Average	Bad
Automobile and parts manufacturing	Good	Very good	Very bad	Very good
Equipment and machinery	Very good	Very bad	Very good	Good
Food and beverage products	Average	Good	Good	Average
Chemical products	Very bad	Average	Bad	Very bad

Table 7: Evaluation of industries according to the index of the ratio of price to profit per share (according to the researcher's findings)

Systematic risk	**First expert**	**Second expert**	**Third expert**	**Fourth expert**
Pharmaceutical materials and products	Bad	Bad	Average	Bad
Automobile and parts manufacturing	Good	Very good	Very bad	Very good
Equipment and machinery	Very good	Very bad	Very good	Good
Food and beverage products	Average	Good	Good	Average
Chemical products	Very bad	Average	Bad	Very bad

We use the same weights as in table 2 to convert linguistic attributes (very good, good, average, bad and very bad) into trapezoidal numbers of censoring models. Now, using table 2 and tables 3 to 7, we calculate the trapezoidal numerical weights of industry i in relation to index j. For example, to calculate the numerical weight of the trapezoidal material and pharmaceutical products industry in relation to the stock liquidity index, the numerical equivalent (very good + good + good + very bad) should be calculated in four trapezoidal points (minimum point, two middle points and the maximum point). to do, that is:

$$\begin{aligned} \text{Lowerbound} &= (0.7 + 0.5 + 0.5 + 0)/4 = 0.425 \\ \text{First middle weight} &= (1 + 0.7 + 0.7 + 0)/4 =)0.6 \\ \text{Second middle weight} &= (1 + 0.7 + 0.7 + 0)/4 =)0.6 \\ \text{Upperbound} &= (1 + 1 + 1 + 0.3)/4 = 0.825 \end{aligned}$$

Therefore, the numerical weight of the trapezoid for the pharmaceutical industry and the stock liquidity index (S_{11}) is equal to:

$$S_{11} = (0.425, 0.6, 0.6, 0.825)$$

In the same way, it is possible to calculate the trapezoidal numerical weights of industry i with respect to index j (S_{ij}) for different i and j.

The trapezoidal numerical weights of different industries ($i = 1, ..., 5$) in relation to the stock liquidity index ($j = 1$) are (the outputs provided according to the analysis performed in MATLAB software):

$$\begin{aligned} S_{11} &= (0.425, 0.6, 0.6, 0.825) \\ S_{21} &= (0.225, 0.525, 0.525, 0.7) \\ S_{31} &= (0.175, 0.3, 0.30, 0.6) \\ S_{41} &= (0.4, 0.625, 0.625, 0.775) \\ S_{51} &= (0.175, 0.45, 0.45, 0.7) \end{aligned}$$

The trapezoid numerical weights of different industries ($i = 1, ..., 5$) in relation to the stock return index ($j = 2$) are (the outputs provided according to the analysis performed in MATLAB software):

$$\begin{aligned} S_{12} &= (0.3, 0.55, 0.55, 0.825) \\ S_{22} &= (0.225, 0.45, 0.45, 0.65) \\ S_{32} &= (0.425, 0.675, 0.675, 0.875) \\ S_{42} &= (0.35, 0.5, 0.5, 0.65) \\ S_{52} &= (0.1, 0.325, 0.325, 0.6) \end{aligned}$$

The trapezoidal numerical weights of different industries ($i = 1, ..., 5$) in relation to the profit per share index ($j = 3$) are (the outputs provided according to the analysis performed in MATLAB software):

$$\begin{aligned} S_{13} &= (0.425, 0.675, 0.675, 0.875) \\ S_{23} &= (0, 0.075, 0.075, 0.35) \\ S_{33} &= (0.175, 0.4, 0.4, 0.575) \\ S_{43} &= (0.4, 0.675, 0.675, 0.9) \\ S_{53} &= (0.4, 0.675, 0.675, 0.9) \end{aligned}$$

The trapezoid numerical weights of different industries ($i = 1, ..., 5$) in relation to the index of the ratio of price to profit per share ($j = 4$) are (the outputs provided according to the analysis performed in MATLAB software):

$$\begin{aligned} S_{14} &= (0.05, 0.35, 0.35, 0.575) \\ S_{24} &= (0.475, 0.675, 0.675, 0.825) \\ S_{34} &= (0.475, 0.675, 0.675, 0.825) \\ S_{44} &= (0.35, 0.6, 0.6, 0.9) \\ S_{54} &= (0.05, 0.2, 0.2, 0.475) \end{aligned}$$

The trapezoidal numerical weights of different industries ($i = 1, ..., 5$) in relation to the systematic risk index

($j = 5$) are (the outputs presented according to the analysis performed in MATLAB software):

$$S_{15} = (0.525, 0.8, 0.8, 0.95)$$
$$S_{25} = (0.05, 0.275, 0.275, 0.525)$$
$$S_{35} = (0, 0.15, 0.15, 0.4)$$
$$S_{45} = (0.225, 0.425, 0.425, 0.725)$$
$$S_{55} = (0.6, 0.85, 0.85, 1)$$

Figure 4 shows the hierarchy of censorship models for ranking industries.

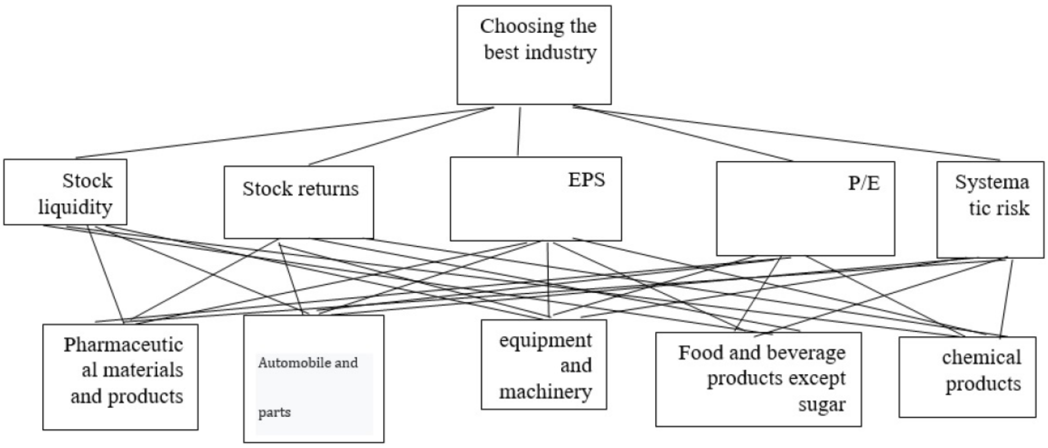

Figure 4: The hierarchy of censorship models for ranking industries

In the third step of censorship models, we must calculate the utility of censorship models of each industry $\left(F_i = \frac{1}{n}\sum_{j=1}^{n} S_{ij} \times W_j\right)$. For example, the desirability of censorship models of pharmaceutical industry and pharmaceutical products (F_1) is calculated as follows:

$$\begin{aligned} F_1 &= \frac{1}{5}(S_{11} \times W_1 + S_{12} \times W_2 + S_{13} \times W_3 + S_{14} \times W_4 + S_{15} \times W_5) \\ &= \frac{1}{5}[(0.425, 0.6, 0.6, 0.825) \times (0.6, 0.85, 0.85, 1) + (0.3, 0.55, 0.55, 0.825) \times (0.1, 0.25, 0.25, 0.55) \\ &+ (0.425, 0.675, 0.675, 0.875) \times (0, 0.15, 0.15, 0.4) + (0.05, 0.35, 0.35, 0.575) \times (0.425, 0.675, 0.675, 0.875) \\ &+ (0.525, 0.8, 0.8, 0.95) \times (0.275, 0.575, 0.575, 0.775)] = (0.0901, 0.2890, 0.2890, 0.5736) \end{aligned}$$

In the same way, the usefulness of censorship models can be calculated for different industries (Table 8).

Table 8: The desirability of censorship models for different industries (researcher's findings)

Industry	**The usefulness of censorship models**
Pharmaceutical materials and products	$F_1 = (0.0901, 0.2890, 0.2890, 0.5736)$
Automobile and parts manufacturing	$F_2 = (0.0746, 0.2368, 0.2368, 0.4653)$
Equipment and machinery	$F_3 = (0.0699, 0.2051, 0.2051, 0.4686)$
Food and beverage products	$F_4 = (0.0971, 0.2814, 0.2814, 0.5684)$
Chemical products	$F_5 = (0.0603, 0.2378, 0.2378, 0.5161)$

Finally, in the last step of censoring models, the total score of each industry (A_i) should be calculated. The total score of each industry is equal to the sum of four trapezoidal points (minimum point, two middle points and maximum point) of the desirability of censorship models. For example, the total score of the pharmaceutical industry (A_1) is calculated as follows:

$$A_1 = 0.0901 + 0.2890 + 0.2890 + 0.5736 = 1.2418$$

In the same way, the total score can be calculated for different industries (Table 9).

Table 9: total score for different industries (researcher's findings)

Industry	Total score
Pharmaceutical materials and products	1.2418
Automobile and parts manufacturing	1.0134
Equipment and machinery	0.9488
Food and beverage products	1.2283
Chemical products	1.0519

As it can be seen, the industries of materials and pharmaceutical products, food and beverage products except sugar and sugar, chemical products, automobile and parts manufacturing, and machinery and equipment have the highest total points, respectively. Therefore, the ranking of industries will be the same (Table 10).

Table 10: Ranking of industries (researcher's findings)

Industry	Rank
Pharmaceutical materials and products	1
Food and beverage products	2
Chemical products	3
Automobile and parts manufacturing	4
Equipment and machinery	5

4.2 Comparison of model performance based on censoring models and invasive weed algorithm

The summary of the results of the cuckoo optimization algorithm for the proposed model based on censoring models for different basket sizes is presented in tables 11. The utility function, in fact, is the objective function of the model, which has an additional part that is considered as a penalty for constraint 2 (budget constraint) so that in addition to maximizing the objective function (maximizing the return and minimizing the variance of the portfolio), the budget constraint is also balanced. On the other hand, the demand was based on the assumption that at least 99.8% of the available budget should be spent in the proposed portfolio. As can be seen, the budget limit has been met for different basket sizes. Therefore, according to the fulfillment of the budget constraint, as can be seen, for different basket sizes, the value of the objective function has not changed much and is stable. Therefore, it can be concluded that in the proposed model, the size of the portfolio has a very small effect on the value of the objective function, and the investor can invest by choosing a desired portfolio size.

Table 11: Summary of the results of the proposed model based on censoring models for different basket sizes

Basket size	5	7	10	12	15	17	20
The value of the utility function	0.2092	0.2055	0.2044	0.2013	0.2067	0.2073	0.2
Percentage of budget spent	99.9339	99.9867	99.8338	99.9597	99.9951	99.9908	99.9968
Portfolio return	0.4226	0.4133	0.414	0.414	0.4145	0.4162	0.4002
Portfolio variance	0.0028	0.0019	0.0018	0.0031	0.0009	0.0013	0.0002
Standard deviation (risk) of the portfolio	0.0528	0.0435	0.042	0.0559	0.0304	0.0365	0.0139
The value of the objective function	0.2099	0.2057	0.2061	0.2017	0.2068	0.0028	0.2

In order to test the research hypotheses regarding the significance of the difference in the performance of the proposed model based on censorship models compared to the performance of the invasive weed algorithm, the two-sample t-test is used. In this section, we use this test to compare the results of the proposed model based on censorship models and the invasive weed algorithm for different portfolio sizes, which we explain below. For this purpose, the summary of the invasive weed algorithm results for different basket sizes is presented in Table 12.

Table 12: Summary of invasive weed algorithm results for different basket sizes

Basket size	5	7	10	12	15	17	20
The value of the utility function	0.197	0.1929	0.1889	0.1843	0.1895	0.1948	0.1831
Percentage of budget spent	99.8923	99.9834	99.932	92.7136	99.8737	99.861	99.8188

Portfolio return	0.321	0.3409	0.3087	0.3357	0.3262	0.2962	0.3424
Portfolio variance	0.0029	0.002	0.0018	0.0032	0.0009	0.0014	0.0002
Standard deviation (risk) of the portfolio	0.0545	0.0493	0.0444	0.057	0.0324	0.0409	0.0153
The value of the objective function	0.1977	0.1917	0.194	0.1855	0.1949	0.1897	0.1837

The test results related to the comparison of the average of the proposed model based on the censoring models and the invasive weed algorithm

A- At the level of the 5-share portfolio

In order to test the equality of the average of two communities (the average efficiency and variance of the proposed model based on censorship models and the invasive weed algorithm), it is necessary to first examine the variance test of the two communities. In other words, the test of equality of variances precedes the test of equality of means:

$H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2$

Table 13: independent samples test of the average yield of two samples (basket of 5)

		Levene's test for equality of variances		**t-test equality of means**						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									lower	upper
Average portfolio return	Equal variances assumed	7.704	0.041	−3.121	35	0.013	−0.0931	0.1067	−0.13671	−0.04951
	Equal variances not assumed			−2.909	28.868	0.021	−0.0931	0.1134	−0.12782	−0.0584

Examining the results of table 13 shows that, because the value of Levene's test is F=7.704 at the 5% error level, it is smaller than 5% (sig=0.041). Therefore, we will use the second line of the t-test in the following investigations. The results of the test show that the value of the t statistic (t = -2.909) is greater than -2 and its significance level (sig = 0.021) is less than 5%.

Therefore, we reject the hypothesis H_0 that there is no significant difference between the average efficiency of the model based on the censoring models and the invasive weed algorithm, and we accept the hypothesis H_1. Also, due to the negativity of the upper and lower limits of the two groups of companies, the average difference of the two communities will be less than zero, in which case the average performance of the model based on censorship models is higher than the average of the invasive weed algorithm.

Table 14: Average risk test of two samples (basket of 5)

		Levene's test for equality of variances		**t-test equality of means**						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									lower	upper
Average portfolio risk	Equal variances assumed	9.121	0.033	−2.121	35	0.0048	0.1087	0.1356	0.0734	0.12154
	Equal variances not assumed			−2.432	28.868	0.041	0.1087	0.1423	0.06432	0.17871

Examining the results of table 14 shows that, because the value of Levene's test is F=9.121 at the 5% error level, it is smaller than 5% (sig=0.033). Therefore, we will use the second line of the t-test in the following investigations. The results of the test show that the value of the t statistic ($t = |-2.432|$) is greater than $|-2|$ and its significance level (sig = 0.033) is less than 5%. Therefore, we reject the hypothesis H_0 that there is no significant difference between the average variance (risk) of the model based on censorship models and the invasive weed algorithm and accept the hypothesis H_1. Also, due to the positivity of the upper and lower limits of the two groups of companies, the average difference of the two communities will be less than zero, in which case the average risk (variance) of the portfolio of the model based on censorship models is lower than the average of the invasive weed algorithm.

B- At the level of the 7-share portfolio

Table 15: Average return test of two societies (basket of 7)

		Levene's test for equality of variances		t-test equality of means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									lower	upper
Average portfolio return	Equal variances assumed	10.611	0.011	−4.509	35	0.0054	−0.1098	0.08723	−0.109432	−0.0267
	Equal variances not assumed			−3.602	28.868	0.0087	−0.1098	0.1423	−0.16231	−0.0321

Examining the results of table 15 shows that, because the value of Levene's test is F=10.611 at the 5% error level, it is smaller than 5% (sig=0.011). Therefore, we will use the second line of the t-test in the following investigations. The test results show that the value of t statistic ($t = |-3.602|$) is greater than $|-2|$ and its significance level (sig=0.0087) is less than 5%. Therefore, we reject the hypothesis H_0 that there is no significant difference between the average efficiency of the model based on the censoring models and the invasive weed algorithm, and we accept the hypothesis H_1. Also, due to the negativity of the upper and lower limits of the two groups of companies, the average difference of the two communities will be less than zero, in which case the average performance of the model based on censorship models is higher than the average of the invasive weed algorithm.

Table 16: Average risk test of two societies (basket of 7)

		Levene's test for equality of variances		t-test equality of means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									lower	upper
Average portfolio risk	Equal variances assumed	1.612	0.342	−2.509	35	0.032	0.0891	0.1732	0.0643	0.11543
	Equal variances not assumed			−2.623	28.868	0.029	0.0891	0.1289	0.0791	0.11765

Examining the results of table 16 shows that the value of Levene's test F=1.612 at the 5% error level is greater than 5% (sig=0.342). Therefore, we use the first line of the t-test in the following investigations. The results of the test show that the value of the t statistic (t = -2.509) is greater than -2 and its significance level (sig = 0.032) is less than 5%. Therefore, we reject the hypothesis H_0 that there is no significant difference between the average variance (risk) of the model based on censorship models and the invasive weed algorithm and accept the hypothesis H_1. Also, due to the positivity of the upper and lower limits of the two groups of companies, the average difference of the two communities will be less than zero, in which case the average risk (variance) of the portfolio of the model based on censorship models is lower than the average of the invasive weed algorithm.

C- At the level of the 10-share portfolio

Examining the results of table 17 shows that, because the value of Levene's test is F=121.8 at the 5% error level, it is smaller than 5% (sig=0.048). Therefore, we will use the second line of the t-test in the following investigations. The results of the test show that the value of the t statistic (t = -3.656) is greater than -2 and its significance level (sig = 0.012) is less than 5%. Therefore, we reject the hypothesis H_0 that there is no significant difference between the average efficiency of the model based on the censoring models and the invasive weed algorithm, and we accept the hypothesis H_1. Also, due to the negativity of the upper and lower limits of the two groups of companies, the average difference of the two communities will be less than zero, in which case the average performance of the model based on censorship models is higher than the average of the invasive weed algorithm.

Examining the results of table 18 shows that the value of Levene's test F=1.986 is greater than 5% at the error level of 5% (sig=0.071). Therefore, we use the first line of the t-test in the following investigations. The results of the test show that the value of the t statistic ($t = |-2.543|$) is greater than $|-2|$ and its significance level (sig = 0.013)

Table 17: Average return test of two samples (basket of 10)

		Levene's test for equality of variances		**t-test equality of means**						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									lower	upper
Average portfolio return	Equal variances assumed	8.121	0.048	−4.432	35	0.010	−0.0889	0.11543	−0.15876	−0.06432
	Equal variances not assumed			−3.656	28.868	0.012	−0.0889	0.11765	−0.16098	−0.07231

is less than 5%. Therefore, we reject the hypothesis H_0 that there is no significant difference between the average variance (risk) of the model based on censorship models and the invasive weed algorithm and accept the hypothesis H_1. Also, due to the positivity of the upper and lower limits of the two groups of companies, the average difference of the two communities will be less than zero, in which case the average risk (variance) of the portfolio of the model based on censorship models is lower than the average of the invasive weed algorithm.

Table 18: Average risk test of two societies (basket of 10)

		Levene's test for equality of variances		**t-test equality of means**						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									lower	upper
Average portfolio risk	Equal variances assumed	1.986	0.071	−2.543	35	0.013	0.13309	0.11709	0.0563	0.17951
	Equal variances not assumed			−2.813	28.868	0.021	0.13421	0.11511	0.0612	0.1884

D- At the level of the 12-share portfolio

Table 19: Test of the average efficiency of two samples (basket of 12)

		Levene's test for equality of variances		**t-test equality of means**						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									lower	upper
Average portfolio return	Equal variances assumed	12.907	0.0017	−2.913	35	0.021	0.1087	0.13675	−0.11205	−0.0567
	Equal variances not assumed			−2.618	28.868	0.032	0.1087	0.12785	−0.10712	−0.0412

Examining the results of table 19 shows that the value of Levene's test = 12.907 F at the 5% error level is smaller than 5% (sig = 0.0017). Therefore, we will use the second line of the t-test in the following investigations. The results of the test show that the value of the t statistic (t = -2.618) is greater than -2 and its significance level (sig = 0.032) is less than 5%. Therefore, we reject the hypothesis H_0 that there is no significant difference between the average efficiency of the model based on the censoring models and the invasive weed algorithm, and we accept the hypothesis H_1. Also, due to the negativity of the upper and lower limits of the two groups of companies, the average difference of the two communities will be less than zero, in which case the average performance of the model based on censorship models is higher than the average of the invasive weed algorithm.

Examining the results of table 20 shows that, because the value of Levene's test is F=2.129 at the 5% error level, it is greater than 5% (sig=0.197). Therefore, we use the first line of the t-test in the following investigations. The results of the test show that the value of the t statistic ($t = |-4.187|$) is greater than $|-2|$ and its significance level (sig = 0.00097) is less than 5%. Therefore, we reject the hypothesis H_0 that there is no significant difference between the average variance (risk) of the model based on censorship models and the invasive weed algorithm and accept the

Table 20: Average risk test of two samples (basket of 12)

		Levene's test for equality of variances		t-test equality of means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									lower	upper
Average portfolio risk	Equal variances assumed	2.129	0.197	−4.187	35	0.00097	0.0931	0.1067	0.07689	0.15897
	Equal variances not assumed			−3.087	28.868	0.016	0.0931	0.1134	0.06976	0.13998

hypothesis H_1. Also, due to the positivity of the upper and lower limits of the two groups of companies, the average difference of the two communities will be less than zero, in which case the average risk (variance) of the portfolio of the model based on censorship models is lower than the average of the invasive weed algorithm.

E- At the level of the 15-share portfolio

Table 21: Average return test of two samples (basket of 15)

		Levene's test for equality of variances		t-test equality of means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									lower	upper
Average portfolio return	Equal variances assumed	9.126	0.027	−2.887	35	0.029	0.11982	0.0954	−0.10732	−0.06921
	Equal variances not assumed			−3.194	28.868	0.021	0.11982	0.0913	−0.10745	−0.07231

Examining the results of table 21 shows that, because the value of Levene's test is F=9.126 at the 5% error level, it is smaller than 5% (sig=0.027). Therefore, we will use the second line of the t-test in the following investigations. The test results show that the value of t statistic (t=-3.194) is greater than -2 and its significance level (sig=0.021) is less than 5%. Therefore, we reject the hypothesis H_0 that there is no significant difference between the average efficiency of the model based on the censoring models and the invasive weed algorithm, and we accept the hypothesis H_1. Also, due to the negativity of the upper and lower limits of the two groups of companies, the average difference of the two communities will be less than zero, in which case the average performance of the model based on censorship models is higher than the average of the invasive weed algorithm.

Table 22: Average risk test of two samples (basket of 15)

		Levene's test for equality of variances		t-test equality of means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									lower	upper
Average portfolio risk	Equal variances assumed	1.134	0.461	−2.472	35	0.039	0.10911	0.08341	0.13671	0.04951
	Equal variances not assumed			−2.853	28.868	0.028	0.10911	0.08341	0.12782	0.0584

Examining the results of table 22 shows that, because the value of Levene's test is F=1.134 at the 5% error level, it is greater than 5% (sig=0.461). Therefore, we use the first line of the t-test in the following investigations. The results of the test show that the value of the t statistic ($t = |-2.472|$) is greater than $|-2|$ and its significance level (sig = 0.039) is less than 5%. Therefore, we reject the hypothesis H_0 that there is no significant difference between the average variance (risk) of the model based on censorship models and the invasive weed algorithm and accept the hypothesis H_1. Also, due to the positivity of the upper and lower limits of the two groups of companies, the average difference of the two communities will be less than zero, in which case the average risk (variance) of the portfolio of the model based on censorship models is lower than the average of the invasive weed algorithm.

F- At the level of the 17-share portfolio

Table 23: Test of the average efficiency of two samples (basket of 17)

		Levene's test for equality of variances		t-test equality of means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									lower	upper
Average portfolio return	Equal variances assumed	9.654	0.031	−3.987	35	0.013	−0.0992	0.1298	−0.1176	−0.06781
	Equal variances not assumed			−3.165	28.868	0.019	−0.0992	0.1298	−0.1176	−0.07911

Examining the results of table 23 shows that, because the value of Levene's test is F=9.654 at the 5% error level, it is smaller than 5% (sig=0.031). Therefore, we will use the second line of the t-test in the following investigations. The results of the test show that the value of the t statistic ($t = |-3.165|$) is greater than $|-2|$ and its significance level (sig $= 0.019$) is less than 5%. Therefore, we reject the hypothesis H_0 that there is no significant difference between the average efficiency of the model based on the censoring models and the invasive weed algorithm, and we accept the hypothesis H_1. Also, due to the negativity of the upper and lower limits of the two groups of companies, the average difference of the two communities will be less than zero, in which case the average performance of the model based on censorship models is higher than the average of the invasive weed algorithm.

Table 24: Average risk test of two societies (basket of 17)

		Levene's test for equality of variances		t-test equality of means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									lower	upper
Average portfolio risk	Equal variances assumed	1.502	0.186	−3.197	35	0.013	0.1078	0.10654	0.11765	0.13897
	Equal variances not assumed			−2.711	28.868	0.021	0.1078	0.0976	0.11481	0.13411

Examining the results of table 24 shows that the value of Levene's test F=1.502 at the 5% error level is greater than 5% (sig=0.186). Therefore, we use the first line of the t-test in the following investigations. The results of the test show that the value of the t statistic ($t = |-3.917|$) is greater than $|-2|$ and its significance level (sig $= 0.013$) is less than 5%. Therefore, we reject the hypothesis H_0 that there is no significant difference between the average variance (risk) of the model based on censorship models and the invasive weed algorithm and accept the hypothesis H_1. Also, due to the positivity of the upper and lower limits of the two groups of companies, the average difference of the two communities will be less than zero, in which case the average risk (variance) of the portfolio of the model based on censorship models is lower than the average of the invasive weed algorithm.

G- At the level of the 20-share portfolio

Examining the results of table 25 shows that, because the value of Levene's test is F=14.804 at the 5% error level, it is smaller than 5% (sig=0.0074). Therefore, we will use the second line of the t-test in the following investigations. The results of the test show that the value of the t statistic ($t = |-3.398|$) is greater than $|-2|$ and its significance level (sig $= 0.0092$) is less than 5%. Therefore, we reject the hypothesis H_0 that there is no significant difference between the average efficiency of the model based on the censoring models and the invasive weed algorithm, and we accept the hypothesis H_1. Also, due to the negativity of the upper and lower limits of the two groups of companies, the average difference of the two communities will be less than zero, in which case the average performance of the model based on censorship models is higher than the average of the invasive weed algorithm.

Examining the results of table 26 shows that, because the value of Levene's test is F=1.134 at the error level of 5%, it is greater than 5% (sig=0.461). Therefore, we use the first line of the t-test in the following investigations. The

Table 25: Average return test of two samples (basket of 20)

		Levene's test for equality of variances		**t-test equality of means**						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									lower	upper
Average portfolio return	Equal variances assumed	14.804	0.0074	−2.693	35	0.033	−0.08431	0.11421	−0.10451	−0.06931
	Equal variances not assumed			−3.398	28.868	0.0092	−0.08431	0.10811	−0.10891	−0.007464

Table 26: Average risk test of two societies (basket of 20)

		Levene's test for equality of variances		**t-test equality of means**						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
									lower	upper
Average portfolio risk	Equal variances assumed	1.134	0.461	−2.472	35	0.039	0.10911	0.08341	0.13671	0.04951
	Equal variances not assumed			−2.853	28.868	0.028	0.10911	0.08341	0.12782	0.0584

results of the test show that the value of the t statistic ($t = |-2.472|$) is greater than $|-2|$ and its significance level (sig $= 0.039$) is less than 5%. Therefore, we reject the hypothesis H_0 that there is no significant difference between the average variance (risk) of the model based on censorship models and the invasive weed algorithm and accept the hypothesis H_1. Also, due to the negativity of the upper and lower limits of the two groups of companies, the average difference of the two communities will be less than zero, in which case the average risk (variance) of the portfolio of the model based on censorship models is lower than the average of the invasive weed algorithm.

5 Conclusion

According to the results of tables 13 to 26, as can be seen, all the stock portfolios formed (with different sizes) by the model based on censoring models have higher returns and lower variance (risk) than the invasive weed algorithm, which shows the superior performance of the proposed Behtz model compared to the invasive weed algorithm, thus confirming the first hypothesis (the extracted indices have a significant effect on stock selection) and the second hypothesis of the research (the model based on censoring models performs better than the invasive weed algorithm) at a 95% confidence level. Examining the results of the implementation of censored models and the meta-heuristic algorithm for portfolio optimization, shows that the censored models and the meta-heuristic algorithm used have a very good performance in portfolio optimization.

Therefore, the results of the implementation of censored models and meta-heuristic algorithm show the very good performance of this network in portfolio optimization. In this research, the censored model and meta-heuristic algorithm are proposed in the form of a non-linear multi-objective integer programming model mixed with cardinal constraints, threshold limit, investment sector, entropy and considering the transaction cost. The problem model has a mixed structure. Therefore, according to the NP-hard feature of the problem, the meta-heuristic algorithm of harmony search with the Pareto approach is used to solve the model. The research findings fill the study gap in investment portfolio optimization and also show that the proposed model has been able to consider the investment portfolio selection conditions well and determine an optimal investment portfolio.

References

[1] R. Aghamohammadi, R. Tehrani, and M. Khademi, *Investigating the effect of study period selection on solving portfolio optimization based on different risk criteria using meta-heuristic algorithms*, J. Financ. Manag. Persp. **12** (2022), no. 37.

[2] M.A. Akbay, C.B. Kalayci, and O. Polat, *A parallel variable neighborhood search algorithm with quadratic programming for cardinality constrained portfolio optimization*, Knowledge-Based Syst. **198** (2020), 105944.

[3] O. Ertenlice and C.B. Kalayci, *A survey of swarm intelligence for portfolio optimization: Algorithms and applications*, Swarm Evol. Comput. **39** (2018), 36–52.

[4] N. Grishina, C.A. Lucas, and P. Date, *Prospect theory–based portfolio optimization: An empirical study and analysis using intelligent algorithms*, Quant. Finance **17** (2017), no. 3, 353–367.

[5] C. John, *High speed hill climbing algorithm for portfolio optimization*, Tanzania J. Sci. **47** (2021), no. 3, 1236–1242.

[6] C.B. Kalayci, O. Ertenlice, and M.A. Akbay, *A comprehensive review of deterministic models and applications for mean-variance portfolio optimization*, Expert Syst. Appl. **125** (2019), 345–368.

[7] C.B. Kalaycı, Ö. Ertenlice, H. Akyer, and H. Aygören, *A review on the current applications of genetic algorithms in mean-variance portfolio optimization*, Pamukkale Univ. J. Eng. Sci. **23** (2017), 470–476.

[8] A.T. Khan, X. Cao, S. Li, B. Hu, and V.N. Katsikis, *Quantum beetle antennae search: A novel technique for the constrained portfolio optimization problem*, Sci. China Info. Sci. **64** (2021), no. 5, 1–14.

[9] Y. Kim, D. Kang, M. Jeon, and C. Lee, *GAN-MP hybrid heuristic algorithm for non-convex portfolio optimization problem*, Eng. Econ. **64** (2019), no. 3, 196–226.

[10] H.M. Markowitz, *Foundations of portfolio theory*, J. Finance **46** (1991), no. 2, 469–477.

[11] S.S. Meghwani and M. Thakur, *Multi-objective heuristic algorithms for practical portfolio optimization and rebalancing with transaction cost*, Appl. Soft Comput. **67** (2018), 865–894.

[12] B.A. Mercangöz and E. Eroglu, *The genetic algorithm: an application on portfolio optimization*, Research anthology on multi-industry uses of genetic programming and algorithms, IGI Global, 2021, pp. 790–810.

[13] D.A. Milhomem and M.J.P. Dantas, *Analysis of new approaches used in portfolio optimization: A systematic literature review*, Production **30** (2020).

[14] M.J. Naik and A.L. Albuquerque, *Hybrid optimization search-based ensemble model for portfolio optimization and return prediction in business investment*, Prog. Artif. Intell. **11** (2022), no. 4, 315–331.

[15] M. Rahmani, M. Khalili Eraqi, and H. Nikoomaram, *Portfolio optimization by means of meta heuristic algorithms*, Adv. Math. Financ. Appl. **4** (2019), no. 4, 83–97.

[16] I.B. Salehpoor and S. Molla-Alizadeh-Zavardehi, *A constrained portfolio selection model at considering risk-adjusted measure by using hybrid meta-heuristic algorithms*, Appl. Soft Comput. **75** (2019), 233–253.

[17] E.P. Setiawan, *Comparing bio-inspired heuristic algorithm for the mean-CVaR portfolio optimization*, J. Phys.: Conf. Ser. IOP Pub. **1581** (2020), no. 1, 012014.

[18] J. Wang, *A novel firefly algorithm for portfolio optimization problem*, IAENG Int. J. Appl. Math. 49 (2019), no. 1, 1–6.

[19] Z. Wang, X. Zhang, Z. Zhang, and D. Sheng, *Credit portfolio optimization: A multi-objective genetic algorithm approach*, Borsa Istanbul Rev. **22** (2022), no. 1, 69–76.

[20] M. Zanjirdar, *Overview of portfolio optimization models*, Adv. Math. Financ. Appl. 5 (2020), no. 4, 419–435.

[21] W. Zhou, W. Zhu, Y. Chen, and J. Chen, *Dynamic changes and multi-dimensional evolution of portfolio optimization*, Econ. Res. Ekon. Istraž. **35** (2021), no. 1, 1–26.