An integrated fuzzy-grey relational analysis approach to portfolio optimization

Mukesh Kumar Mehlawat[1] · Pankaj Gupta[1,2] · Ahmad Zaman Khan[1]

Abstract

This paper combines two approaches (Fuzzy set theory and Grey Relational Analysis) for modelling an investor's imprecise linguistic expectations and the uncertain returns of assets. We propose a novel maximization-type risk measure capable of incorporating the investor's individual preferences. The investor provides the expectations of what is considered the "ideal" return from the portfolio. We use Credibility theory to capture the investors' subjective and imprecise expectations in a precise mathematical form. We construct a portfolio return sequence using the assets' actual return data and an ideal sequence based on investors' preferences. Subsequently, we calculate the Grey similitude and the closeness incidence degree between the two sequences. The closer the portfolio return is to the ideal return, the better. In this manner, we develop a new risk measure that can quantify an investor's perception of risk. This measure is intuitive and easy to calculate. It does not involve estimating many parameters, something which would increase the estimation risk. We use a genetic algorithm to solve the resulting portfolio optimization model. We illustrate this method with two case studies: (i) a case study of 100 assets of the U.S. stock market's NASDAQ-100 index and (ii) a case study of 50 assets of the Indian stock market's NIFTY-50 index. We comprehensively analyze the model's out-of-sample performance and discuss its implications. The portfolios obtained using the proposed approach exhibit healthy growth outside the in-sample period. We also compare the out-of-sample performance of the proposed model with several approaches in the literature to establish its superiority.

Keywords Portfolio optimization · Credibility theory · Grey incidence analysis · Multi-objective programming · Fuzzy set theory · Genetic algorithm

✉ Pankaj Gupta
  pankajgpta@gmail.com

  Mukesh Kumar Mehlawat
  mukesh0980@yahoo.com

  Ahmad Zaman Khan
  ahmad.du.or.17@gmail.com

[1] Department of Operational Research, University of Delhi, New Delhi, 110007, India

[2] Present address: Flat No. - 01, Kamayani Kunj Apartment, Indra Prastha Extension, Delhi, 110092, India

1 Introduction

A stock market is an elaborate organization that facilitates firms to raise capital through the issuance of *stocks*. A stock generally represents fractional ownership of a business. The stock market also enables the trading of stocks. A typical investor in any stock market endeavours to create a *portfolio* of assets with a maximum possible investment return with little variability. However, the prices of assets are hard to predict, with many random and human factors affecting them, including company-specific factors, like the imminent outlook of a company's business plan, general factors like the supply and demand of stocks, the social and economic laws of the country, etc. Common wisdom dictates that investors diversify their capital over many assets so that losses in some assets can be offset by profits in others. Thus, there are various aspects of portfolio optimization, viz., the investor's individual preferences and the volatilities of the stock market, which one must consider before

deciding on the allocation of capital over different assets. Mathematically, this problem, with its various conflicting objectives, is called *multi-objective portfolio optimization*. Initially, Markowitz [1] formulated a mathematical model of portfolio optimization. Numerous authors have improved the original mean-variance portfolio optimization model (proposed by Markowitz) to make it more relevant to stock markets' real-world scenarios. Speranza [2] employed mean absolute semi-deviation (MASD), and Konno and Yamazaki [3] used mean absolute deviation (MAD) to measure the risk and transformed the quadratic mean-variance model into a linear programming problem.

Different investors have diverse perceptions of the current scenarios prevalent in the stock market. This is vital to the facilitation of asset trading in the stock market. The contrasting investor attitudes are the prominent reason for asset transactions. Therefore, any portfolio optimization model must incorporate investor attitudes. Various authors have explored the inclusion of investor attitude into the portfolio optimization model. Li and Yi [4] have proposed coherent fuzzy numbers in a possibilistic environment to model the asset returns. An adaptive index, l, distinguishes between optimistic, pessimistic, and neutral investors. Interested readers can refer to [5–7], where a portfolio optimization model that incorporates investor attitude is proposed.

This paper proposes a novel maximization-type risk measure for selecting the optimal portfolio for an investor. The investor provides inputs about the expectations of the returns from the portfolio (which can be understood as a target return). These expectations are generally imprecise and linguistic. Therefore, we use Fuzzy set theory to capture these subjective expectations in a precise mathematical form. The target return is expressed as a fuzzy number and is called the *ideal portfolio return*. We use fuzzy simulation to generate a sequence of returns based on the investor's fuzzy ideal returns. This sequence is the "ideal" sequence, which the portfolio sequence should emulate. We calculate the portfolio sequence using the asset weights and the historical return data. We then calculate the Grey similitude and closeness incidence degree between the ideal sequence and the portfolio sequence. The closer the portfolio sequence is to the ideal, the less risky the portfolio is for that investor. Therefore, the objective is to maximize the weighted average of the closeness and the similitude degree. Our method of modelling the risk takes care of only the downside deviations because no rational investor would suggest losses (i.e., negative returns) as ideal portfolio behaviour.

Numerous approaches in the literature have tackled the portfolio optimization problem. A natural question is: what is the need for combining Fuzzy set theory and Grey Systems theory to obtain an optimal portfolio? The success of a portfolio optimization model depends heavily on the accuracy of the prediction of the asset returns. Asset prices are affected by a plethora of factors. Therefore, any endeavour to completely describe such a system is not easy. It may even render the entire exercise meaningless [8]. In such a scenario, the portfolio optimization model can benefit from the experience and intuition of the experts. The model should provide ways to integrate the advice and knowledge of experts in the optimization process. In our previous work [5, 6], we have used the mean absolute semi-deviation (MASD) to estimate the portfolio risk. Although MASD works well, we feel that the investor attitude should play a role in the estimation of the risk. The proposed approach aims to achieve this objective. The investor or the expert provides inputs on the ideal portfolio return in any given scenario. It is a problem of cognitive uncertainty since most investors don't know the optimal course of action. Therefore, we use fuzzy set theory to model the ideal portfolio return (i.e., the level of return that the investor aims to achieve from the portfolio). The next task is to calculate the risks involved in securing the ideal return. We cannot use the existing risk measures such as the variance or MASD in our model (since we only have a fuzzy estimate of the target return). One could argue that we can use the fuzzy ideal return's expected value to calculate the variance or MASD, but we would again be using crisp values for estimating the risk. This would defeat the purpose of incorporating the subjective preferences of the investors in assessing the risk of the portfolio. As explained above, asset returns are affected by several factors. A complete enumeration of all these factors is not possible. However, the returns themselves reflect the total effect of these factors. For instance, if an asset's returns are positive, for whatever reasons, this must imply that the demand for the asset exceeds its supply, i.e., its prospects must be bullish, and vice versa. Therefore, modelling the risk of a portfolio can be simplified if we focus our attention on asset returns. On account of this objective, we propose a risk measure that estimates the closeness and similarity of the portfolio returns with the ideal returns. Grey Systems theory [9] effectively deals with problems of "small sets and poor information" [8]. It focuses on excavating partially known information using the possibility function and sequence operators. It uncovers the hidden patterns in any system for its accurate understanding and description. One advantage of Grey Systems theory is that it constructs models with small amounts of data. Grey Incidence Analysis uses the degree of similarity of geometric curves of the available data sequences (in this case, the asset returns). Therefore, we propose a way of using fuzzy simulation and Grey Incidence Analysis to define a risk measure that quantifies the similarity and the closeness between the ideal returns and

M. K. Mehlawat et al.

those of the portfolio. The proposed risk measure is intuitive and easy to understand. It simply gives a measure of the point-to-point closeness and similarity between the ideal returns and the returns of the portfolio. It does not require the estimation of many parameters, something which would compound errors [10]. The next step is to integrate the investor attitude in the portfolio optimization model. We provide two different methods to accomplish this task using fuzzy set theory. The final step is to determine the optimal asset allocations (i.e., the optimal portfolio weights). We apply a genetic algorithm (a machine learning technique) for this purpose. To summarize, the proposed approach combines human intelligence and artificial intelligence to determine the optimal portfolio that works for the investor. We believe that no model in the literature quantifies the investor's perception of risk. Consequently, the investor or the expert makes, in the proposed method, an increased contribution in selecting the best portfolio.

We summarize the paper's novelty and key highlights as follows: (i) We propose a novel maximization-type risk measure based on the Grey similitude and closeness incidence degree between the portfolio returns and the ideal returns (as preferred by the investor). This is a novel contribution to the portfolio optimization problem, to the best of our knowledge. (ii) An aspect of portfolio optimization is that risk is a perception-based, individual-specific concept. The same asset may be considered risky by one investor, while it may be considered less risky by another. From the above discussion, we can conclude that the proposed risk measure can incorporate investor risk perceptions. The optimistic, pessimistic, and neutral investors can specify the ideal portfolio behaviour, modelled using the appropriate fuzzy numbers. (iii) Two separate approaches to integrate investor attitude in the portfolio optimization model using Fuzzy set theory have been discussed. (iv) We provide a method to simulate coherent fuzzy numbers using probability theory. In particular, we modify the triangular probability distribution to generate sequences based on coherent triangular fuzzy numbers. (v) We add several constraints, namely bound and cardinality constraints, in the model to make it more relevant to the practical requirements of investors. We apply the proposed model to a case study involving the 100 assets of the U.S. stock market's NASDAQ-100 index. We use a genetic algorithm to solve the proposed model. We also perform another case study of the 50 assets listed in the Indian stock market's NIFTY-50 index to further establish the validity of the proposed model. (vi) We demonstrate the out-of-sample performance of the optimal portfolios and discuss its implications. (vii) Škrinjarić [11] used Grey Incidence Analysis to rank the assets based on the first four moments, viz. expected return, variance, skewness, and kurtosis. These rankings were subsequently used to construct the portfolios using 22 predefined investment strategies.

Our model calculates the optimal asset allocations based on historical data using the techniques of multi-criteria decision-making. Thus, our model relies on optimization techniques rather than predefined investment strategies. Our approach combines Fuzzy set theory, Grey Systems theory, and multi-criteria decision-making to provide optimal portfolios based on investors' preferences.

The rest of this paper is organized as follows: Section 2 provides a brief literature review of the recent studies in portfolio optimization. Section 3 introduces the basic concepts used throughout the paper. In Section 4, we formulate the mathematical model for portfolio optimization, which employs the proposed risk measure. We also develop a genetic algorithm to solve the proposed model. Section 5 provides two case studies: one involving the 100 assets of the U.S. stock market's NASDAQ-100 index, and the second involving the 50 assets of the Indian stock market's NIFTY-50 index. We also conduct a comprehensive analysis of the out-of-sample performance of the obtained optimal portfolios and discuss their implications. Finally, Section 6 is the conclusion section with remarks on the possible extensions of this paper.

2 Literature review

An essential ingredient in the success of any mathematical model for obtaining optimal portfolios is an accurate representation of the stock market's real-life scenarios. As discussed above, asset prices are highly volatile and subject to various human-made and random factors. Thus, the mathematical model for portfolio optimization must be flexible enough to capture the uncertainty inherent in the stock market. One method is to model the asset returns as random variables. There are many studies where the asset returns have been modelled as random variables. Markowitz used a normal distribution to model the asset returns. However, various authors have challenged this assumption [16, 17]. Asset returns exhibit extreme values more frequently than predicted by the normal distribution [18]. This prompted the authors to explore other methods of modelling asset returns. Kamali et al. [19] employed various probability distributions to model asset returns and select the best-fitting distribution for the return data. They employ the MATLAB solver to solve the optimization model. They discover that the kernel distribution is the most appropriate for modelling the asset returns. However, their study is not exhaustive, and there could be other distributions that are a better fit for the return data. Akbay et al. [20] proposed a parallel variable neighbourhood search algorithm for cardinality constrained portfolio optimization problems. They propose a Variable Neighbourhood Search (VNS) algorithm to find the best combination of assets. The

asset proportions are determined using Quadratic programming. The algorithm's computational complexity is $O(K^2 (poolsize_L - K))$, where K is the number of assets held in the portfolio, and $poolsize_L$ is the number of solutions in the local search phase. The proposed algorithm provides superior results on five datasets (Hang Seng, DAX100, FTSE100, S&P100, and Nikkei) over various competing algorithms. However, their algorithm is based on the Mean-variance (M-V) framework. We have already discussed the disadvantages of the M-V model. Yu et al. [21] used Autoregressive Integrated Moving Average (ARIMA) to forecast the asset returns. They study the importance of incorporating return forecasting in the portfolio optimization process. Through various models (Mean-variance, Mean-absolute deviation, and so on), they establish that portfolio models with return forecasting outperform their corresponding counterparts (without return forecasting) by 54%–61%. However, the success of their model is heavily dependent on the accuracy of the forecast. Other works include [12, 22–26]. Modelling asset returns as random variables, although popular, requires a significant amount of data for fitting probability distributions with the requisite accuracy in the estimation. It may not always be practical, especially for newly-listed assets. The variations in asset returns are also affected by various human-made factors, which are not random. In such a scenario, it is more prudent to use the judgments of stock market experts. However, these opinions and judgments are usually qualitative rather than quantitative, hence incorporating them into mathematical models is difficult. Fuzzy set theory [27, 28] is an effective method to formulate imprecise judgments in a mathematical form. It is a popular method for modelling the uncertainty in portfolio optimization problems [13, 14, 29–33]. Credibility theory [34, 35], which employs a credibility measure to determine the chance of an event's occurrence, has the axiomatic properties of self-duality and normality. It is a consistent method for modelling investors' imprecise and linguistic judgments. Many researchers have used credibility theory to model volatile asset returns [36–42].

Multi-objective programming determines the optimal tradeoff between return and risk in portfolio optimization. It is an essential branch of Operations Research that transcends disciplines in its application. It has applications in finance, accounting, marketing, supply chain management, and various other fields. Tirkolaee et al. [43] use multi-objective optimization for a reliable pollution routing with cross-dock selection. They propose a bi-objective mixed-integer linear programming problem with cost and supply reliability as the objectives. Two multi-objective metaheuristic algorithms are used to solve the model: Non-dominating Sorting Genetic algorithm-II (NSGA-II) and Multi-Objective Simulated Annealing (MOSA). The computational complexity of NSGA-II is $O(mN^2)$, where N is the size of the population, and m is the number of objectives. They discovered that the objective functions are heavily dependent on the demand; thus, the management can evaluate the required resources by analyzing the demand fluctuations. In [44], the authors study the problem of allocation and scheduling rescue units during a natural disaster. They formulate a bi-objective mixed-integer linear programming model to simultaneously minimize the total time to complete all the rescue operations and the total delay time. Multi-choice goal programming was used to solve the model. The model was validated using several comparative studies. Reference [45] is a study of medical waste management during the COVID-19 pandemic. The authors develop a multi-objective mixed linear programming model for a multi-trip location routing problem with time windows. They use fuzzy chance-constrained programming to study the uncertainty of the demand parameter. Based on their case study, the total trip time is reported to be 19.733 hours using three vehicles. The runtime of the solution algorithm is 1029.739 seconds. Thus, we can conclude from the above discussion that multi-objective programming has a wide variety of applications in various branches of Operations Research.

Julong Deng developed Grey Systems theory [9] for dealing with uncertain systems with partially known information. It deals with insufficient information by "generating and extracting useful information from what is available" [8]. In this manner, the operational behaviour of an uncertain system can be effectively monitored and described. There are many instances of small samples and poor information in the real world. Therefore, Grey Systems theory can be applied to a wide variety of problems. Various authors have used the Grey Systems theory to model the uncertainty of asset returns [11, 46–48]. As discussed above, asset returns are affected by different kinds of factors. The mutual interaction of these factors determines the behaviour of the asset's returns. Statistical methods are usually employed for analyzing asset returns. However, these methods suffer from various weaknesses [8]: (i) Large data samples are needed for reliable results. (ii) The available data is assumed to follow a probability distribution, and factors are assumed to have linear relationships. These requirements are hard to satisfy. When historical data is limited, applying the traditional statistical methods is challenging, as a small amount of data is not usually compatible with a conventional probability distribution. Grey Incidence Analysis is a credible alternative to analyze a system when statistical methods are impractical. It applies to large and small samples and involves simple computations. The idea is "to use the degree of similarity of geometric curves formed by the available data sequences" [8]. The higher the degree, the more similarity between the curves and vice-versa [8]. It may seem that Grey Systems

theory is similar to Fuzzy set theory; however, as Liu et al. [8] mentioned, "Fuzzy mathematics emphasizes the investigation of problems with cognitive uncertainty, where research objects possess the characteristic of clear intension and unclear extension." For instance, "tall man" is a fuzzy concept because everybody vaguely understands a "tall man". However, suppose we want to evaluate the particular range within which every person is tall, and those outside it are not tall. In that case, we will have many difficulties because "tall man" has an unclear extension. This is a problem of cognitive uncertainty, which is resolved using individual judgment (i.e., using the membership function).

The method to measure the risk involved in an investment is another important characteristic of the portfolio optimization model. A quantitative estimate of a portfolio's prospective losses helps the investor assess its suitability (vis-á-vis returns). The variance was used to measure the portfolio's risk in the original mean-variance model (Markowitz [1]). However, the variance is known to be a two-sided measure of risk, i.e., it penalizes both upside and downside deviations from the expected return. Only the downside deviations are of concern for the investor; most investors would, in fact, *like* the upside deviations. Nevertheless, the variance is a prevalent risk measure among researchers [4, 15, 23, 49]. *Downside risk measures*, which penalize only the negative deviations from the expected return, are more realistic for the portfolio optimization problem. Mean absolute semi-deviation (MASD), Value-at-Risk (VaR), and Conditional Value-at-Risk (CVaR) are some examples of downside risk measures. Each has its advantages and disadvantages. Refer to [50] for an explanation of these. Downside risk measures have also been popular amongst authors [5–7, 42, 51, 52]. We have compared the existing literature approaches in Table 1 based on the above discussion.

3 Preliminaries

This section provides a primer on the basic concepts used in the paper. The basic theory of each concept is presented in the appendix.

3.1 A brief introduction to coherent Fuzzy Numbers

Definition 1 (Mehlawat et al. [6]) "The credibility function of a coherent triangular fuzzy number $\xi_l = (d - \delta, d, d + \eta)_l$, $v_l(z) : \mathbb{R} \to [0, 1]$, is given by

$$v_l(z) = \begin{cases} \frac{1}{2} \left(\frac{z - d + \delta}{\delta} \right)^{\frac{1}{l}}, & \text{if } d - \delta \leq z \leq d, \\ \frac{1}{2} \left(\frac{d + \eta - z}{\eta} \right)^l, & \text{if } d \leq z \leq d + \eta, \\ 0, & \text{otherwise}, \end{cases} \quad (1)$$

where $l \in \mathbb{R}, l > 0$ (See Fig. 1)."

Remark 1 The adaptive index l specifies the stock market perception of an investor (see Fig. 1). When $1 < l < \infty$, the credibility is higher for the left spread ($d - \delta \leq z \leq d$), while it is lower for the right spread ($d \leq z \leq d + \eta$) in comparison to a standard triangular fuzzy number (since $0 \leq v_l(z) \leq 1$). Thus, if the investors assess the current stock market prospects to be pessimistic, they can assign l in $[1, \infty]$. When $0 < l < 1$, the credibility is lower for the left spread ($d - \delta \leq z \leq d$), while it is higher for the right spread ($d \leq z \leq d + \eta$) in comparison to a standard triangular fuzzy number (since $0 \leq v_l(z) \leq 1$). Thus, if the investors assess the current stock market prospects to be optimistic, they can assign l in $(0, 1)$. For $l = 1$, the coherent fuzzy number reduces to a standard triangular fuzzy number with similar credibilities assigned to both the spreads. Therefore, a neutral investor can assign $l = 1$. See Fig. 2 for visual explanation of this concept.

Proposition 1 *"Suppose that $\xi_{1,l} = (d_1 - \delta_1, d_1, d_1 + \eta_1)_l$, $\xi_{2,l} = (d_2 - \delta_2, d_2, d_2 + \eta_2)_l$ are two coherent triangular fuzzy numbers. Then, for $\kappa > 0$, we have*

$$\xi_{1,l} + \xi_{2,l} = ((d_1 + d_2) - (\delta_1 + \delta_2), d_1 + d_2, (d_1 + d_2)$$
$$+ (\eta_1 + \eta_2))_l, \quad (Addition)$$
$$\kappa \xi_{1,l} = (\kappa d_1 - \kappa \delta_1, \kappa d_1, \kappa d_1 + \kappa \eta_1)_l.$$
$$(Scalar\ multiplication)"$$

Proposition 2 (Credibilistic expectation [6]) *"Let $\xi_l = (d - \delta, d, d + \eta)_l$ be a coherent fuzzy number. Then, its expected value is given by*

$$E[\xi_l] = d + \frac{1}{2(l + 1)} (\eta - l\delta)." \quad (2)$$

3.2 Grey incidence analysis

Proposition 3 *"Let $Z_i = (z_i(1), z_i(2), \ldots, z_i(n))$ be the data sequence of a system's behaviour, $Z_i - z_i(1)$ denote the zigzag line $(z_i(1) - z_i(1), z_i(2) - z_i(1), \ldots, z_i(n) - z_i(1))$, and*

$$s_i = \int_1^n (Z_i - z_i(1)) \, dt,$$

when Z_i increases, $s_i \geq 0$, when Z_i decreases, $s_i \leq 0$, and when Z_i vibrates, the sign of s_i varies."

Proposition 4 *"Assume that the images of the zero-starting point of two behavioural sequences Z_i and Z_j are respectively, $Z_i^0 = (z_i^0(1), z_i^0(2), \ldots, z_i^0(n))$ and $Z_j^0 = (z_j^0(1), z_j^0(2), \ldots, z_j^0(n))$. Let,*

$$s_i - s_j = \int_1^n \left(Z_i^0 - Z_j^0 \right) dt,$$
$$S_i - S_j = \int_1^n (Z_i - Z_j) \, dt,$$

Table 1 Comparison with existing methodologies

Feature	[11]	[4]	[12]	[13]	[14]	[7]	[15]	Proposed Approach
Mean	✓	✓	✓	✓	✓	✓	✓	✓
Risk measure	Variance	Variance	VaR	Minimax	Variance	MASD	Variance	GSD and GCD
Environment	Random	Possibilistic	Random	Uncertain	Possibilistic	Crediblistic	Uncertain random	Credbilistic
Investor attitude	×	✓	×	×	×	✓	×	✓
MCDM techniques	×	✓	✓	✓	✓	✓	✓	✓
Solution method	N.A.	NLP	Improved NSGA-II	NLP	DE	PBEMO	NLP	GA
Bound constraints	×	×	✓	✓	✓	✓	×	✓
No Short-selling	×	✓	✓	✓	✓	✓	✓	✓
Cardinality constraint	×	×	✓	×	✓	✓	×	✓
Fuzzy simulation	×	×	×	×	×	×	×	✓

Acronyms: VaR: Value-at-risk; MASD: Mean absolute semi-deviation; GSD: Grey similitude incidence degree ; GCD: Grey closeness incidence degree; N.A.: Not applicable; DE: Differential evolution; NLP: Non-linear programming; NSGA-II: Nondominated sorting genetic algorithm II; PBEMO: Preference-based evolutionary multi-objective optimization; GA: Genetic algorithm

when Z_i^0 is entirely above Z_j^0, $s_i - s_j \geq 0$, when Z_i^0 is entirely below Z_j^0, $s_i - s_j \leq 0$, and when Z_i^0 and Z_j^0 alternate their positions, the sign of $s_i - s_j$ is not fixed. Similarly, the sign of $S_i - S_j$ can be discussed."

Proposition 5 (Liu [53]) *"Assume that Z_i and Z_j are 1-time interval sequences (i.e., the time difference between each observation is 1) of the same length, and the following are the zero-starting point images of Z_i and Z_j,*

$$Z_i^0 = (z_i^0(1), z_i^0(2), \ldots, z_i^0(n)).$$
$$Z_j^0 = (z_j^0(1), z_j^0(2), \ldots, z_j^0(n)),$$

then,

$$|s_i| = \left| \sum_{k=2}^{n-1} z_i^0(k) + \frac{1}{2} z_i^0(n) \right|.$$

$$|s_j| = \left| \sum_{k=2}^{n-1} z_j^0(k) + \frac{1}{2} z_j^0(n) \right|.$$

$$|s_i - s_j| = \left| \sum_{k=2}^{n-1} \left(z_i^0(k) - z_j^0(k) \right) + \frac{1}{2} \left(z_i^0(n) - z_j^0(n) \right) \right|.$$

$$|S_i| = \left| \frac{1}{2} z_i(1) + \sum_{k=2}^{n-1} z_i(k) + \frac{1}{2} z_i(n) \right|.$$

$$|S_j| = \left| \frac{1}{2} z_j(1) + \sum_{k=2}^{n-1} z_j(k) + \frac{1}{2} z_j(n) \right|.$$

$$\left| S_i - S_j \right| = \left| \frac{1}{2} (z_i(1) - z_j(1)) + \sum_{k=2}^{n-1} \left(z_i(k) - z_j(k) \right) \right.$$
$$\left. + \frac{1}{2} \left(z_i(n) - z_j(n) \right) \right|."$$

Fig. 1 Coherent fuzzy numbers' credibility distribution for different values of l
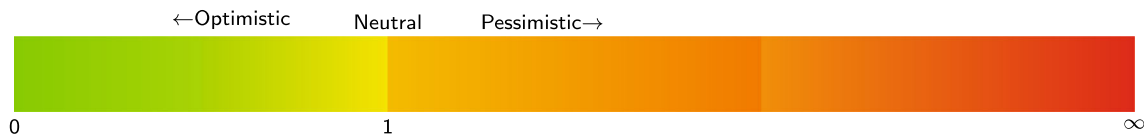

(a) $l < 1$

(b) $l > 1$
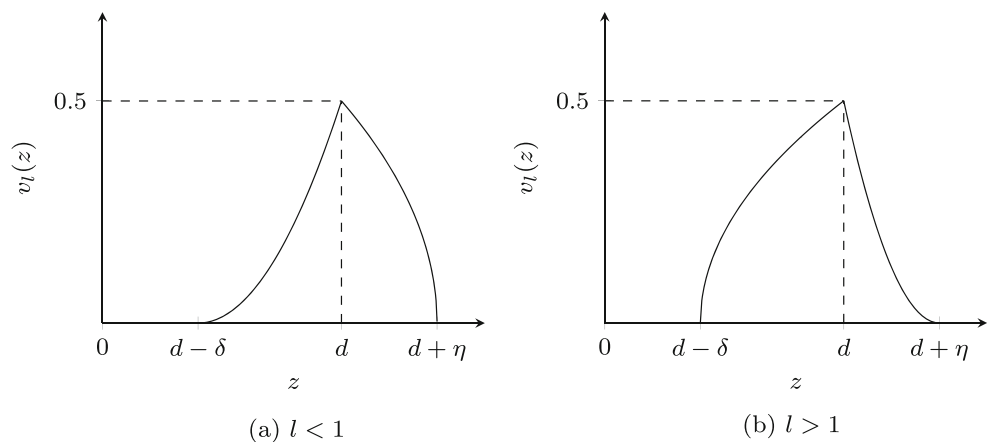
Fig. 2 Change in investor attitude with l

Definition 2 (Liu and Xie [54]) "Let Z_i and Z_j be sequences of identical length, and $s_i - s_j$ be as defined in Proposition 5. Then the Grey similitude incidence degree is given by,

$$\varepsilon_{ij} = \frac{1}{1 + \left|s_i - s_j\right|}."\qquad(3)$$

The similitude degree of Grey incidence measures the geometric similarity of the shapes of the sequences Z_i and Z_j. A greater value of ε_{ij} indicates a higher similarity of the geometric shapes of Z_i and Z_j and vice-versa

Remark 2 Some properties of the similitude degree are as follows,

1. $0 \le \varepsilon_{ij} \le 1$.
2. ε_{ij} is determined by the geometric shape of the sequences Z_i and Z_j and is independent of the relative spatial positions of the sequences. Therefore, any translation of sequences Z_i and Z_j does not affect ε_{ij}.
3. $\varepsilon_{ii} = \varepsilon_{jj} = 1$, and $\varepsilon_{ij} = \varepsilon_{ji}$.

Definition 3 (Liu and Xie [54]) "Let Z_i and Z_j be sequences of identical length, and $S_i - S_j$ be as defined in Proposition 5. Then the Grey closeness incidence degree is given by,

$$\rho_{ij} = \frac{1}{1 + \left|S_i - S_j\right|}."\qquad(4)$$

Closeness degree of Grey incidence measures the spatial closeness of sequences Z_i and Z_j. A greater the value of ρ_{ij} indicates that the sequences Z_i and Z_j are spatially closer and vice versa.

Remark 3 Some properties of the closeness degree are as follows,

1. $0 \le \rho_{ij} \le 1$.
2. The value of ρ_{ij} depends on both the geometric shapes of the sequences X_i and X_j and their relative spatial positions. Thus, any translation of the sequences will change the value of ρ_{ij}.
3. $\rho_{ii} = \rho_{jj} = 1$, and $\rho_{ij} = \rho_{ji}$.

3.3 Fuzzy simulation

In the proposed approach, the investor provides the expectations on the "ideal" portfolio behaviour. Since the investor

expectations will most likely be imprecise and linguistic, we use a credibilistic triangular fuzzy number to capture the expectations in a mathematical form. To apply Grey Incidence Analysis, we need two sequences to calculate the closeness and the similitude degree. The portfolio sequence can be readily constructed using the assets' historical data and the portfolio's asset weights. Next, we need the "ideal" sequence, i.e., the sequence that follows the credibility distribution of the triangular fuzzy number given by the investor. The portfolio sequence should emulate the ideal sequence as closely as possible.

3.3.1 Generating sequences of credibilistic triangular fuzzy numbers

Our task is to generate a random sequence of numbers distributed according to a given credibility function (i.e., the given ideal triangular fuzzy number). However, this is not a straightforward process as credibility functions do not obey the laws of probability. Therefore, we use a triangular probability distribution to generate the ideal sequence. In this manner, the shape of the generated sequence will be similar to that of a triangular credibility function (See Fig. 3).

Definition 4 A random variable $Y \in [a, c]$ follows triangular distribution if it has the following probability density ($P(.)$) and cumulative distribution function ($F(.)$),

$$P(Y = y) = \begin{cases} 0, & \text{if } y \le a, \\ \frac{2(y-a)}{(c-a)(b-a)}, & \text{if } a < y \le b, \\ \frac{2(c-y)}{(c-a)(c-b)}, & \text{if } b < y \le c, \\ 0, & y > c. \end{cases}$$

$$F(Y \le y) = \begin{cases} 0, & \text{if } y \le a, \\ \frac{(y-a)^2}{(c-a)(b-a)}, & \text{if } a < y \le b, \\ 1 - \frac{(c-y)^2}{(c-a)(c-b)}, & \text{if } b < y \le c, \\ 1, & \text{otherwise}, \end{cases}\qquad(5)$$

where b is the mode of the distribution.

We summarize the procedure for generating the ideal sequence $< I_s >= (i_s(1), i_s(2), \dots, i_s(L))$ for a triangular fuzzy number (a, b, c) in the following steps (L is the historical data length),

Step 1: For $t = 1$ to L, do;
Step 2: Generate $r \in [0, 1]$ randomly.
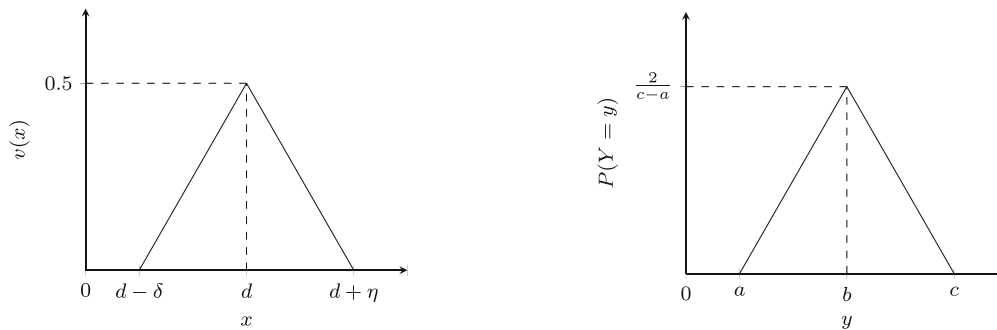
(a) Credibility distribution of a triangular fuzzy number

(b) Probability density function of a triangular distribution

Fig. 3 Comparison of triangular credibility and probability functions

Step 3: If $0 \leq r \leq \frac{b-a}{c-a}$, then $i_s(t) = a + \sqrt{r(c-a)(b-a)}$.

Step 4: If $\frac{b-a}{c-a} \leq r \leq 1$, then $i_s(t) = c - \sqrt{(1-r)(c-a)(c-b)}$.

The basic idea here is to calculate $F^{-1}(r)$. The above process is repeated until the sequence's requisite length has been achieved (i.e., the historical data length).

3.3.2 Generating sequences of credibilistic coherent triangular fuzzy numbers

We need to modify the probability density function of the triangular distribution (see Definition 4) such that it retains the properties of a probability function and has a similar shape, as shown in Fig. 1. The probability density function can be modified as follows,

$$
P_l(Y = y) = \begin{cases} 0, & \text{if } y \leq a, \\ \frac{(l+1)(y-a)^{\frac{1}{l}}}{(c+b(l-1)-al)(b-a)^{\frac{1}{l}}}, & \text{if } a \leq y \leq b, \\ \frac{(l+1)(c-y)^l}{(c+b(l-1)-al)(c-b)^l}, & \text{if } b \leq y \leq c, \\ 0, & y > c. \end{cases} \tag{6}
$$

Proposition 6 *The function defined in (6) is a probability density function for all $l > 0$.*

Proof It is a straightforward exercise to show that $\int_{-\infty}^{\infty} P_l(y) dy = 1 \;\; \forall l > 0$. Therefore, we only need to show that $c + b(l-1) - al > 0 \;\; \forall l$. Since $c > b \Rightarrow c + b(l-1) > b + b(l-1) \Rightarrow c + b(l-1) > bl \Rightarrow c + b(l-1) > al$ (since $b > a$) $\Rightarrow c + b(l-1) - al > 0$, and the proof is complete. $\square$

Thus, P_l is a probability density function. The cumulative distribution function of P_l is given by,

$$
F_l(Y \leq y) = \begin{cases} 0, & \text{if } y \leq a, \\ \frac{l(y-a)^{\frac{l+1}{l}}}{(c+b(l-1)-al)(b-a)^{\frac{1}{l}}}, & \text{if } a \leq y \leq b, \\ 1 - \frac{(c-y)^{l+1}}{(c+b(l-1)-al)(c-b)^l}, & \text{if } b \leq y \leq c, \\ 1, & \text{otherwise.} \end{cases} \tag{7}
$$

Thus, the ideal sequence $< I_{s,l} > = (i_{s,l}(1), i_{s,l}(2), \ldots, i_{s,l}(L))$ for the coherent triangular fuzzy number $(a, b, c)_l$ can be generated using the following steps (L is the historical data length),

Step 1: For $t = 1$ to L, do;

Step 2: Generate $r \in [0, 1]$ randomly.

Step 3: If $0 \leq r \leq \frac{l(b-a)}{c+b(l-1)-al}$, then $i_{s,l}(t) = a + \left(\frac{r(c+b(l-1)-al)(b-a)^{\frac{1}{l}}}{l} \right)^{\frac{l}{l+1}}$.

Step 4: If $\frac{l(b-a)}{c+b(l-1)-al} \leq r \leq 1$, then $i_{s,l}(t) = c - ((1-r)(c+b(l-1)-al)(c-b)^l)^{\frac{1}{l+1}}$.

As a check, we plot the frequency distribution of the generated sequence for the triangular fuzzy number $(1, 2, 5)$, for $l = 0.3, 2$ using the proposed method to verify if the generated sequences are indeed following the required distribution. We have kept the length of the generated sequence sufficiently high to ensure convergence. Figure 4 establishes the validity of the proposed method (see Fig. 1 for a comparison).

4 Methodology

4.1 A mathematical formulation of the portfolio optimization model

Suppose that the asset universe comprises N assets ($i = 1, 2, \ldots, N$). The investor aims to invest the available capital in these assets optimally. Asset returns are modelled using credibilistic triangular fuzzy numbers. We discuss two methodologies for incorporating investor attitude into the model.

– **Using standard triangular fuzzy numbers**
  In this method, the pessimistic, optimistic, and neutral investors provide different fuzzy numbers for ideal return, each based on their market assessment. The asset returns are also modelled using standard triangular

fuzzy numbers. Suppose that $\xi_i = (d_i - \delta_i, d_i, d_i + \eta_i)$ is the triangular fuzzy number that represents the ith asset's return. Let x_i be the proportion of allocation of capital in the ith asset. The portfolio return is modelled as, $\sum_{i=1}^{N} x_i \xi_i = \left(\sum_{i=1}^{N} x_i d_i - x_i \delta_i, \sum_{i=1}^{N} x_i d_i, \sum_{i=1}^{N} x_i d_i + x_i \eta_i \right)$.

– **Using coherent triangular fuzzy numbers**
If it is not possible to provide different fuzzy numbers for ideal return due to the nature of the historical data (i.e., past performance of the assets), then the coherent triangular fuzzy numbers can be used to provide the ideal return for the optimistic, pessimistic, and the neutral investors. The coherent fuzzy numbers' adaptive index, l, can be used to incorporate the investor attitude (as explained in Section 3.1). Suppose that $\xi_{i,l} = (d_i - \delta_i, d_i, d_i + \eta_i)_l$ is the coherent triangular fuzzy number that represents the return of the ith asset. By Proposition 1, the portfolio return is modelled as, $\sum_{i=1}^{N} x_i \xi_{i,l} = \left(\sum_{i=1}^{N} x_i d_i - x_i \delta_i, \sum_{i=1}^{N} x_i d_i, \sum_{i=1}^{N} x_i d_i + x_i \eta_i \right)_l$.

Remark 4 We can integrate the investor attitude into the ideal portfolio return by setting appropriate fuzzy numbers. When using coherent fuzzy numbers for ideal return, the adaptive index l is used for investor attitude (an optimistic investor sets $0 < l < 1$, a pessimistic investor sets $1 < l < \infty$, and a neutral investor sets $l = 1$). When using standard triangular fuzzy numbers, a typical optimistic investor focused on return maximization will choose a higher fuzzy number with a larger variance. Thus, the generated ideal sequence will also exhibit a larger variance. Therefore, the resulting portfolio will be riskier. On the other hand, a pessimistic investor focused on risk minimization will choose a lower fuzzy number with a smaller variance.

After we get the fuzzy numbers for the ideal return, we use the method described in Section 3.3 to generate the ideal sequence $< I_{s,l} >= \left(i_{s,l}(1), i_{s,l}(2), \ldots, i_{s,l}(L) \right)$, where L is the length of the available historical data. When using standard triangular fuzzy numbers for modelling, the case of $l = 1$ will suffice. Therefore, in the sequel, asset returns are modelled using coherent fuzzy numbers. We now formulate the model's objectives and constraints.

4.1.1 Model objectives

– **Portfolio's expected return**
This objective maximizes the portfolio's expected return. It is modelled as follows:

$$\text{Maximize} \quad E\left[\sum_{i=1}^{N} x_i \xi_{i,l} \right]$$

$$= \sum_{i=1}^{N} x_i d_i + \frac{1}{2(l+1)} \left(\sum_{i=1}^{N} x_i \eta_i - l \sum_{i=1}^{N} x_i \delta_i \right).$$

– **Weighted Grey closeness and similitude incidence degree**
This objective minimizes the portfolio's risk. We use a weighted sum of grey closeness and similitude incidence degree to measure the risk. Suppose that $< I_{s,l} >= \left(i_{s,l}(1), i_{s,l}(2), \ldots, i_{s,l}(L) \right)$ is the sequence generated from the fuzzy number representing the ideal return, and that $< P_s >= \left(p_s(1), p_s(2), \ldots, p_s(L) \right)$ is the sequence of the portfolio return. As discussed above, the portfolio return is calculated from the historical data of assets and the asset allocations, $x_i, \quad i = 1, 2, \ldots, N$. The Grey closeness incidence degree is calculated from (4) as,

$$\rho_{IP} = \frac{1}{1 + |S_I - S_P|} = \frac{1}{1 + \left| \frac{1}{2}(i_{s,l}(1) - p_s(1)) + \sum_{t=2}^{L-1} \left(i_{s,l}(t) - p_s(t) \right) + \frac{1}{2} \left(i_{s,l}(L) - p_s(L) \right) \right|}, \tag{8}$$

where L is the length of the historical data available. The Grey similitude incidence degree is calculated from (3) as,

$$\varepsilon_{IP} = \frac{1}{1 + |s_I - s_P|} = \frac{1}{1 + \left| \sum_{t=2}^{L-1} \left(i_{s,l}{}^0(t) - p_s{}^0(t) \right) + \frac{1}{2} \left(i_{s,l}{}^0(L) - p_s{}^0(L) \right) \right|}, \tag{9}$$

where $i_{s,l}{}^0(t) = i_{s,l}(t) - i_{s,l}(1)$, $p_s{}^0(t) = p_s(t) - p_s(1)$. As discussed before, the portfolio sequence should emulate the ideal sequence as closely as possible. Therefore, the portfolio with a low closeness and similitude incidence degree with the ideal sequence is considered more risky from the investor's perspective.

Thus, the objective to minimize the risk of the portfolio is formulated as

Maximize $\quad CLSIM = w_1 \rho_{IP} + w_2 \varepsilon_{IP}$,

where $CLSIM$ is the weighted average of the Grey similitude and closeness incidence degree, and w_1, w_2

Fig. 4 Frequency distribution of the sequences obtained using the proposed method

(a) $l = 0.3$

(b) $l = 2$
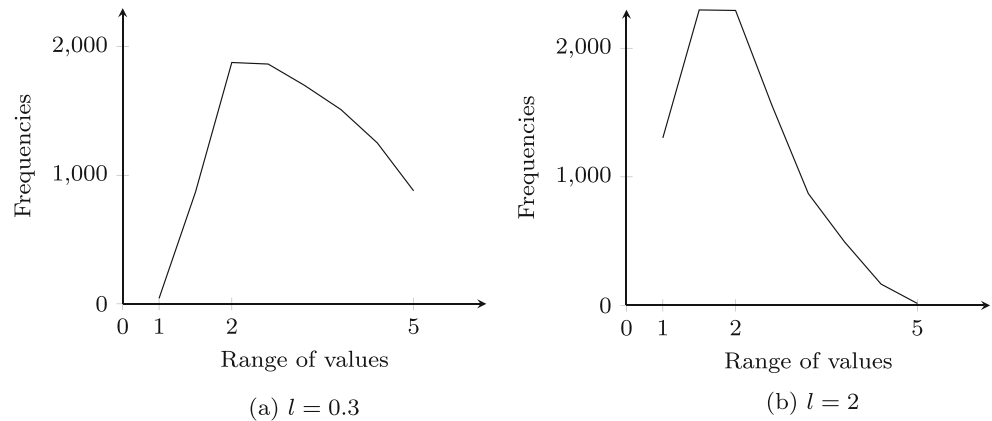
are weights representing the relative importance of Grey closeness and similitude incidence degree, respectively.

Remark 5 The concept of maximizing the risk measure may alarm the investors. It is important to note that the proposed risk measure ($CLSIM$) is different from standard risk measures. Since most risk measures either estimate the volatility of returns or the expected losses, the endeavour to minimize such risk measures is easy to understand. On the other hand, the proposed risk measure estimates the closeness and similarity of the portfolio return and ideal return (i.e., the investor's target return). A high value of $CLSIM$ indicates that the portfolio return is closer to the ideal return, and thus, is less risky from the investor's perspective. Therefore, the objective to minimize the risk of the portfolio is to maximize $CLSIM$ in our method.

4.1.2 Model constraints

– **Cardinality constraint**
The investor may require a certain level of diversification in the portfolio. We define a binary variable, y_i, $i = 1, 2, \ldots, N$ as

$$y_i = \begin{cases} 1, & \text{if the } i^{th} \text{ asset is included in the portfolio,} \\ 0, & \text{otherwise.} \end{cases}$$

Then the cardinality constraint to ensure at least K assets in the portfolio, is formulated as

$$\sum_{i=1}^{N} y_i \geq K. \tag{10}$$

– **Total investment constraint**
The constraint requires the investment of the entire available capital. Thus, this constraint is formulated as

$$\sum_{i=1}^{N} x_i = 1. \tag{11}$$

– **Upper bound on investment constraint**
This constraint places an upper bound, u_i, on the ith asset's capital allocation. Then

$$x_i < y_i u_i \quad \forall i = 1, 2, \ldots, N, \tag{12}$$

where y_i is as defined above.

– **Lower bound on investment constraint**
This constraint places a lower bound, l_i, on the ith asset's capital allocation. Then

$$x_i > l_i y_i \quad \forall i = 1, 2, \ldots, N, \tag{13}$$

where y_i is as defined above.

– **Binary variable constraint**
y_i can only take values 0 or 1. Thus,

$$y_i \in \{0, 1\} \quad \forall i = 1, 2, \ldots, N. \tag{14}$$

The portfolio optimization model is subsequently formulated as

$$\text{Max } Z_1 = E\left[\sum_{i=1}^{N} x_i \xi_{i,l}\right]$$

$$\text{Max } Z_2 = CLSIM = w_1 \rho + w_2 \varepsilon$$

subject to (10)–(14). (M1)

Model ($M1$) is a multi-objective optimization problem. The next logical step is to devise a method to solve the multi-objective optimization problem. A typical multi-objective optimization problem involves a trade-off between its objectives. Thus, the endeavour to improve one objective's value may lead to a degradation in other objectives' values. The primary objective in such a scenario is to find *non-dominated* solutions.

Definition 5 Suppose that F is the set of all the feasible solutions to a multi-objective programming problem. Let f_j, $j = 1, 2, \ldots, M$, be the objectives. Without loss of generality, assume that f_j is a maximization-type objective for each $j = 1, 2, \ldots, M$. A solution $x^* \in F$ is said to be

a non-dominated solution to the multi-objective problem if and only if there does not exist any $x \in F$ such that,

$$f_j(x) \geq f_j(x^*) \ \forall \ j = 1, 2, \ldots, M. \ j \neq b, \ \text{and},$$
$$f_b(x) > f_b(x^*), \ \text{for at least one } b, \ 1 \leq b \leq M.$$

Several literature approaches convert the multi-objective problem to single-objective problem, such as weighted sum and ϵ-constraint methods. These methodologies *guarantee* non-dominated solutions to the multi-objective optimization problem under suitable assumptions. We use the ϵ-constraint method in this paper.

Consider the multi-objective programming problem ($MOPP$),

Minimize $f_1(x)$

Minimize $f_2(x)$

$$\vdots$$

Minimize $f_k(x)$

subject to $x \in S$,

where S is the set of feasible solutions of the problem and $x \in \mathbb{R}^n$ is an n-dimensional vector of decision variables. In the ϵ-constraint method, a single-objective problem is constructed from the multi-objective problem by treating all but one of the objectives in the constraints. A maximization-type objective is converted to a $\geq$-type constraint, while the minimization-type objective is converted to a $\leq$-type constraint. We discuss only the minimization case here, as the maximization-type objective can be converted to a minimization-type by multiplying it with -1. Typically, the essential objectives are treated in the constraints to control their optimization. Thus the ϵ-constraint formulation of the $MOPP$ is given by,

Minimize $\quad f_1(x)$

subject to $\quad f_2(x) \leq \epsilon_2$

$$\vdots$$

$$f_k(x) \leq \epsilon_k$$

$$x \in S.$$

The threshold values for the objectives (i.e., ϵ_i, $i = 2, 3, \ldots, k$) in the constraints can be calculated by solving the single-objective problem while ignoring the remaining objectives. For our situation, the investor can provide the threshold values for the objectives in the constraints. A flow chart of the proposed method is shown in Fig. 5.

An investor usually has several targets associated with the investments. These targets can easily be expressed in a numerical form (for instance, return expectation, risk tolerance, etc.). The ϵ-constraint method has the advantage that it can directly incorporate the investor's goals. Other methods like the weighted sum approach require that the
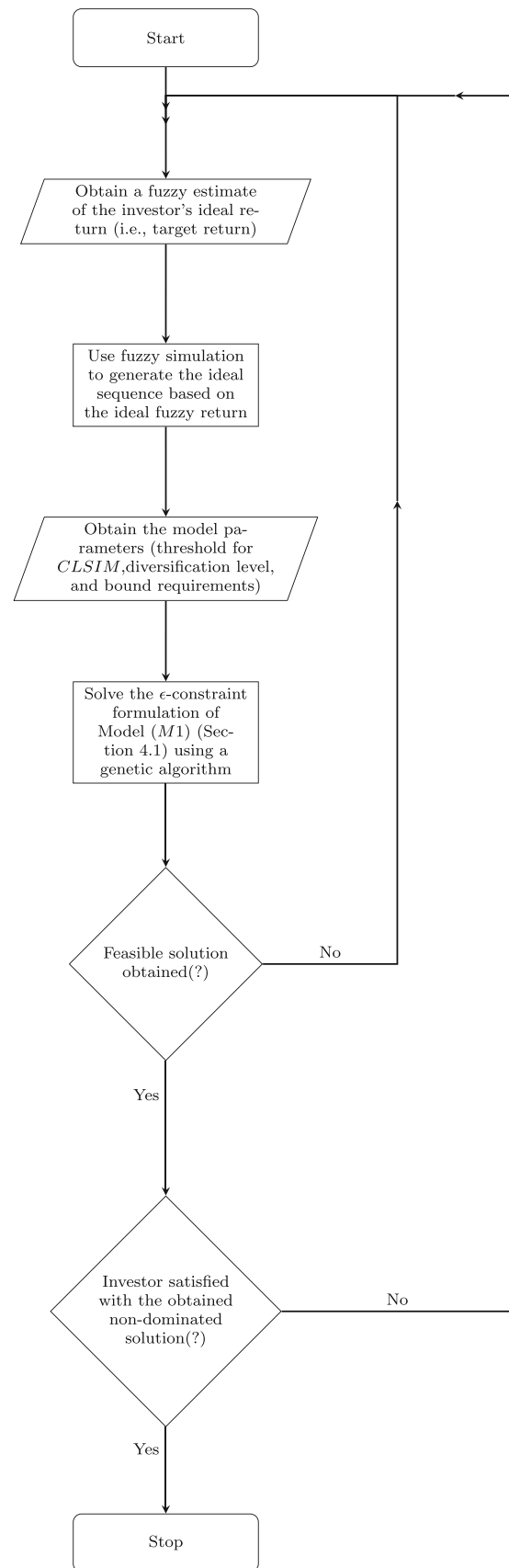

Fig. 5 Flow chart of the proposed solution method

investor specify the relative preferences of the objectives in a numerical form. However, it may not always be possible for the investor to numerically specify the extent to which one objective is preferred over others. *No-preference methods* are another class of methods for solving a multi-objective optimization problem that do not consider the decision-maker's preferences. Such methods aim to minimize the distance of the objective function vector from an appropriate reference point. However, it does not fill the bill here as investor preferences play an important role in portfolio optimization. *Interactive methods* are the most advanced class of methods for solving a multi-objective optimization problem. In such a method, the analyst or an interactive computer program tries to determine the exact preferences of the decision-maker iteratively. First, an initial feasible solution is calculated and presented to the decision-maker. If the decision-maker is satisfied with the solution, the procedure is stopped; otherwise, the process continues until a solution is found that satisfies the decision-maker's preferences. If we observe Fig. 5, we can immediately deduce that we have used a hybrid of interactive and the ϵ-constraint method for our problem.

4.2 Solution methodology

We begin this section by illustrating the calculation of $CLSIM$ using an example.

Example 1 Suppose the fuzzy ideal return of an investor is given by, $(0.0005, 0.0015, 0.0035)$. Suppose the length of the available historical data is 10 ($L = 10$). Using the method described in Section 3.3, we generate 10 observations based on the credibility distribution of $(0.0005, 0.0015, 0.0035)$. Let us assume $l = 2.5$. Table 2 summarizes the calculations. We have, $a = 0.0005$, $b = 0.0015$, $c = 0.0035$. Thus

$$\frac{l(b-a)}{c+b(l-1)-al} = 0.55556.$$

Consider the first random number in Table 2 ($r = 0.56056$). Since $r > \frac{l(b-a)}{c+b(l-1)-al}$, we have

$$i_{s,l}(1) = c - \left((1-r)(c+b(l-1)-al)(c-b)^l\right)^{\frac{1}{l+1}}$$
$$= 0.001507.$$

The second random number in Table 2 is 0.041383. Since $r < \frac{l(b-a)}{c+b(l-1)-al}$, we have

$$i_{s,l}(2) = a + \left(\frac{r(c+b(l-1)-al)(b-a)^{\frac{1}{l}}}{l}\right)^{\frac{l}{l+1}}$$
$$= 0.000656.$$

Table 2 Generation of the fuzzy ideal sequence

Random Number, r	Ideal sequence, $< I_{s,l} >$
0.560656	0.001507
0.041383	0.000656
0.387799	0.001274
0.054567	0.000691
0.220832	0.001017
0.868099	0.002087
0.050264	0.00068
0.945982	0.002405
0.990356	0.002831
0.988708	0.0028

The rest of the ideal sequence can be generated in an similar manner. Now, let us assume that there are only two assets (A_1 and A_2). We compare the values of $CLSIM$ for three different portfolios, namely, $x_A = \{0.5, 0.5\}$, $x_B = \{0.3, 0.7\}$, $x_C = \{0.8, 0.2\}$. The 10-period historical return data and the return sequence generated by the three portfolios (i.e., the portfolio sequences) have been shown in Table 3. Now that we have the ideal and the portfolio sequences, we can use (8) and (9) to calculate the Grey similitude and closeness incidence degrees. Finally, using $w_1 = 0.8$, $w_2 = 0.2$, $CLSIM$ for the three portfolios can be calculated as follows,

$$\rho_{IP}(A) = 0.972, \varepsilon_{IP}(A) = 0.8902,$$
$$CLSIM(A) = 0.8 * 0.972013 + 0.2 * 0.890188 = 0.9557,$$
$$\rho_{IP}(B) = 0.9749, \varepsilon_{IP}(B) = 0.9004,$$
$$CLSIM(B) = 0.8 * 0.974921 + 0.2 * 0.900437 = 0.96,$$
$$\rho_{IP}(C) = 0.9677, \varepsilon_{IP}(C) = 0.8753,$$
$$CLSIM(C) = 0.8 * 0.967683 + 0.2 * 0.875245 = 0.9492.$$

Therefore, based on $CLSIM$, we get that $x_B > x_A > x_C$.

We now describe the method to solve Model ($M1$) using a genetic algorithm. There are several complexities involved in the proposed approach that render impractical traditional solvers as a solution method. For instance, each iteration requires calculating the Grey closeness and similitude incidence degree between the portfolio and the ideal sequence. The presence of cardinality constraint further exacerbates the complexity of the problem (since the cardinality constraint is an integer valued function). For a problem of higher dimensions (i.e., a large asset universe), obtaining a *globally optimal solution* becomes a challenge. Therefore, we must *forgo* the unrealistic expectation of a globally optimal solution and look for alternate solution methodologies to provide an *acceptable* solution in a *reasonable* run time.

Table 3 Calculation of the portfolio return sequences

A_1	A_2	$P_A = 0.5 * A_1 + 0.5 * A_2$	$P_B = 0.3 * A_1 + 0.7 * A_2$	$P_C = 0.8 * A_1 + 0.2 * A_2$
0.022816	0.014008	0.018412	0.016651	0.021055
-0.00972	-0.00783	-0.00878	-0.0084	-0.00934
0.007968	0.005726	0.006847	0.006399	0.00752
-0.0047	-0.00096	-0.00283	-0.00208	-0.00395
0.016086	0.013438	0.014762	0.014232	0.015557
0.021241	0.007636	0.014438	0.011718	0.01852
0.002261	-0.00188	0.00019	-0.00064	0.001433
0.021364	0.017127	0.019246	0.018398	0.020517
-0.0135	-0.00289	-0.0082	-0.00608	-0.01138
-0.00429	-0.0049	-0.00459	-0.00472	-0.00441

Genetic algorithms [55] are a class of metaheuristics for solving optimization problems when traditional methods are impractical. Unlike conventional methods, they simultaneously optimize a population of solutions. Genetic algorithms employ selection, crossover, and mutation operators to *explore* and *exploit* the search space for optimal or near-optimal solutions. Genetic algorithms are stochastic, i.e., the rules of transition in the algorithm are dependent upon probabilistic conditions. Observant readers would have guessed by now that the natural process of evolution has inspired the operators and the transition rules of genetic algorithms. The "Survival of the fittest" principle is implemented in genetic algorithms to improve the solutions and weed out *unfit* solutions iteratively. The fitness of a solution is characterized by its objective function value. In the natural process of evolution, both parents' fitness characteristics are passed to their offspring, creating even fitter individuals in the process. The selection operator creates a *mating pool* of solutions, i.e., the set of solutions (parents) that produce the subsequent generation's solutions (i.e., the solutions for the next iteration). The offspring solutions are created by applying the crossover operator to the parent solutions. The mutation operator involves a random alteration of solutions. The mutation itself is a random walk in the solution space. When occasionally used with crossover and selection, it provides an alternative path to the algorithm if it gets stuck in a locally optimal solution (introducing new solutions in the population). As explained above, the transition rules are probabilistic; thus, a probability condition is associated with each algorithm operator. As with any machine-learning algorithm, we can improve the genetic algorithm's performance by tuning its parameters (in this case, the crossover and mutation probabilities). Selection and crossover direct the search towards optimal solutions, i.e., these operators exploit the feasible space. However, relying heavily on these operators can lead to a premature convergence (i.e., local optimal solutions). As explained above, mutation aids in

exploring the search space, i.e., it provides an *alternate route* to the algorithm. The trade-off is that the mutation operator slows down convergence. Therefore, one needs to strike a delicate balance between exploitation and exploration operators. This is done by controlling the crossover and mutation probabilities. The size of the population also plays an important role. A large population ensures a sufficient diversity of solutions in the population; however, it increases the algorithm's runtime. A small population quickens the algorithm's convergence, but it also risks a locally optimal solution. A trade-off of convergence versus runtime is also involved with the maximum number of generations (iterations). Now, we discuss the operators of a genetic algorithm in detail.

Remark 6 Typically, the solutions are encoded as bit strings in the genetic algorithm. The selection, crossover, and the mutation operator work with an encoding of solutions. It works well for discrete search spaces with countable solutions. In the case of continuous search spaces with uncountable solutions, a lot of bits are required to maintain the requisite precision in the solutions. It hampers the algorithm's performance. We avoid encoding the solutions and use them directly to overcome this disadvantage (i.e., a real-coded genetic algorithm). The operators described in the sequel apply to a real-coded genetic algorithm only.

– **Selection**
  As explained above, the selection operator is used to create a mating pool of solutions (the set of solutions (parents) used to develop the next generation's solutions). To emulate natural selection, we introduce a bias towards fitter parent solutions. We employ the roulette wheel selection. Each parent receives a slot on the roulette wheel relative to its fitness. Thus, the selection becomes biased towards parents with a higher fitness value. In this manner, the parent solutions are selected

to produce the subsequent generation's solutions (i.e., offspring). The procedure is implemented in the following steps

Step 1: Make the fitness value of each solution in the population positive. To achieve this, determine the population's minimum fitness value. Deduct this value from every solution's fitness value.

Step 2: Evaluate the population's average fitness, $f\_av = \frac{\sum_i fitness_i}{POP}$. Here $fitness_i$ is the objective value (fitness value) of the ith solution and POP is the size of the population.

Step 3: Calculate the expected copies, $ex\_cpy_i = \frac{fitness_i}{f\_av}$, of each solution.

Step 4: Round ex_cpy to the nearest integer. It is the number of copies of the ith solution in the mating pool.

– **Crossover**
Various literature implementations of the crossover operators are available, viz. flat crossover, BLX-α crossover, Fuzzy connectives-based crossover (FCB), simple crossover. The crossover operator is applied to the selected solutions in the mating pool to produce the subsequent generation's solutions. In this paper, the simulated binary crossover (SBX) [56, 57] is used. It works well for multi-modal problems. We first determine a spread factor as follows

$$\zeta = \left| \frac{ch_1 - ch_2}{pr_1 - pr_2} \right|,$$

where pr_1, pr_2 are the solutions of the previous generation (parents), and ch_1, ch_2 are the offspring solutions of the current generation. Then, the probability of producing offspring solutions with a particular ζ is evaluated. It obtained as follows

$$P(\zeta) = \begin{cases} 0.5(r+1)\zeta^m, & \text{if } \zeta \le 1, \\ 0.5(r+1)\frac{1}{\zeta^{m+2}}, & \text{otherwise}, \end{cases}$$

where $m \ge 0$, $m \in \mathbb{R}$. If the magnitude of m is small, the obtained offspring solutions are far away from the parents. If the magnitude of m is large, it leads to the offspring solutions that are closer to the parent solutions. The implementation of creating offspring solutions in a particular range, say, $ch_L \le ch \le ch_U$, is as follows:

Step 1: Obtain the following:

$$\zeta_L = \frac{pr_1 + pr_2 - 2ch_L}{|pr_2 - pr_1|}, \quad \zeta_U = \frac{2ch_U - pr_1 - pr_2}{|pr_2 - pr_1|}.$$

Step 2: Evaluate the following probabilities,

$$P_1' = \int_0^{\zeta_L} P(\zeta)d\zeta, \quad P_2' = \int_0^{\zeta_U} P(\zeta)d\zeta.$$

Step 3: Obtain $u \in [0, 1]$, where u is a random number.

Step 4: Obtain ζ_1' and ζ_2' for which,

$$\frac{\int_0^{\zeta_1'} P(\zeta)d\zeta}{P_1'} = u, \quad \frac{\int_0^{\zeta_2'} P(\zeta)d\zeta}{P_2'} = u.$$

Step 5: Finally, the offspring solutions are obtained as follows,

$$ch_1 = 0.5\left[(pr_1 + pr_2) - \zeta_1' |pr_2 - pr_1|\right],$$
$$ch_2 = 0.5\left[(pr_1 + pr_2) + \zeta_2' |pr_2 - pr_1|\right].$$

– **Mutation**
There are various mutation operator implementations, viz. non-uniform mutation, Mühlenbein's mutation, real number creep, etc. If only the crossover operator is present in the algorithm, the solutions will become identical after a specific number of iterations (generations). Thus, the algorithm may get stuck in a locally optimal solution. Mutation provides an alternate path to the algorithm by randomly introducing new solutions in the population. The paper uses the non-uniform mutation operator (Michalewicz [58]). Let $y \in [lb, ub]$ be a solution, CG be the current generation, and mut_y be the mutated solution. Then

$$mut\_y = \begin{cases} y + g(CG, ub - y), & \text{if } RAND = 0, \\ y - g(CG, y - lb), & \text{if } RAND = 1, \end{cases}$$

where $RAND$ is chosen at random from the set $\{0, 1\}$, and

$$g(z, r) = r\left(1 - b^{\left(1 - \frac{z}{MG}\right)^F}\right),$$

where $b \in \mathbb{R}$, $0 \le b \le 1$, is chosen at random, MG is the maximum number of generations, and F is a user-specified parameter, which sets the level of dependency on the number of generations. The function is defined such mut_y is distant from y in the initial generations to encourage exploration, while in the later generations, mut_y comes closer to y for convergence.

– **Elitism**
Elitism operator stores the solutions with the best objective value so that they are preserved during the search

– **Constraint handling**
To dissuade the search from exploring an infeasible region, the objective function is penalized for any constraint violations. The penalty is usually levied in proportion to the extent of the constraint violation. A static penalty function, which remains constant during the entire search process, is used in this paper.

Algorithm 1 Proposed method's pseudo code, explained in Section 4.2

Input : Number of assets (N), Fuzzy return data of all the assets, Fuzzy ideal return (I_a, I_b, I_c), Maximum number of generations (MG), Historical asset return data (HIS), Length of historical data (L), Probability of Crossover and Mutation (p_m and p_c), Population size (POP).

Output: Optimal allocation of assets, x^*.

```
1  for i = 0 to POP − 1 do
2      sum = 0
3      for j = 0 to N − 1 do
4          x[i][j] =Rand(0,1)        ▷ Generate a random
                number between 0 and 1 for each asset
                allocation
5          sum = sum+x[i][j]
6      end
7      for j = 0 to N − 1 do
8          x[i][j] = x[i][j]/sum ▷ For each i, normalize
                x
9      end
10 end
11 GenIdealSequence((Iₐ, I_b, I_c))     ▷ Generate the ideal
       sequence based on the fuzzy ideal return (Section 3.3)
12 for gen = 0 to MG − 1 do
13     for i = 0 to POP − 1 do
14         for t = 0 to L − 1 do
15             port_seq[i][t] = 0
16             for m = 0 to N − 1 do
17                 port_seq[i][t] =
                       port_seq[i][t] + x[i][m] ∗ HIS[m][t]
                       ▷ Get the return sequence of the iᵗʰ
18                                            ▷ solution
19             end
20         end
21     end
22     CalCLSIM(port_seq)     ▷ Calculate CLSIM for
           each member of the population using (8) and (9)
23     GetReturn(x)    ▷ Calculate the expected return for
           each member of the population
24     Penalize(x)       ▷ Levy a penalty on the infeasible
           solutions
25     Elitism(x)        ▷ Store 10% of the best solutions
26     RouletteWheelSelection(x)     ▷ Selection operator
27     SBX(x)                   ▷ SBX crossover operator
28     Mutation(x)
29     Elitism(x)       ▷ Restore the best solutions of the
           previous generation
30 end
```

Before closing this section, let us discuss the computational complexity of the proposed genetic algorithm. The computations can be broadly classified into three categories,

1. Evaluation of the objective function along with the penalties.
2. Applying the roulette wheel selection, SBX crossover, and the non-uniform mutation operator.
3. Sorting the solutions according to objective function values for the elitism operator.

The above computations are performed in each generation. Thus, if N is the number of assets (i.e., problem dimension), POP is the population size, MG is the maximum number of generations, and L is the length of the historical data, then the worst case time complexity of each step is given by

1. $O(MG \times N \times POP)$ (for evaluating the return)+$O(MG \times N \times POP \times L)$ (for evaluating the risk) $= O(MG \times N \times POP \times L)$.
2. $O(MG \times N \times POP)+O(MG \times N \times POP)+O(MG \times N \times POP) = O(MG \times N \times POP)$ (for selection, mutation, and crossover).
3. $O(MG \times POPlog(POP))$.

Thus, the worst case time complexity of the algorithm is given by $O(MG \times N \times POP \times L) + O(MG \times POPlog(POP)) < O(MG \times N \times POP \times L) + O(MG \times POP^2) = O(MG \times POP(N \times L + POP)) = O(MG \times POP \times X)$, where $X = N \times L + POP$. Therefore, the worst-case time complexity of the algorithm is bounded by a polynomial.

5 Numerical experiments

We illustrate the proposed method on a case study of 100-asset of the U.S. stock market's NASDAQ-100 index. The portfolio sequence is constructed using each asset's daily return data from 1st January 2016 to 31st December 2019 (1004 observations). The asset returns have been modelled as coherent fuzzy numbers. See Appendix C for a method on fuzzy number construction. We demonstrate both methods discussed in Section 4.1. Standard fuzzy numbers can be handled by using the coherent fuzzy numbers with $l = 1$. The fuzzy return data for each asset is shown in Table 4.

5.1 Solving the model

5.1.1 Using standard fuzzy numbers for the ideal return

As discussed in Section 4.1, the optimistic, pessimistic, and neutral investors provide different fuzzy numbers, conveying their respective perceptions of the ideal return.

Table 4 Fuzzy return data of the assets (all figures in percentages)

S. No.	Symbol	Fuzzy return	S. No.	Symbol	Fuzzy return	S. No.	Symbol	Fuzzy return
1	"AAPL"	(−9.9607,0.1412,7.0422)	35	"CVX"	(−5.2964,0.0512,6.3373)	68	"MU"	(−11.0905,0.039,13.3415)
2	"ADBE"	(−8.164,0.1289,12.2418)	36	"DLTR"	(−15.3913,0.032,12.7744)	69	"MXIM"	(−8.1473,0.1295,12.303)
3	"ADI"	(−7.0447,0.0998,6.3986)	37	"DXCM"	(−32.6516,0.087,30.6475)	70	"NFLX"	(−13.1262,−0.0143,19.0281)
4	"ADP"	(−6.2494,0.1272,9.1055)	38	"EA"	(−13.3052,0.0538,16.0523)	71	"NTES"	(−14.803,0.0696,14.0844)
5	"ADSK"	(−15.8599,0.1185,15.3254)	39	"EBAY"	(−12.4527,0.0857,13.8245)	72	"NVDA"	(−18.7559,0.1267,29.8067)
6	"AEP"	(−3.5357,0.0832,2.297)	40	"EXC"	(−5.8622,0.0616,4.5511)	73	"NXPI"	(−10.0857,−0.0032,16.8774)
7	"ALGN"	(−26.9879,0.183,16.2381)	41	"EXPE"	(−27.3862,0.0676,9.5456)	74	"ORLY"	(−18.8921,0.0823,13.0584)
8	"ALXN"	(−12.8644,−0.0246,14.504)	42	"FAST"	(−8.0453,0.0723,17.1502)	75	"PAYX"	(−4.4491,0.1044,5.985)
9	"AMAT"	(−8.2469,0.0851,13.8122)	43	"FB"	(−18.9609,0.1105,15.5214)	76	"PCAR"	(−7.7728,0.0652,5.3829)
10	"AMD"	(−24.2291,0.0231,52.2901)	44	"FISV"	(−5.4724,0.0939,6.1644)	77	"PEP"	(−3.5233,0.0587,4.7606)
11	"AMGN"	(−9.5846,0.0666,5.6086)	45	"GILD"	(−9.0619,−0.0033,7.2485)	78	"PYPL"	(−8.1106,0.1284,10.1272)
12	"AMZN"	(−7.7148,0.1363,13.2164)	46	"GOOG"	(−7.6966,0.1455,10.4485)	79	"QCOM"	(−12.7226,0.0458,23.2074)
13	"ANSS"	(−6.5787,0.154,9.646)	47	"GOOGL"	(−7.5019,0.1445,9.6202)	80	"REGN"	(−7.6965,−0.0453,13.7722)
14	"ASML"	(−8.5516,0.0915,6.4674)	48	"IDXX"	(−7.2251,0.1812,13.2389)	81	"ROST"	(−9.376,0.0825,10.6694)
15	"ATVI"	(−12.3905,0.0902,18.8774)	49	"ILMN"	(−24.8093,0.0686,16.6031)	82	"SBUX"	(−9.1578,0.0585,9.7049)
16	"AVGO"	(−13.7447,0.0812,8.3682)	50	"INCY"	(−22.9325,−0.0061,12.9883)	83	"SGEN"	(−15.368,0.0712,18.9775)
17	"BAC"	(−7.4074,0.0609,7.1176)	51	"INTC"	(−9.102,0.0714,10.5519)	84	"SIRI"	(−10.3152,0.0598,8.1633)
18	"BIDU"	(−16.5192,−0.0322,13.5153)	52	"INTU"	(−7.3507,0.1201,8.4621)	85	"SNPS"	(−6.0185,0.123,6.4997)
19	"BIIB"	(−29.2305,0.0566,26.1107)	53	"ISRG"	(−7.0068,0.1578,8.1709)	86	"SPLK"	(−23.0622,0.115,17.8932)
20	"BKNG"	(−13.5197,0.1161,11.2436)	54	"JD"	(−10.6354,0.0304,12.8866)	87	"SWKS"	(−10.6533,0.0867,13.013)
21	"BMRN"	(−9.6552,−0.0269,10.0683)	55	"KDP"	(−82.0556,0.044,22.3941)	88	"TCOM"	(−19.0186,0.0194,19.7771)
22	"CDNS"	(−7.1565,0.1571,15.3657)	56	"KHC"	(−27.4595,0.0081,13.4386)	89	"TEAM"	(−15.875,0.1681,24.646)
23	"CDW"	(−7.5726,0.1557,12.148)	57	"KLAC"	(−10.2571,0.1518,10.461)	90	"TMUS"	(−7.8553,0.0447,9.4973)
24	"CERN"	(−10.8633,0.0844,10.2995)	58	"LRCX"	(−8.6797,0.1343,15.6965)	91	"TSLA"	(−13.7576,0.0503,17.5082)
25	"CHKP"	(−12.4883,0.0898,7.5103)	59	"LULU"	(−23.4389,0.0584,16.316)	92	"TTWO"	(−13.7571,0.0873,12.2182)
26	"CHTR"	(−11.6787,0.0335,14.1906)	60	"MAR"	(−7.3096,0.0985,6.3609)	93	"TXN"	(−8.3595,0.1506,7.645)
27	"CMCSA"	(−7.3775,0.0528,5.495)	61	"MCHP"	(−10.8789,0.1389,7.1433)	94	"VRSK"	(−6.439,0.109,5.5156)
28	"COST"	(−8.5868,0.1092,5.0879)	62	"MDLZ"	(−6.4839,0.0259,5.5492)	95	"VRSN"	(−7.1385,0.1206,17.2017)
29	"CPRT"	(−13.3863,0.1236,11.7439)	63	"MELI"	(−10.4441,0.0623,20.4579)	96	"VRTX"	(−7.2256,−0.0034,20.6418)
30	"CSCO"	(−8.6149,0.1097,9.6402)	64	"MNST"	(−14.439,0.0802,12.7859)	97	"WBA"	(−12.8052,0.0537,5.3642)
31	"CSX"	(−10.2703,0.0557,23.4002)	65	"MRVL"	(−6.4368,0.0242,14.1117)	98	"WDAY"	(−16.3479,0.1007,18.6766)
32	"CTAS"	(−6.5203,0.1126,9.4344)	66	"MSFT"	(−7.171,0.1172,7.5705)	99	"XEL"	(−4.5956,0.0951,2.3824)
33	"CTSH"	(−13.2545,0.0854,6.6192)	67	"MTCH"	(−17.7686,0.0835,14.3129)	100	"XLNX"	(−17.077,0.086,18.4366)
34	"CTXS"	(−6.5005,0.0699,7.0857)						

The ideal sequence is generated based on these fuzzy numbers (Table 5). The investor preferences for the ideal return, bounds, cardinality constraint, and the remaining model parameters have been summarized in Tables 6 and 7. The fuzzy numbers and the ideal return in Tables 4 and 6 are in terms of percentages. The optimistic investor considers higher returns ideal and does not bother about the extra risks (increased variance of the subsequent ideal return sequence). On the contrary, the primary objective of the pessimistic investor is risk minimization. Thus, the pessimistic investor settles for a low positive return. We have kept the remaining model parameters identical for each investor. Since we have chosen w_1, w_2 so that $w_1 + w_2 = 1$, by definition of $CLSIM$, we have $0 \leq CLSIM \leq 1$. Thus, the investors want at least a 75% resemblance to the ideal return (Table 7). It is reasonable to place relatively more importance on the closeness degree (i.e., $w_1 > w_2$) because the spatial closeness of the portfolio return to the ideal return is more important to the investor than the shape similarity. We generate the ideal sequences based on the fuzzy numbers and solve different portfolio optimization problems (Section 4.2). Example 1 illustrates the procedure to calculate the ideal and the portfolio return sequences. We solve an ϵ-constraint formulation of Model ($M1$) with the expected return from the portfolio as the objective and treat $CLSIM$ as a constraint in the single-objective problem. We chose the genetic algorithm parameters shown in Table 5 after extensive experimentation (parameter tuning). As discussed in Section 4.2, the metaheuristic algorithms like genetic algorithms cannot find globally optimum solutions. They optimize a population of solutions simultaneously and use operators to arrive at good solutions in a reasonable run time. The performance of a genetic algorithm is influenced by the parameter setting of the algorithm (probabilities of crossover and mutation, population size, etc.). Theoretically, a universal combination of parameters that works for each problem does not exist, just as there is no single genetic algorithm for every problem (although some general guidelines are available for a good parameter setting). After testing several combinations of parameters for stability and consistency of solutions, it is usually obtained by trial and error. It is beyond the scope of this text to present the results of all the experiments. The optimal allocations for each type of investor have been shown in Table 8. Since the threshold value for $CLSIM$ is the same for each investor, optimistic investors expect the highest return. However, it also entails larger risks as the portfolio sequence of the

Table 5 Ideal return for different types of investors

	Optimistic	Pessimistic	Neutral
Ideal return	**(0.1,0.2,0.5)**	**(0,0.05,0.1)**	**(0.05,0.15,0.3)**

Table 6 Parameters for the ϵ-constraint formulation of Model $M1$ (Section 5.1.1)

Model parameter	Value	Model parameter	Value
$CLSIM$ ($\geq$)	0.75	l_i, $(i = 1, 2, \ldots, 100)$	0.01
u_i, $(i = 1, 2, \ldots, 100)$	1	K ($\geq$)	30
w_1	0.8	w_2	0.2

optimistic investor is similar to the portfolio sequence of a higher variance (i.e., the sequence generated by the fuzzy number (0.1,0.2,0.5)). How this strategy pans out in the real world is another matter entirely. We will take up this discussion during the out-of-sample analysis.

Remark 7 As explained in Section 1, the ideal return is the target return that the investor aims to achieve with the portfolio (expressed as a fuzzy number). Careful consideration must be given in selecting the ideal return. The investors can consult experts or set the ideal return based on the assets' historical performance. Setting an unrealistically high ideal return will only result in infeasible solutions. The bound constraints have been added to avoid unrealistic asset allocations in the final solution (10^{-3} or lesser). The cardinality constraint ensures the investors' desired level of diversification. A pessimistic or conservative investor may want investments in many assets so that profits could offset losses in other assets. An optimistic investor focused on return maximization may not bother about too much diversification. Here, the experience of the investor or expert plays an important role. The threshold value of $CLSIM$ indicates the level of similarity between the ideal return and the investors' desired portfolio return. A pessimistic or conservative investor may want a very high level of similarity, but this may result in infeasible solutions. In such a scenario, the investor can adjust the preferences (see Fig. 5) until a suitable solution is found.

5.1.2 Using coherent fuzzy numbers for the ideal return

Suppose the evidence from the historical asset data is such that using different fuzzy numbers for the ideal return is not practical. In that case, the same fuzzy number can be

Table 7 Parameter setting of the genetic algorithm

Parameter	Value	Parameter	Value
Probability of crossover, p_c	0.8	Number of generations (MG)	2000
Probability of mutation, p_m	0.2	Population size (POP)	200
F (Non-Uniform Mutation)	5	m (SBX Crossover)	2

Table 8 Optimal asset allocations (Section 5.1.1)

Symbol	Pessimistic Allocation (Expectedreturn = 1.9085%)	Neutral Allocation (Expectedretun = 3.7179%)	Optimistic Allocation (Expectedreturn = 4.8967%)
"AAPL"	0	0	0.0108
"ADBE"	0	0	0
"ADI"	0	0	0.0107
"ADP"	0.0105	0	0.0118
"ADSK"	0	0	0
"AEP"	0	0	0
"ALGN"	0	0	0.0103
"ALXN"	0.0443	0.0102	0
"AMAT"	0.0108	0	0
"AMD"	0.0207	0.3288	0.6658
"AMGN"	0	0	0.0133
"AMZN"	0	0.0104	0
"ANSS"	0	0	0.0107
"ASML"	0	0	0
"ATVI"	0.0105	0	0
"AVGO"	0	0	0.0107
"BAC"	0	0.0102	0.0107
"BIDU"	0.0105	0	0
"BIIB"	0	0	0
"BKNG"	0	0	0
"BMRN"	0.0106	0.0104	0.0128
"CDNS"	0	0.0102	0.011
"CDW"	0.0103	0.0101	0
"CERN"	0	0	0
"CHKP"	0	0	0.0103
"CHTR"	0.0102	0	0
"CMCSA"	0	0.012	0
"COST"	0	0	0
"INTC"	0	0.0102	0.0108
"INTU"	0	0	0
"ISRG"	0	0	0
"JD"	0	0.0102	0
"KDP"	0	0	0
"KHC"	0.0105	0	0
"KLAC"	0	0	0
"LRCX"	0.0105	0.01	0.0111
"LULU"	0	0	0
"MAR"	0	0.0108	0.0111
"MCHP"	0	0	0
"MDLZ"	0	0	0
"MELI"	0	0.01	0.0111
"MNST"	0	0	0
"MRVL"	0.0152	0.0107	0
"MSFT"	0	0	0
"MTCH"	0	0	0
"MU"	0	0	0
"MXIM"	0.0102	0	0.015
"NFLX"	0	0	0
"NTES"	0	0	0
"NVDA"	0	0	0.0113
"NXPI"	0.0101	0.0185	0.0118
"ORLY"	0	0	0
"PAYX"	0.0105	0	0
"PCAR"	0	0	0.0103
"PEP"	0.0106	0.0103	0
"PYPL"	0	0	0.0105

Table 8 (continued)

3822 M. K. Mehlawat et al.

Pessimistic				Optimistic				Neutral			
Expectedreturn = 1.9085%				Expectedreturn = 4.8967%				Expectedreturn = 3.7179%			
Symbol	Allocation	Symbol	Allocation	Symbol	Allocation	Symbol	Allocation	Symbol	Allocation	Symbol	Allocation
"CPRT"	0	"QCOM"	0.0205	"CPRT"	0	"QCOM"	0.0134	"CPRT"	0	"QCOM"	0.0112
"CSCO"	0	"REGN"	0.2908	"CSCO"	0.0105	"REGN"	0	"CSCO"	0	"REGN"	0
"CSX"	0.1183	"ROST"	0	"CSX"	0	"ROST"	0	"CSX"	0	"ROST"	0
"CTAS"	0	"SBUX"	0	"CTAS"	0	"SBUX"	0	"CTAS"	0	"SBUX"	0
"CTSH"	0.0143	"SGEN"	0.0105	"CTSH"	0	"SGEN"	0	"CTSH"	0	"SGEN"	0.0103
"CTXS"	0	"SIRI"	0	"CTXS"	0	"SIRI"	0	"CTXS"	0.0101	"SIRI"	0
"CVX"	0.0104	"SNPS"	0	"CVX"	0	"SNPS"	0	"CVX"	0	"SNPS"	0
"DXCM"	0	"SWKS"	0	"DXCM"	0	"SWKS"	0	"DXCM"	0	"SWKS"	0
"DLTR"	0	"SPLK"	0	"DLTR"	0	"SPLK"	0	"DLTR"	0.0105	"SPLK"	0
"EA"	0	"TCOM"	0.0106	"EA"	0	"TCOM"	0	"EA"	0	"TCOM"	0.0103
"EBAY"	0	"TEAM"	0	"EBAY"	0	"TEAM"	0	"EBAY"	0	"TEAM"	0
"EXC"	0	"TMUS"	0.0105	"EXC"	0	"TMUS"	0	"EXC"	0	"TMUS"	0
"EXPE"	0	"TSLA"	0	"EXPE"	0	"TSLA"	0.0108	"EXPE"	0	"TSLA"	0
"FAST"	0.0665	"TTWO"	0	"FAST"	0	"TTWO"	0	"FAST"	0	"TTWO"	0
"FB"	0	"TXN"	0	"FB"	0	"TXN"	0	"FB"	0	"TXN"	0.0101
"FISV"	0	"VRSK"	0	"FISV"	0.0105	"VRSK"	0	"FISV"	0	"VRSK"	0.0107
"GILD"	0	"VRSN"	0.0642	"GILD"	0	"VRSN"	0.0152	"GILD"	0.0104	"VRSN"	0.0119
"GOOG"	0	"VRTX"	0.1314	"GOOG"	0.0111	"VRTX"	0.016	"GOOG"	0	"VRTX"	0.3675
"GOOGL"	0.0149	"WBA"	0.0104	"GOOGL"	0	"WBA"	0	"GOOGL"	0.0102	"WBA"	0.0131
"IDXX"	0.0107	"WDAY"	0	"IDXX"	0	"WDAY"	0	"IDXX"	0	"WDAY"	0
"ILMN"	0	"XEL"	0	"ILMN"	0	"XEL"	0	"ILMN"	0	"XEL"	0
"INCY"	0	"XLNX"	0	"INCY"	0.0103	"XLNX"	0.0105	"INCY"	0.0103	"XLNX"	0.0103

used as the ideal return for each investor type. We introduce the investor attitude in the model using the adaptive index, l, of the coherent fuzzy numbers (Section 3.1). For a fair comparison, the remaining model parameters (Table 9) have been kept identical as in Section 5.1.1. The genetic algorithm parameters are also the same as in Table 5. The optimal asset allocations are shown in Table 10. The expected return is highest for the optimistic investor in this case as well. The difference in the expected return from Section 5.1.1 is due to the modelling of investor attitude. We have chosen $l = 0.5$ for an optimistic investor; any value of $l < 1$ would suffice for an optimistic investor. The same is the case for a pessimistic investor. Any value of $l > 1$ can be used for a pessimistic investor. The magnitude of l only indicates the extent of pessimism or optimism (although, for an optimistic investor, the value l works in an opposite sense, i.e., as $l \to 0$, the optimism increases). We have demonstrated this concept visually in Fig. 2 for a better understanding.

Remark 8 Observing Tables 7 and 9, it may seem unusual that we have chosen identical model parameters for each type of investor (optimistic, pessimistic, and neutral). This has been done deliberately to demonstrate the accuracy of the proposed method of modelling investor attitude using $CLSIM$ and the fuzzy ideal return.

5.2 Out-of-sample analysis

This section is dedicated to the optimal portfolios' out-of-sample performance (Sections 5.1.1 and 5.1.2). We assume that \$1 is allocated to the obtained portfolios on January 1st, 2020, and track the accumulated wealth till December 31st, 2020. No additional investment is made during this period. Figures 6 and 7 illustrate the accumulated wealth of the obtained portfolios of Sections 5.1.1 and 5.1.2. We have also compared the out-of-sample performance with Markowitz's [1] mean-variance (M-V) model and Speranza's [2] mean-MASD model. For a fair comparison, we have imposed the cardinality and the bound constraints in both models. It can be observed that the COVID-19

pandemic significantly affected the portfolios' performance during the period March-April 2020. Post that, the portfolios make a significant recovery even when no changes were made in the asset allocations. It can be observed from both Figs. 6 and 7 that the pessimistic portfolio can minimize extreme losses during crisis periods. This behaviour is compatible with the pessimistic investor, whose primary objective is risk minimization. In contrast, the optimistic portfolio suffers from severe losses and lags behind the other portfolios during this period. The M-V model and the Mean-MASD model also exhibit poor performance during this period. During the period of the recovery of the stock market (July 2020 onwards), the optimistic portfolio shows a dramatic surge and overtakes all the other portfolios. This illustrates the importance of recognizing the bull and bear stock market trends. Adopting an aggressive trading strategy during periods of stock market crisis can lead to significant losses. When the stock market is on an upward trend, the pessimistic portfolio lags behind the other portfolios because of the pessimistic investors' conservative attitude. Adopting a conservative attitude when the stock market is in a bullish trend leads to lost opportunities. This also illustrates the criticality of an accurate stock market assessment. From Figs. 6 and 7, we can observe that the pessimistic portfolio performs worse than even the M-V and Mean-MASD portfolios, while the neutral portfolio exhibits near-similar performance. This demonstrates that both of the proposed methods to model the investor attitude are effective and accurate.

The sample statistics of the optimal portfolios' out-of-sample performance (Sections 5.1.1 and 5.1.2) are shown in Tables 11 and 12. Apart from the usual statistics, we have also shown two performance indicators: the Sharpe ratio and the Sortino ratio. Sharpe ratio [59], also known as a reward to variability is defined as

$$\text{Sharpe ratio} = \frac{r_p - r_F}{\sigma_p},$$

where r_p is the portfolio return, r_F is the risk-free return, and σ_p is the standard deviation of the portfolio return. Sortino ratio [60], a natural extension of the Sharpe ratio, is defined as

$$\text{Sortino ratio} = \frac{r_p - r_T}{\sigma_D},$$

where r_p is the portfolio return, r_T is the target return of the investor, and σ_D is defined as

$$\sigma_D = \sqrt{\sum_{i=1}^{T} \frac{(\min\{r_i - r_T, 0\})^2}{T}},$$

where T is the number of observations, r_i is the realized return of the ith day. This paper uses U.S. treasury bond futures as a risk-free asset. Each investor's target return

Table 9 Parameters for the ϵ-constraint formulation of Model $M1$ (Section 5.1.2)

Ideal return	(0,0.15,0.35)		
Model parameter	Value	Model parameter	Value
$CLSIM$ ($\geq$)	0.75	l_i, $(i = 1, 2, \ldots, 100)$	0.01
u_i, $(i = 1, 2, \ldots, 100)$	1	K ($\geq$)	30
w_1	0.8	w_2	0.2

Table 10 Optimal asset allocations (Section 5.1.2)

Pessimistic ($l = 1.5$)				Optimistic ($l = 0.5$)				Neutral ($l = 1$)			
Expected return=1.4607				Expected return=8.0895				Expected return=3.511			
Symbol	Allocation	Symbol	Allocation	Symbol	Allocation	Symbol	Allocation	Symbol	Allocation	Symbol	Allocation
"AAPL"	0	"INTC"	0.0109	"AAPL"	0	"INTC"	0	"AAPL"	0	"INTC"	0
"ADBE"	0	"INTU"	0.0112	"ADBE"	0	"INTU"	0	"ADBE"	0	"INTU"	0
"ADI"	0.0112	"ISRG"	0.0102	"ADI"	0	"ISRG"	0	"ADI"	0	"ISRG"	0
"ADP"	0	"JD"	0	"ADP"	0	"JD"	0	"ADP"	0	"JD"	0
"ADSK"	0	"KDP"	0	"ADSK"	0	"KDP"	0	"ADSK"	0	"KDP"	0
"AEP"	0.0106	"KHC"	0	"AEP"	0	"KHC"	0	"AEP"	0	"KHC"	0
"ALGN"	0	"KLAC"	0	"ALGN"	0	"KLAC"	0.0103	"ALGN"	0	"KLAC"	0
"ALXN"	0	"LRCX"	0	"ALXN"	0	"LRCX"	0.0109	"ALXN"	0.0141	"LRCX"	0
"AMAT"	0	"LULU"	0	"AMAT"	0.0109	"LULU"	0	"AMAT"	0.0107	"LULU"	0
"AMD"	0.2329	"MAR"	0	"AMD"	0.4757	"MAR"	0.0107	"AMD"	0.2695	"MAR"	0
"AMGN"	0.0107	"MCHP"	0	"AMGN"	0	"MCHP"	0	"AMGN"	0.0102	"MCHP"	0
"AMZN"	0.0105	"MDLZ"	0	"AMZN"	0	"MDLZ"	0	"AMZN"	0	"MDLZ"	0
"ANSS"	0.0191	"MELI"	0	"ANSS"	0.0114	"MELI"	0.0109	"ANSS"	0.0105	"MELI"	0.01
"ASML"	0	"MNST"	0	"ASML"	0	"MNST"	0.0112	"ASML"	0	"MNST"	0.0109
"ATVI"	0.0104	"MRVL"	0.0105	"ATVI"	0.0909	"MRVL"	0	"ATVI"	0	"MRVL"	0.0107
"AVGO"	0.014	"MSFT"	0	"AVGO"	0	"MSFT"	0	"AVGO"	0	"MSFT"	0
"BAC"	0	"MTCH"	0	"BAC"	0	"MTCH"	0	"BAC"	0	"MTCH"	0
"BIDU"	0	"MU"	0	"BIDU"	0.0109	"MU"	0	"BIDU"	0	"MU"	0
"BIIB"	0	"MXIM"	0.0102	"BIIB"	0.0122	"MXIM"	0	"BIIB"	0	"MXIM"	0.0112
"BKNG"	0	"NFLX"	0	"BKNG"	0	"NFLX"	0	"BKNG"	0	"NFLX"	0
"BMRN"	0	"NTES"	0.0101	"BMRN"	0	"NTES"	0	"BMRN"	0	"NTES"	0
"CDNS"	0	"NVDA"	0.0101	"CDNS"	0.0112	"NVDA"	0.0103	"CDNS"	0.0102	"NVDA"	0
"CDW"	0	"NXPI"	0	"CDW"	0.0116	"NXPI"	0	"CDW"	0.0103	"NXPI"	0
"CERN"	0	"ORLY"	0.0101	"CERN"	0.0105	"ORLY"	0	"CERN"	0	"ORLY"	0
"CHKP"	0.0101	"PAYX"	0	"CHKP"	0	"PAYX"	0	"CHKP"	0	"PAYX"	0
"CHTR"	0	"PCAR"	0	"CHTR"	0.0113	"PCAR"	0	"CHTR"	0.011	"PCAR"	0
"CMCSA"	0	"PEP"	0.0105	"CMCSA"	0	"PEP"	0.0109	"CMCSA"	0	"PEP"	0.0125
"COST"	0.0112	"PYPL"	0	"COST"	0	"PYPL	0	"COST"	0.0105	"PYPL"	0
"CPRT"	0	"QCOM"	0	"CPRT"	0.0105	"QCOM"	0	"CPRT"	0	"QCOM"	0.0102
"CSCO"	0	"REGN"	0.0103	"CSCO"	0	"REGN"	0	"CSCO"	0.0108	"REGN"	0

Table 10 (continued)

| Pessimistic (l = 1.5)				Optimistic (l = 0.5)				Neutral (l = 1)			
Expected return=1.4607				Expected return=8.0895				Expected retun=3.511			
Symbol	Allocation	Symbol	Allocation	Symbol	Allocation	Symbol	Allocation	Symbol	Allocation	Symbol	Allocation
"CSX"	0.0636	"ROST"	0.0101	"CSX"	0	"ROST"	0	"CSX"	0.4251	"ROST"	0.0103
"CTAS"	0	"SBUX"	0	"CTAS"	0	"SBUX"	0.0111	"CTAS"	0	"SBUX"	0
"CTSH"	0.0102	"SGEN"	0.01	"CTSH"	0	"SGEN"	0	"CTSH"	0	"SGEN"	0
"CTXS"	0	"SIRI"	0.0109	"CTXS"	0	"SIRI"	0	"CTXS"	0	"SIRI"	0.0115
"CVX"	0	"SNPS"	0	"CVX"	0	"SNPS"	0.0102	"CVX"	0	"SNPS"	0.0103
"DLTR"	0	"SPLK"	0.01	"DLTR"	0	"SPLK"	0	"DLTR"	0.0127	"SPLK"	0
"DXCM"	0	"SWKS"	0	"DXCM"	0	"SWKS"	0	"DXCM"	0	"SWKS"	0
"EA"	0	"TCOM"	0	"EA"	0.011	"TCOM"	0.1334	"EA"	0.0131	"TCOM"	0.0116
"EBAY"	0	"TEAM"	0	"EBAY"	0	"TEAM"	0.0112	"EBAY"	0	"TEAM"	0
"EXC"	0	"TMUS"	0	"EXC"	0	"TMUS"	0.0017	"EXC"	0	"TMUS"	0.0102
"EXPE"	0	"TSLA"	0	"EXPE"	0	"TSLA"	0.0108	"EXPE"	0	"TSLA"	0.0102
"FAST"	0.0109	"TTWO"	0	"FAST"	0.0106	"TTWO"	0.0138	"FAST"	0	"TTWO"	0.0101
"FB"	0	"TXN"	0	"FB"	0	"TXN"	0	"FB"	0	"TXN"	0
"FISV"	0	"VRSK"	0	"FISV"	0	"VRSK"	0	"FISV"	0	"VRSK"	0
"GILD"	0	"VRSN"	0.0111	"GILD"	0	"VRSN"	0	"GILD"	0	"VRSN"	0.0106
"GOOG"	0.0127	"VRTX"	0.4057	"GOOG"	0	"VRTX"	0.0117	"GOOG"	0	"VRTX"	0
"GOOGL"	0	"WBA"	0	"GOOGL"	0	"WBA"	0.0109	"GOOGL"	0	"WBA"	0
"IDXX"	0	"WDAY"	0	"IDXX"	0	"WDAY"	0.0115	"IDXX"	0.0102	"WDAY"	0.0103
"ILMN"	0	"XEL"	0	"ILMN"	0	"XEL"	0	"ILMN"	0	"XEL"	0
"INCY"	0	"XLNX"	0	"INCY"	0	"XLNX"	0	"INCY"	0	"XLNX"	0.0104
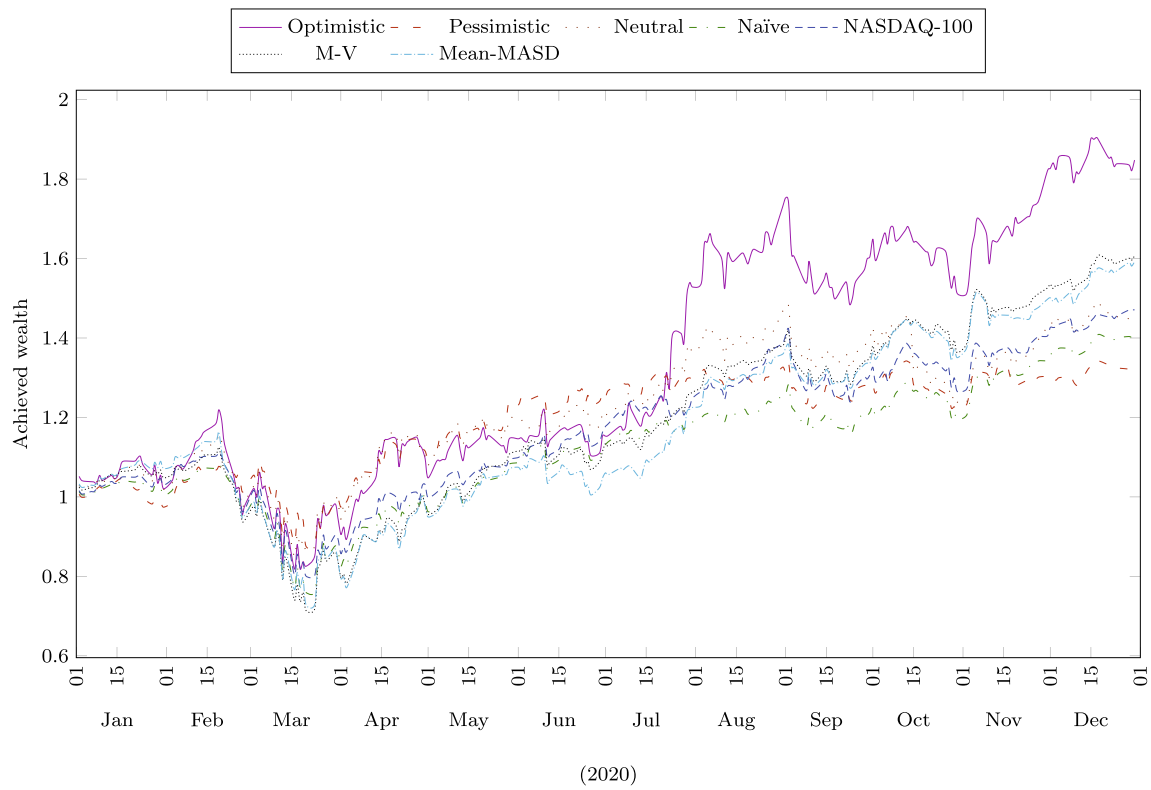
Fig. 6 Wealth accumulated by the obtained portfolios (Section 5.1.1)
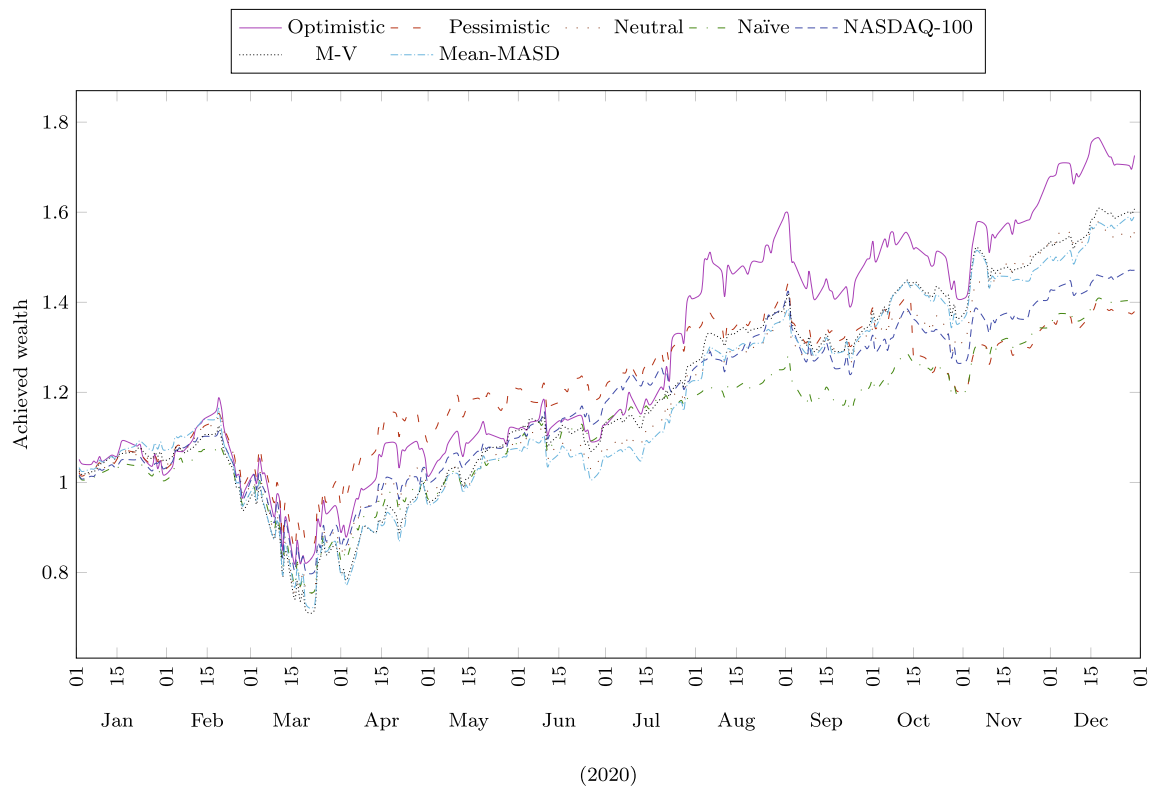

Fig. 7 Wealth accumulated by the obtained portfolios (Section 5.1.2)

Table 11 Sample statistics of the optimal portfolios' daily out-of-sample return(Section 5.1.1)

	Optimistic	Pessimistic	Neutral	Naïve	NASDAQ-100	M-V	Mean-MASD
Target return	0.25%	0.05%	0.1625%	–	–	–	–
Min return	-12.7372%	**-9.8789%**	-10.6109%	-12.2318%	-12.1932%	-12.6723%	-11.3057%
Max return	11.7088%	8.6457%	9.8605%	9.3677%	10.0722%	13.4364%	**14.176%**
Average return	**0.2949%**	0.1322%	0.1826%	0.1606%	0.1797%	0.217%	0.2165%
SD	0.0319	**0.0212**	0.0258	0.0219	0.0229	0.0238	0.0252
Sharpe ratio	**0.079**	0.0423	0.0541	0.0539	0.0598	0.0732	0.069
Skewness	**-0.0768**	-0.1427	-0.2202	-0.6102	-0.5377	-0.3964	-0.1233
Kurtosis	**3.1795**	5.1622	3.6925	7.4398	6.0137	8.1525	6.8149
Sortino ratio	0.0201	**0.056**	0.0108	–	–	–	–

Acronyms: SD: Standard deviation; M-V: Mean-variance; M-MASD: Mean-mean absolute semi-deviation

is the expected value of the fuzzy ideal return (Tables 11 and 12). The optimistic portfolio has the best overall performance during the out-of-sample period. Out of the three different kinds of investor attitudes, the pessimistic portfolio exhibits the least standard deviation during the out-of-sample period. In Table 11, the pessimistic portfolio's standard deviation is less than even the naïve portfolio, with the maximum systematic risk diversification. This is in line with the pessimistic investors' focus on risk minimization. The optimistic portfolio has the best overall reward to variability, as indicated by the Sharpe ratio. Since the Sortino ratio is dependent on the target return, the optimistic portfolio has a lesser value of the Sortino ratio (Table 11). The optimistic portfolio's target return is very high, while the target return for the pessimistic portfolio is low. All the portfolios exhibit a slight negative skewness. This could be due to the significant losses incurred during the slowdown during the March-April period (effects of COVID-19). The kurtosis for the optimistic portfolio is only slightly higher than 3. This indicates that the out-of-sample returns were relatively stable around the expected return. This is in contrast with other portfolios, which exhibit a large kurtosis.

This shows that these portfolios were more volatile during the out-of-sample period. Thus, the optimistic portfolio outperforms the rest of the portfolios in this aspect as well.

5.3 A case study of the Indian stock market

This section applies the proposed method to a case study involving 50 assets listed in the Indian stock market's NIFTY-50 index. This section uses a 4-year historical return data of assets beginning from 1st January 2016 to 31st December 2019. Table 13 provides the fuzzy data of the assets. We will use the adaptive index, l, to incorporate the investor attitude into the model. Table 14 provides the model parameters. In this study, let us focus on achieving closeness to the ideal return. Therefore, we have chosen $w_1 = 1$, $w_2 = 0$, and $CLSIM \geq 0.95$. Next, we solve the model to obtain the optimal asset allocations for each investor. Table 15 provides the optimal asset allocations. We can observe that the optimistic investor expects the highest return, while the pessimistic investor expects the lowest return. Now, let us focus on the out-of-sample performance. As in Section 5.2, we assume that the capital of INR 1

Table 12 Sample statistics of the optimal portfolios' daily out-of-sample return (Section 5.1.2)

	Optimistic	Pessimistic	Neutral	Naïve	NASDAQ-100	M-V	Mean-MASD
Target return	0.1917%	0.145%	0.1625%	–	–	–	–
Min return	-11.3529%	**-11.1793%**	-13.424%	-12.2318%	-12.1932%	-12.6723%	-11.3057%
Max return	9.0138%	10.1734%	12.364%	9.3677%	10.0722%	13.4364%	**14.176%**
Average return	**0.2533%**	0.1586%	0.2087%	0.1606%	0.1797%	0.217%	0.2165%
SD	0.027	0.0248	0.0255	**0.0219**	0.0229	0.0238	0.0252
Sharpe ratio	**0.0781**	0.0466	0.0651	0.0539	0.0598	0.0732	0.069
Skewness	-0.3626	**-0.2267**	-0.3947	-0.6102	-0.5377	-0.3964	-0.1233
Kurtosis	**3.3895**	4.7558	7.4855	7.4398	6.0137	8.1525	6.8149
Sortino ratio	**0.0319**	0.0076	0.025	–	–	–	–

Acronyms: SD: Standard deviation; M-V: Mean-Variance; M-MASD: Mean-mean absolute semi-deviation

Table 13 Fuzzy returns of the assets (Section 5.3)

S. No.	Symbol	Fuzzy return	S. No.	Symbol	Fuzzy return
1	"ADANIPORTS"	(-3.5482,-0.0275,3.5731)	26	"INFY"	(-3.222,0.0225,2.9437)
2	"ASIANPAINT"	(-2.7443,-0.0082,3.0177)	27	"ITC"	(-2.7549,-0.0349,3.049)
3	"AXISBANK"	(-2.9879,-0.0572,3.3505)	28	"JSWSTEEL"	(-3.0825,0.0422,3.2051)
4	"BAJAJ-AUTO"	(-2.867,-0.0357,2.9481)	29	"KOTAKBANK"	(-2.8024,0.0708,3.0547)
5	"BAJFINANCE"	(-3.609,0.0369,3.6028)	30	"LT"	(-2.7109,-0.0596,3.4005)
6	"BAJAJFINSV"	(-3.1301,0.0209,3.2699)	31	"M&M"	(-3.039,-0.0287,3.2358)
7	"BHARTIARTL"	(-3.1056,-0.0411,3.6799)	32	"MARUTI"	(-3.1095,-0.0292,3.1955)
8	"BPCL"	(-3.5538,0.0159,3.3019)	33	"NESTLEIND"	(-2.7641,0.0269,3.2029)
9	"BRITANNIA"	(-2.8129,0.0156,3.1093)	34	"NTPC"	(-2.8733,-0.0427,2.8928)
10	"CIPLA"	(-2.8984,-0.0763,3.0535)	35	"ONGC"	(-3.1085,0.0058,3.0776)
11	"COALINDIA"	(-2.8886,-0.024,3.0222)	36	"POWERGRID"	(-2.8165,0.0079,2.9292)
12	"DIVISLAB"	(-3.5593,-0.0304,3.5155)	37	"RELIANCE"	(-2.8185,0.051,3.1487)
13	"DRREDDY"	(-3.2336,-0.0463,3.2191)	38	"SBIN"	(-3.175,-0.0173,3.7464)
14	"EICHERMOT"	(-3.3216,-0.0597,3.4293)	39	"SBILIFE"	(-3.0904,0.0217,3.2731)
15	"GRASIM"	(-3.1579,-0.0552,3.0835)	40	"SHREECEM"	(-2.9387,0.0256,3.2646)
16	"HCLTECH"	(-2.9908,0.0575,2.8188)	41	"SUNPHARMA"	(-3.3852,-0.1003,3.198)
17	"HDFCBANK"	(-2.7229,0.0502,2.8905)	42	"TATACONSUM"	(-3.225,-0.0121,3.593)
18	"HDFCLIFE"	(-2.8478,-0.1009,3.1803)	43	"TATAMOTORS"	(-3.7208,-0.0433,3.9584)
19	"HEROMOTOCO"	(-2.952,-0.0626,3.3207)	44	"TATASTEEL"	(-3.2681,-0.038,3.5389)
20	"HINDALCO"	(-3.2967,0.0054,3.5039)	45	"TCS"	(-2.8473,0.0521,2.9523)
21	"HINDUNILVR"	(-2.4992,0.0099,2.8047)	46	"TECHM"	(-3.008,0.0118,3.2382)
22	"HDFC"	(-2.9222,0.0325,2.8611)	47	"TITAN"	(-3.189,0.0398,3.4744)
23	"ICICIBANK"	(-2.9103,-0.0431,3.6187)	48	"ULTRACEMCO"	(-2.8434,0.0191,2.9504)
24	"IOC"	(-3.2673,0.0022,3.3503)	49	"UPL"	(-3.2451,0.0002,3.5422)
25	"INDUSINDBK"	(-3.2409,0.0142,3.3102)	50	"WIPRO"	(-2.8505,0.0345,2.7636)

(Indian National Rupee) is invested on 1st January 2020. We track the out-of-sample performance for one year (till 31st December 2020). Figure 8 traces the movements of all the portfolios during the out-of-sample period. One can immediately observe the impact of COVID-19 on all the portfolios. The risk-seeking optimistic investor faces heavy losses during this period. However, the pessimistic portfolio minimizes the losses and grows steadily throughout the year. The standard deviation of returns for the pessimistic portfolio is also the lowest except for the naïve and NIFTY-50 portfolios. This is consistent with the risk-averse attitude of the pessimistic investor. At the latter end of 2020, when the stock market undergoes a recovery (post-COVID-19), we observe that the optimistic portfolio shows a dramatic surge and overtakes all the other portfolios. This shows that the proposed method consistently captures the investor attitude. The optimistic and the pessimistic portfolios significantly outperform the M-V and the Mean-MASD portfolios as well. Table 16 records the sample statistics of each portfolio during the out-of-sample period. Each portfolio exhibits a negative skewness due to the extreme losses during the crisis period of COVID-19. All the portfolios are highly leptokurtic, which is again the

consequence of a higher frequency of extreme returns. This also establishes that each portfolio exhibited a non-normal return during the out-of-sample period. The optimistic and the pessimistic portfolios have the lowest kurtosis, demonstrating that the returns were relatively more stable around the mean than the rest of the portfolios.

5.4 Managerial implication

We demonstrated the effectiveness and accuracy of the proposed approach by the case studies of Section 5. It is the reason most of the model parameters were kept identical so

Table 14 Parameters for the ϵ-constraint formulation of Model $M1$ (Section 5.3)

Ideal return	(0.05,0.15,0.25)			
Model parameter	Value	Model parameter	Value
$CLSIM$ ($\geq$)	0.95	l_i, $(i = 1, 2, \ldots, 50)$	0.02
u_i, $(i = 1, 2, \ldots, 50)$	1	K ($\geq$)	15
w_1	1	w_2	0

Table 15 Optimal asset allocations (Section 5.3)

l = 0.5				l = 1				l = 1.5			
Expected return=0.614898				Expected return=0.0612271				Expected return=-0.213402			
Symbol	allocation	Symbol	allocation	Symbol	allocation	Symbol	allocation	Symbol	allocation	Symbol	allocation
"ADANIPORTS"	0	"INFY"	0	"ADANIPORTS"	0	"INFY"	0	"ADANIPORTS"	0	"INFY"	0
"ASIANPAINT"	0	"ITC"	0	"ASIANPAINT"	0	"ITC"	0	"ASIANPAINT"	0	"ITC"	0.0202
"AXISBANK"	0.0201	"JSWSTEEL"	0.0203	"AXISBANK"	0	"JSWSTEEL"	0.0203	"AXISBANK"	0.0205	"JSWSTEEL"	0
"BAJAJ-AUTO"	0	"KOTAKBANK"	0	"BAJAJ-AUTO"	0	"KOTAKBANK"	0.0799	"BAJAJ-AUTO"	0	"KOTAKBANK"	0.0301
"BAJFINANCE"	0.5002	"LT"	0	"BAJFINANCE"	0.1495	"LT"	0.045	"BAJFINANCE"	0.0688	"LT"	0
"BAJAJFINSV"	0.0209	"M&M"	0	"BAJAJFINSV"	0.393	"M&M"	0	"BAJAJFINSV"	0.0327	"M&M"	0
"BHARTIARTL"	0	"MARUTI"	0.0203	"BHARTIARTL"	0.0309	"MARUTI"	0	"BHARTIARTL"	0.0211	"MARUTI"	0
"BPCL"	0.02	"NESTLEIND"	0.0769	"BPCL"	0	"NESTLEIND"	0	"BPCL"	0	"NESTLEIND"	0.042
"BRITANNIA"	0.0201	"NTPC"	0	"BRITANNIA"	0	"NTPC"	0	"BRITANNIA"	0	"NTPC"	0
"CIPLA"	0	"ONGC"	0	"CIPLA"	0	"ONGC"	0	"CIPLA"	0	"ONGC"	0
"COALINDIA"	0	"POWERGRID"	0	"COALINDIA"	0	"POWERGRID"	0	"COALINDIA"	0	"POWERGRID"	0
"DIVISLAB"	0	"RELIANCE"	0	"DIVISLAB"	0.0201	"RELIANCE"	0.033	"DIVISLAB"	0	"RELIANCE"	0.5891
"DRREDDY"	0	"SBIN"	0	"DRREDDY"	0	"SBIN"	0.0281	"DRREDDY"	0	"SBIN"	0
"EICHERMOT"	0	"SBILIFE"	0	"EICHERMOT"	0	"SBILIFE"	0	"EICHERMOT"	0	"SBILIFE"	0
"GRASIM"	0	"SHREECEM"	0	"GRASIM"	0	"SHREECEM"	0	"GRASIM"	0	"SHREECEM"	0.02
"HCLTECH"	0	"SUNPHARMA"	0	"HCLTECH"	0	"SUNPHARMA"	0	"HCLTECH"	0	"SUNPHARMA"	0
"HDFCBANK"	0.0216	"TATACONSUM"	0.0611	"HDFCBANK"	0	"TATACONSUM"	0.0278	"HDFCBANK"	0	"TATACONSUM"	0
"HDFCLIFE"	0	"TATAMOTORS"	0	"HDFCLIFE"	0	"TATAMOTORS"	0	"HDFCLIFE"	0	"TATAMOTORS"	0
"HEROMOTOCO"	0	"TATASTEEL"	0.0327	"HEROMOTOCO"	0	"TATASTEEL"	0.0201	"HEROMOTOCO"	0	"TATASTEEL"	0
"HINDALCO"	0.0401	"TCS"	0	"HINDALCO"	0.0225	"TCS"	0	"HINDALCO"	0	"TCS"	0
"HINDUNILVR"	0.021	"TECHM"	0	"HINDUNILVR"	0	"TECHM"	0	"HINDUNILVR"	0	"TECHM"	0.021
"HDFC"	0	"TITAN"	0.0953	"HDFC"	0.0839	"TITAN"	0	"HDFC"	0.0201	"TITAN"	0.0213
"ICICIBANK"	0.0296	"ULTRACEMCO"	0	"ICICIBANK"	0	"ULTRACEMCO"	0.02	"ICICIBANK"	0.0528	"ULTRACEMCO"	0
"IOC"	0	"UPL"	0	"IOC"	0.0261	"UPL"	0	"IOC"	0	"UPL"	0.0203
"INDUSINDBK"	0	"WIPRO"	0	"INDUSINDBK"	0	"WIPRO"	0	"INDUSINDBK"	0.02	"WIPRO"	0
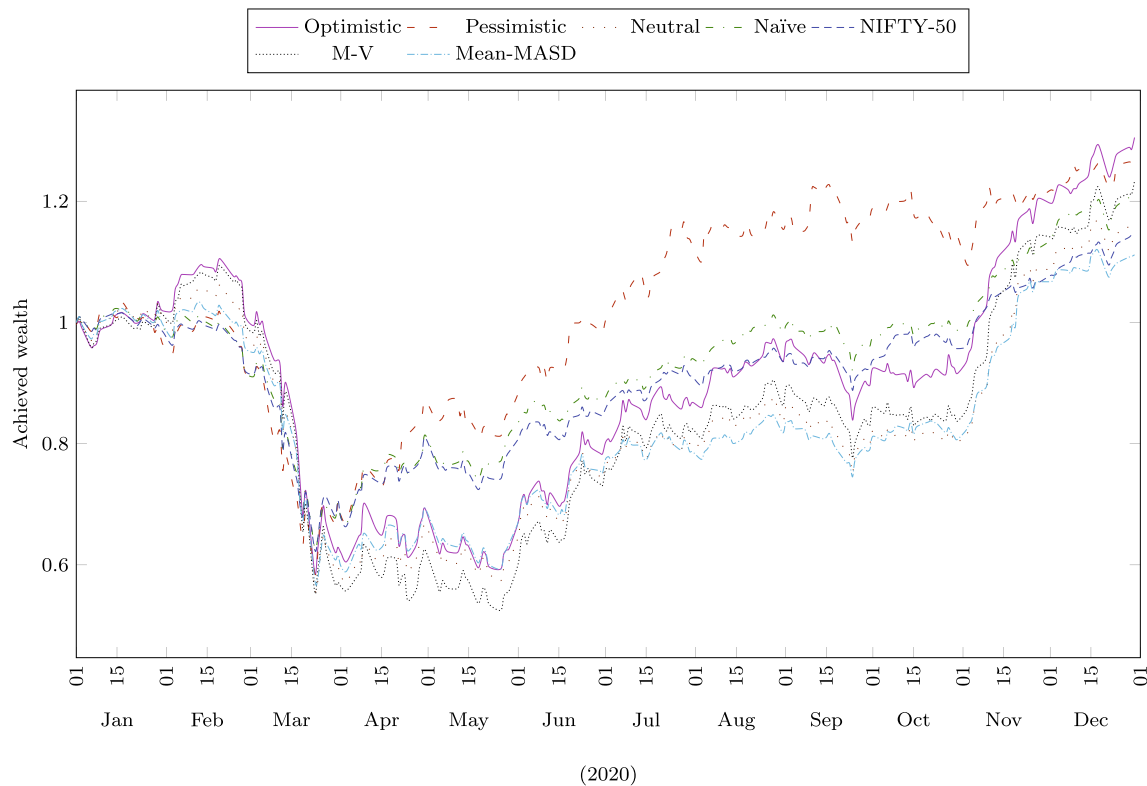
Fig. 8 Wealth accumulated by the obtained portfolios (Section 5.3)

that impact of modelling the risk and the investor attitude using $CLSIM$ and fuzzy ideal return could be studied (see Remark 8). This section discusses implementing results and conclusions derived in the previous sections in real-life scenarios. The construction of fuzzy asset returns using expert opinion or historical data is straightforward. The method of constructing fuzzy numbers for asset return is described in Appendix C. The next step is to set the ideal return. There is no specific procedure for establishing the ideal return. It depends on the assessment of the investor (whether optimistic, pessimistic, or neutral) based on the historical performance of the assets and the current situation prevailing in the stock market. Using experts' help is also a good option. We have provided only some general

guidelines on setting the investor attitude (using either of the two methodologies described in Section 4.1). The extent of optimism or pessimism depends on the subjective assessment of the investor. In practice, the investor needs to adjust the attitude in response to changing market conditions (either by changing the ideal return or the adaptive index, l). The portfolio also needs to be rebalanced accordingly as new asset information and return data become available. Incorrect assessment of the overall stock market prospects can lead to losses (as illustrated in Figs. 6 and 7). Asset returns are volatile and subject to various random and non-random factors. A universal parameter setting optimal for all problems and time frames does not exist. We cannot provide a theoretical justification for the chosen parameter

Table 16 Sample statistics of the optimal portfolios' daily out-of-sample return (Section 5.3)

	Optimistic	Neutral	Pessimistic	M-V	Mean-MASD	Naïve	NIFTY-50
Average return	**0.001479**	0.000985	0.001263	0.001343	0.000746	0.000969	0.000757
Standard deviation	0.028168	0.026963	0.024551	0.031095	0.024901	**0.019294**	0.019735
Maximum return	0.073524	0.07432	**0.111782**	0.07724	0.074986	0.088389	0.087632
Minimum return	-0.18565	-0.19694	-0.14774	-0.20593	-0.18631	-0.13043	**-0.1298**
Skewness	-1.49294	-1.74732	**-0.50264**	-1.49521	-1.86105	-1.55751	-1.40985
Kurtosis	8.1663	11.86239	9.045949	**8.11381**	13.46162	11.8822	11.04732
Sharpe's ratio	**0.048696**	0.032563	0.047074	0.039742	0.02565	0.044667	0.032903

setting. No study in portfolio optimization can claim that it has developed a model that provides optimal results in all scenarios. It all depends on the skill and experience of the investor. The investor needs to *readjust* and *recalibrate* the portfolios in the wake of changing market conditions. Our model provides a mathematical framework for an investor's subjective and linguistic assessments. Using a numerical study, we have established the model's accuracy in capturing diverse investor attitudes. The next step is to determine the optimal portfolio weights. Since our method is based on optimization techniques, a discussion of the advantages of optimizing the portfolio is warranted here. Many authors have commented on this subject. Michaud [10] showed that optimizing the portfolio weights multiplies the estimation error of the model parameters. However, Jagannathan and Ma [61] showed that imposing certain constraints (such as the bound constraints) alleviates this problem. Optimizing the portfolio provides an excellent initial recommendation to the investor on the capital allocation based on the natural assumption that assets that have performed well in the past will continue to do so in the future. Some critics also claim that risky assets become less risky if the investments are held for a long time. While this is true in general, it requires a lot of patience and discipline from the investor. Interim losses can trigger irrational or impulsive decisions. Another criticism is that the naïve portfolio performs just as well, if not better than the optimized portfolio. This claim has been proven false in our case study. Each portfolio performs significantly better than the naïve portfolio and the NASDAQ-100 index (the portfolio based on market capitalization). This was observed when there is no rebalancing of the portfolios in response to market conditions. Even the mean-variance (M-V) and the Mean-MASD portfolio outperform the naïve portfolio. Thus, we can conclude that optimization is crucial in the portfolio selection process.

6 Conclusion

Various authors have studied numerous approaches to portfolio optimization, each with its unique perspective. This paper attempts to integrate two techniques of modelling the stock market's uncertainty and the investor's imprecise preferences (namely, Fuzzy set theory and Grey Relational Analysis). We use Grey Relational Analysis to model a new measure of risk based on the investor's subjective requirements. We also incorporate the investor attitude in the model using Fuzzy set theory. We provide a method to simulate coherent triangular fuzzy numbers by modifying the triangular probability distribution. A genetic algorithm is used to solve the resulting optimization model. Numerical examples

have been provided to illustrate the proposed method. We tested the real-world performance of the model by conducting an out-of-sample analysis. The obtained results prove the validity of the proposed approach. The proposed method accurately captures the optimistic and pessimistic behaviour of the investors (See Sections 5.2 and 5.3). We can observe from Figs. 6, 7, and 8 that the pessimistic portfolio can minimize heavy losses during periods of recession (i.e., COVID-19). We can also observe that the optimistic portfolio makes above-average gains when the stock market is bullish. The proposed approach also outperforms several competing portfolios such as the naïve, the NIFTY-50, the M-V, and the Mean-MASD portfolios. The conclusion drawn is that the proposed approach effectively captures different investor behaviours observed in the stock market. The main limitation of the study is the heavy dependence on human input. It requires an accurate assessment of the prevailing stock market conditions. As observed in Sections 5.2 and 5.3, an incorrect stock market assessment can lead to underperformance. Therefore, our model cannot be used by novice investors. Our study also does not consider other realistic aspects of the stock market, such as dividends and transaction costs. The accrued wealth shown in Sections 5.2 and 5.3 assumes that all the assets held in the portfolio are liquid and can be quickly sold or bought. Sometimes, it is not easy to sell a particular asset. Therefore, we should consider the liquidity of assets in the portfolio optimization model. The proposed model is difficult to solve using traditional approaches. In non-traditional approaches such as genetic algorithms, the parameter setting influences the performance. Therefore, we need to test the algorithm for several parameter settings before arriving at a stable and consistent solution.

For further research, there can be various extensions of this paper. A multi-period problem with transaction costs can be considered, which would further add another dimension of reality. We can modify the Grey closeness incidence degree to consider only the downside deviations from the ideal return. We can also consider investor attitudes regarding each asset rather than the stock market as a whole.

Appendix A: Credibility theory

Refer [62] for a detailed explanation of Credibility theory. To summarize, credibility theory is defined as "the study of fuzzy events in which a *credibility measure* determines the chances of a fuzzy event occurrence". Let Θ be a set with at least one element, and U be the set containing all the subsets of Θ. A set function from U to $[0, 1]$ is called a credibility measure. Each member of U is called an event. The chance of occurrence of an event E in U is indicated by

its credibility. A credibility measure satisfies the following four axioms (See Li and Liu [34]),

- **Normality**
- **Monotonicity**
- **Duality**
- **Maximality**

Definition 6 "Let Θ be a non-empty set, U be its power set, and Cr is a credibility measure. Then the triplet (Θ, U, Cr) is called a credibility space."

Definition 7 [62] "A fuzzy variable ξ is a function from a credibility space (Θ, U, Cr) to the set of real numbers $\mathbb{R}$."

Definition 8 "Let ξ be a fuzzy variable defined on the credibility space (Θ, U, Cr). Then, its credibility function v is given by: $v(x) = Cr\{\xi = x\}, \forall x \in \mathbb{R}$."

Definition 9 "The following credibility function defines a triangular fuzzy number, $\xi = (d - \delta, d, d + \eta)$,

$$v(x) = \begin{cases} \frac{x - d + \delta}{2\delta}, & \text{if } d - \delta \leq x \leq d, \\ \frac{d + \eta - x}{2\eta}, & \text{if } d \leq x \leq d + \eta, \\ 0, & \text{otherwise}, \end{cases}$$

where d is the core value of the fuzzy number and $\delta > 0$, $\eta > 0$ are the left and right spreads, respectively (See Fig. 9)."

Definition 10 "Fuzzy variables $\xi_1, \xi_2, ..., \xi_n$ are said to be independent if and only if $v(x_1, x_2, ..., x_n) = \min\limits_{1 \leq i \leq n} v(x_i)$ for all $(x_1, x_2, ..., x_n) \in \mathbb{R}^n$."

Proposition 7 [62] "Let $\xi_1 = (d_1 - \delta_1, d_1, a_1 + \eta_1)$ and $\xi_2 = (d_2 - \delta_2, d_2, d_2 + \eta_2)$ be two independent
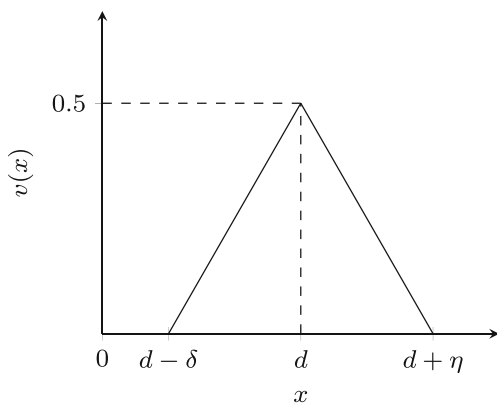
Fig. 9 Triangular fuzzy number's credibility distribution

triangular fuzzy numbers with credibility functions v_1 and v_2, respectively. Then, for any $\lambda_1 > 0$, $\lambda_2 > 0$, we have

$$\begin{aligned} \lambda_1 \xi_1 + \lambda_2 \xi_2 = & ((\lambda_1 d_1 + \lambda_2 d_2) - (\lambda_1 \delta_1 + \lambda_2 \delta_2), \\ & (\lambda_1 d_1 + \lambda_2 d_2), (\lambda_1 d_1 + \lambda_2 d_2) \\ & + (\lambda_1 \eta_1 + \lambda_2 \eta_2)).\text{"} \end{aligned}$$

Definition 11 "(Credibilistic expectation of a fuzzy variable) Let ξ be a fuzzy variable on a credibility space (Θ, U, Cr). Then its expected value is defined by (Liu and Liu [35])

$$E[\xi] = \int_0^{+\infty} Cr\{\xi \geq r\}dr - \int_{-\infty}^0 Cr\{\xi \leq r\}dr,$$

provided that at least one of the two integrals is finite."

Appendix B: Grey Incidence Analysis

A brief introduction to Grey Incidence Analysis has already been provided in Section 2. Refer to [8] for a detailed explanation of this topic.

Definition 12 ([8]) "Assume that X_i is a system factor and its observation value at the ordinal position k is $x_i(k)$, $k = 1, 2, \ldots, n$, then $X_i = (x_i(1), x_i(2), \ldots, x_i(n))$ is referred as the behavioural sequence of the factor X_i."

Remark 9 If k stands for the time order, then $x_i(k)$ is referred to as the observational value of factor X_i at time moment k, and $X_i = (x_i(1), x_i(2), \ldots, x_i(n))$ is the behavioural time sequence of X_i.

Definition 13 ([8]) "Assume that X is a data sequence of a system's behaviour, D is an operator to work on X, and after being applied by the operator D, X becomes the following sequence:

$$XD = (x(1)d, x(2)d, \ldots, x(n)d),$$

then D is referred as the sequence operator."

Definition 14 ([8]) "Let $X_i = (x_i(1), x_i(2), \ldots, x_i(n))$ be a data sequence of a system's behaviour, and D be the sequence operator which satisfies $X_i D = (x_i(1)d, x_i(2)d, \ldots, x_i(n)d)$, and $x_i(k)d = x_i(k) - x_i(1)$, $k = 1, 2, \ldots, n$. Then, D is referred to as the zero-starting point operator, and $X_i D$ is the image of X_i. $X_i D$ is often written as $X_i D = X_i^0 = (x_i^0(1), x_i^0(2), \ldots, x_i^0(n))$."

Table 17 Frequency distribution of AAPL's returns

Return range	Frequency	Return average	Return range	Frequency	Return average
(-10%)-(-9%)	1	-9.9607	(0%)-(1%)	323	0.4562
(-7%)-(-6%)	3	-6.4872	(1%)-(2%)	155	1.4354
(-6%)-(-5%)	3	-5.3614	(2%)-(3%)	37	2.4020
(-5%)-(-4%)	7	-4.4411	(3%)-(4%)	17	3.5402
(-4%)-(-3%)	11	-3.4682	(4%)-(5%)	8	4.4386
(-3%)-(-2%)	48	-2.4094	(5%)-(6%)	2	5.6039
(-2%)-(-1%)	85	-1.4422	(6%)-(7%)	3	6.4760
(-1%)-(0%)	300	-0.4180	(7%)-(8%)	1	7.0422

Appendix C: Construction of fuzzy numbers for asset returns using historical data

In this section, we illustrate the method to construct the fuzzy numbers for the asset return. We determine the fuzzy number that represents AAPL's returns (Table 4). We create frequency distribution of the return as shown in Table 17. The least return class observed during the estimation period was $(-10\%) - (-9\%)$. Therefore, the average return of this class (-9.9607) becomes the left endpoint (i.e., $d - \delta = -9.9607$). Similarly, the highest return class observed during the estimation period was $(7\%) - (8\%)$. Therefore, the average return of this class (7.0422) becomes the right endpoint (i.e., $d + \eta = 7.0422$). For constructing the core of the triangular fuzzy number, we have used a weighted average of the average return of the four classes with the highest frequency. This is worked out to be 0.1412. Thus, $(-9.9607, 0.1412, 7.0422)$ represents the return of AAPL. The fuzzy numbers of the returns of the remaining assets can be constructed in a similar manner.

References

1. Markowitz H (1952) Portfolio selection. J Finance 7(1):77–91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x
2. Speranza MG (1993) Linear programming models for portfolio optimization. Finance 14:107–123
3. Konno H, Yamazaki H (1991) Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. Manag Sci 37(5):519–531. https://doi.org/10.1287/mnsc.37.5.519
4. Li H-Q, Yi Z-H (2019) Portfolio selection with coherent investor?s expectations under uncertainty. Expert Syst Appl 133:49–58. https://doi.org/10.1016/j.eswa.2019.05.008
5. Gupta P, Mehlawat MK, Khan AZ (2020) Multi-period portfolio optimization using coherent fuzzy numbers in a credibilistic environment. Expert Syst. Appl., p 114135. https://doi.org/10.1016/j.eswa.2020.114135
6. Mehlawat MK, Gupta P, Khan AZ (2021) Multiobjective portfolio optimization using coherent fuzzy numbers in a credibilistic environment. Int J Intell Syst 36(4):1560–1594. https://doi.org/10.1002/int.22352
7. Ruiz AB, Saborido Infantes R, Bermúdez J, Luque M, Vercher E (2019) Preference-based evolutionary multi-objective optimization for portfolio selection: A new credibilistic model under investor preferences. J Glob Optim 76:295–315. https://doi.org/10.1007/s10898-019-00782-1
8. Liu S, Yang Y, Forrest J (2017) Grey data analysis : Methods, models and applications. Springer Singapore, Singapore
9. Deng J-L (1982) Control problems of grey systems. Syst Control Lett 1(5):288–294. https://doi.org/10.1016/S0167-6911(82)80025-X
10. Michaud RO (1989) The markowitz optimization enigma: Is ?optimized?optimal? Financial Anal J 45(1):31–42. https://doi.org/10.2469/faj.v45.n1.31
11. Škrinjaric T (2020) Dynamic portfolio optimization based on grey relational analysis approach. Expert Syst Appl 147:113207. https://doi.org/10.1016/j.eswa.2020.113207
12. Babazadeh H, Esfahanipour A (2019) A novel multi period mean-var portfolio optimization model considering practical constraints and transaction cost. J Comput Appl Math 361:313–342. https://doi.org/10.1016/j.cam.2018.10.039
13. Li B, Sun Y, Aw G, Teo KL (2019) Uncertain portfolio optimization problem under a minimax risk measure. Appl Math Model 76:274–281. https://doi.org/10.1016/j.apm.2019.06.019
14. Liu Y-J, Zhang W-G (2019) Possibilistic moment models for multi-period portfolio selection with fuzzy returns. Comput Econ 53(4):1657–1686. https://doi.org/10.1007/s10614-018-9833-6
15. Qin Z (2015) Mean-variance model for portfolio optimization problem in the simultaneous presence of random and uncertain returns. Eur J Oper Res 245(2):480–488. https://doi.org/10.1016/j.ejor.2015.03.017
16. Bonato M (2012) Modeling fat tails in stock returns: a multivariate stable-garch approach. Comput Stat 27(3):499–521
17. Kittiakarasakun J, Tse Y (2011) Modeling the fat tails in asian stock markets. Int Rev Econ Finance 20(3):430–440. https://doi.org/10.1016/j.iref.2010.11.013. Institutional Characteristics and Market Impediments in Asian Capital Markets
18. Campbell JY, Lo AW, MacKinlay AC (2012) The econometrics of financial markets. Princeton University press, Princeton
19. Kamali R, Mahmoodi S, Jahandideh M-T (2019) Optimization of multi-period portfolio model after fitting best distribution. Finance Res Lett 30:44–50. https://doi.org/10.1016/j.frl.2019.03.027
20. Akbay MA, Kalayci CB, Polat O (2020) A parallel variable neighborhood search algorithm with quadratic programming for cardinality constrained portfolio optimization. Knowl-Based Syst 198:105944. https://doi.org/10.1016/j.knosys.2020.105944
21. Yu J-R, Chiou W-JP, Lee W-Y, Lin S-J (2020) Portfolio models with return forecasting and transaction costs. Int Rev Econ Finance 66:118–130. https://doi.org/10.1016/j.iref.2019.11.002

22. Ahmadi-Javid A, Fallah-Tafti M (2019) Portfolio optimization with entropic value-at-risk. Eur J Oper Res 279(1):225–241. https://doi.org/10.1016/j.ejor.2019.02.007

23. Chen B, Zhong J, Chen Y (2020) A hybrid approach for portfolio selection with higher-order moments: Empirical evidence from Shanghai Stock Exchange. Expert Syst Appl 145:113104. https://doi.org/10.1016/j.eswa.2019.113104

24. Chen W, Jiang M, Zhang W-G, Chen Z (2020) A novel graph convolutional feature based convolutional neural network for stock trend prediction. Inf Sci 556:67–94. https://doi.org/10.1016/j.ins.2020.12.068

25. Nesaz HH, Jasemi M, Monplaisir L (2020) A new methodology for multi-period portfolio selection based on the risk measure of lower partial moments. Expert Syst Appl 144:113032. https://doi.org/10.1016/j.eswa.2019.113032

26. Yu J-R, Chiou W-JP, Lee W-Y, Yu K-C (2017) Does entropy model with return forecasting enhance portfolio performance? Comput Ind Eng 114:175–182. https://doi.org/10.1016/j.cie.2017.10.007

27. Zadeh LA (1965) Fuzzy sets. Inf Control 8(3):338–353. https://doi.org/10.1016/S0019-9958(65)90241-X

28. Zadeh LA (2011) Generalized theory of uncertainty: Principal concepts and ideas. In: Fundamental Uncertainty: Rationality and Plausible Reasoning. Palgrave Macmillan UK, pp 104–150

29. Chen W, Xu W (2019) A hybrid multiobjective bat algorithm for fuzzy portfolio optimization with real-world constraints. Int J Fuzzy Syst 21(1):291–307. https://doi.org/10.1007/s40815-018-0533-0

30. Gupta P, Mehlawat M, Yadav S, Kumar A (2019) A polynomial goal programming approach for intuitionistic fuzzy portfolio optimization using entropy and higher moments. Appl Soft Comput 85:105781. https://doi.org/10.1016/j.asoc.2019.105781

31. Gupta P, Mehlawat MK, Kumar A, Yadav S, Aggarwal A (2020) A credibilistic fuzzy DEA approach for portfolio efficiency evaluation and rebalancing toward benchmark portfolios using positive and negative returns. Int J Fuzzy Syst 22(3):824–843. https://doi.org/10.1007/s40815-020-00801-4

32. Liagkouras K, Metaxiotis K (2018) Multi-period mean?variance fuzzy portfolio optimization model with transaction costs. Eng Appl. Artif Intell 67:260–269. https://doi.org/10.1016/j.engappai.2017.10.010

33. Mehlawat MK, Gupta P, Kumar A, Yadav S, Aggarwal A (2020) Multiobjective fuzzy portfolio performance evaluation using data envelopment analysis under credibilistic framework. IEEE Tran Fuzzy Syst 28(11):2726–2737. https://doi.org/10.1109/TFUZZ.2020.2969406

34. Li X, Liu B (2006) A sufficient and necessary condition for credibility measures. Int J Uncert Fuzziness Knowl-Based Syst 14(05):527–535. https://doi.org/10.1142/S0218488506004175

35. Liu B, Liu Y-K (2002) Expected value of fuzzy variable and fuzzy expected value models. IEEE Tran Fuzzy Syst 10(4):445–450. https://doi.org/10.1109/TFUZZ.2002.800692

36. Gupta P, Inuiguchi M, Mehlawat MK, Mittal G (2013) Multiobjective credibilistic portfolio selection model with fuzzy chance-constraints. Inf Sci 229:1–17. https://doi.org/10.1016/j.ins.2012.12.011

37. Jalota H, Thakur M, Mittal G (2017) A credibilistic decision support system for portfolio optimization. Appl Soft Comput 59:512–528. https://doi.org/10.1016/j.asoc.2017.05.054

38. Liu N, Chen Y, Liu Y (2018) Optimizing portfolio selection problems under credibilistic cvar criterion. J Intell Fuzzy Syst 34:335–347. https://doi.org/10.3233/JIFS-171298

39. Liu Y-J, Zhang W-G, Zhang Q (2016) Credibilistic multi-period portfolio optimization model with bankruptcy control and affine recourse. Appl Soft Comput 38:890–906. https://doi.org/10.1016/j.asoc.2015.09.023

40. Mehlawat MK (2016) Credibilistic mean-entropy models for multi-period portfolio selection with multi-choice aspiration levels. Inf Sci 345:9–26. https://doi.org/10.1016/j.ins.2016.01.042

41. Mehlawat MK, Gupta P (2014) Credibility-based fuzzy mathematical programming model for portfolio selection under uncertainty. Int J Inf Technol Decis Mak 13(01):101–135. https://doi.org/10.1142/S0219622014500059

42. Vercher E, Bermúdez JD (2015) Portfolio optimization using a credibility mean-absolute semi-deviation model. Expert Syst Appl 42(20):7121–7131. https://doi.org/10.1016/j.eswa.2015.05.020

43. Tirkolaee EB, Goli A, Faridnia A, Soltani M, Weber G-W (2020) Multi-objective optimization for the reliable pollution-routing problem with cross-dock selection using pareto-based algorithms. J Clean Prod 276:122927. https://doi.org/10.1016/j.jclepro.2020.122927

44. Tirkolaee EB, Aydin NS, Ranjbar-Bourani M, Weber G-W (2020) A robust bi-objective mathematical model for disaster rescue units allocation and scheduling with learning effect. Comput Ind Eng 149:106790. https://doi.org/10.1016/j.cie.2020.106790

45. Tirkolaee EB, Abbasian P, Weber G-W (2021) Sustainable fuzzy multi-trip location-routing problem for medical waste management during the covid-19 outbreak. Sci Total Environ 756:143607. https://doi.org/10.1016/j.scitotenv.2020.143607

46. Bhattacharyya R (2015) A grey theory based multiple attribute approach for r&d project portfolio selection. Fuzzy Inf Eng 7(2):211–225. https://doi.org/10.1016/j.fiae.2015.05.006

47. Huang KY, Jane C-J (2009) A hybrid model for stock market forecasting and portfolio selection based on ARX, grey system and RS theories. Expert Syst Appl 36(3, Part 1):5387–5392. https://doi.org/10.1016/j.eswa.2008.06.103

48. Škrinjaric T, Šego B (2019) Using grey incidence analysis approach in portfolio selection. Int J Financ Stud 7(1):1. https://doi.org/10.3390/ijfs7010001

49. Tsaur R-C (2013) Fuzzy portfolio model with different investor risk attitudes. Eur J Oper Res 227(2):385–390. https://doi.org/10.1016/j.ejor.2012.10.036

50. Sortino FA, Forsey HJ (1996) On the use and misuse of downside risk. J Portf Manag 22(2):35. https://doi.org/10.3905/jpm.1996.35

51. Goel A, Sharma A (2020) Mixed value-at-risk and its numerical investigation. Physica A 541:123524. https://doi.org/10.1016/j.physa.2019.123524

52. Qin Z, Dai Y, Zheng H (2017) Uncertain random portfolio optimization models based on value-at-risk. J Intell Fuzzy Syst 32:4523–4531. https://doi.org/10.3233/JIFS-169216

53. Liu SF (1991) The three axioms of buffer operator and their application. J Grey Syst 3(1):39–48

54. Liu SF, Xie NM (2011) New grey evaluation method based on reformative triangular whitenization weight function. J Syst Eng 26(2):244–250

55. Holland JH (1992) Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control and artificial intelligence. MIT Press

56. Deb K, Agrawal RB (1995) Simulated binary crossover for continuous search space. Complex Syst 9(2):115–148

57. Kumar A, Deb K (1995) Real-coded genetic algorithms with simulated binary crossover: Studies on multimodal and multiobjective problems. Complex Syst 9:431–454

58. Michalewicz Z (2013) Genetic algorithms+ data structures= evolution programs. Springer Berlin Heidelberg, Berlin

59. Sharpe WF (1966) Mutual fund performance. J Bus 39(1):119–138. https://doi.org/10.1086/294846
60. Sortino FA, Van Der Meer R (1991) Downside risk. J Portf Manag 17(4):27. https://doi.org/10.3905/jpm.1991.409343
61. Jagannathan R, Ma T (2003) Risk reduction in large portfolios: Why imposing the wrong constraints helps. J Finance 58(4):1651–1683
62. Li X (2013) Credibility theory. In: Credibilistic Programming: An Introduction to Models and Applications. Springer Berlin Heidelberg, pp 1–29

Mukesh Kumar Mehlawat received a Ph.D. degree in operational research from the University of Delhi, New Delhi, India, in 2011. He is currently teaching with the Department of Operational Research, University of Delhi. He has published several research papers in reputed international journals, including *Fuzzy Sets and Systems, Information Sciences, Expert Systems with Applications, the Journal of Global Optimization, Optimization Letters, TOP, Knowledge-Based Systems, Optimization, the International Journal of Fuzzy Systems, Insurance: Mathematics and Economics, the International Journal of Advanced Manufacturing Technology, the IEEE Transactions on Fuzzy Systems, Applied Intelligence, Applied Soft Computing, Annals of Operations Research, Computers & Operations Research and the International Journal of Information Technology and Decision Making*. In addition, he has co-authored the monograph entitled *Fuzzy Portfolio Optimization* (Springer) under the *Studies in Fuzziness and Soft Computing* series. His current research interests include optimization: theory and applications, soft computing, fuzzy optimization, and financial optimization. Dr.Mehlawat is a recipient of several teaching and research awards.

Pankaj Gupta received a Ph.D. degree in mathematics from the University of Delhi, New Delhi, India, in 2000. He is currently a Full Professor with the Department of Operational Research, University of Delhi. He has published significantly in reputed international journals, including the *European Journal of Operational Research, Fuzzy Sets and Systems, Information Sciences, Expert Systems with Applications, Applied Soft Computing, Applied Mathematics and Computation, the Journal of Global Optimization, the Journal of Optimization Theory and Applications, Optimization Letters, TOP, Applied Intelligence, the Journal of Nonlinear and Convex Analysis, Optimization, Applied Mathematical Modelling, Knowledge-Based Systems, the International Journal of Fuzzy Systems, Insurance: Mathematics and Economics, the International Journal of Advanced Manufacturing Technology, the IEEE Transactions on Fuzzy Systems, Annals of Operations Research, Computers & Operations Research* and the *International Journal of Information Technology and Decision Making*. He has visited several universities abroad for research collaborations. He has co-authored the monograph entitled *Fuzzy Portfolio Optimization* (Springer) under the *Studies in Fuzziness and Soft Computing* series. He has supervised several Ph.D. theses andM.Phil. dissertations. His current research interests include multiple criteria optimization, fuzzy optimization, and portfolio optimization. Dr. Gupta is a recipient of prestigious national and international fellowships. He is an Associate Editor of the *Information Sciences, the International Journal of Fuzzy Systems*, and the *IEEE Transactions on Fuzzy Systems*. In addition, he is an Editor of the journal *Applied Soft Computing*. He has also been a Guest Editor of special issues on optimization and finance in several reputed journals.

Ahmad Zaman Khan is currently pursuing a Ph.D. degree in operational research with the Department of Operational Research, University of Delhi, New Delhi, India. He has published several research papers in reputed international journals, including *Information Sciences, Expert Systems with Applications, and the International Journal of Intelligent Systems*. His current research interests include portfolio optimization, fuzzy optimization, and evolutionary algorithms.