

Research Article

Stochastic Portfolio Selection Problem with Reliability Criteria

Xiangsong Meng¹ and Lixing Yang²

¹Department of Economic Management, North China Electric Power University, Baoding 071003, China

²State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Lixing Yang; lyyang@bjtu.edu.cn

Received 8 October 2015; Accepted 7 February 2016

Academic Editor: Kamel Barkaoui

Copyright © 2016 X. Meng and L. Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Portfolio selection focuses on allocating the capital to a set of securities such that the profit or the risks can be optimized. Due to the uncertainty of the real-world life, the return parameters always take uncertain information in the realistic environments because of the scarcity of the a priori knowledge or uncertain disturbances. This paper particularly considers a portfolio selection process in the stochastic environment, where the return parameters are characterized by sample-based correlated random variables. To decrease the decision risks, three evaluation criteria are proposed to generate the reliable portfolio selection plans, including max-min reliability criterion, percentile reliability criterion, and expected disutility criterion. The equivalent linear (mixed integer) programming models are also deduced for different evaluation strategies. A genetic algorithm with a polishing strategy is designed to search for the approximate optimal solutions of the proposed models. Finally, a series of numerical experiments are implemented to demonstrate the effectiveness and performance of the proposed approaches.

1. Introduction

Portfolio selection problems deal with how to allocate one capital to a given number of securities so that the involved return can be maximized in this process. When the return of each security is a constant, this problem can typically be formulated as a linear programming model and efficiently solved through the simplex method. Due to the uncertainty of the real-world applications, the actual return of each security usually cannot be prespecified in advance. In this case, how to effectively choose portfolio section strategy is a key problem for the investors in order to generate the least-risk plans and produce the expected return as much as possible. Along this line, Markowitz [1, 2] first proposed the mean-variance models in stochastic environments, in which the variance is used to quantify the existing risks in the uncertain return. In this method, a tolerance threshold is usually given for the portfolios. A portfolio selection plan is a safe (or a low-risk) strategy if the variance of its corresponding random return is not greater than this threshold. Based on this approach, a variety of researches about the portfolio selection with either random parameters or fuzzy parameters have been proposed

in the literature, such as Markowitz et al. [3], Gao et al. [4], Yi et al. [5], Xing et al. [6], H. Levy and M. Levy [7], Chiu and Wong [8], A. Palczewski and J. Palczewski [9], Castellano and Cerqueti [10], Fu et al. [11], and Zhang et al. [12]. The second approach to represent the risk proposed by Markowitz [1, 2] is the semivariance. In comparison to the mean-variance method, the mean-semivariance approach can better handle the risks in case of asymmetrical security return distributions. Along this line, interested readers can refer to Zhang et al. [13], Huang [14], Najafi and Mushakhian [15], Yan et al. [16], Yang et al. [17], and so forth. Besides the aforementioned two approaches, some effective chance-constrained methods can also be adopted to characterize the risks such as Huang [14], Huang and Zhao [18], and Li et al. [19].

Different from the variance, semivariance and chance-constrained risk based methods, this paper aims to introduce the reliability term into the portfolio selection problem by using some reliability evaluation indexes, which can be used to optimize the low-risk portfolio strategies according to the real-world applications. We note that the current studies in the literature mainly focus on two specific uncertainties, that is, randomness and fuzziness, in which the returns of the

involved securities are often assumed to be independent random variables or fuzzy variables (without correlations). Differently, this paper intends to propose a new representation method for the random return of each security on the basis of sample-based random data framework. This representation can allow for the correlations among returns of different securities in different periods. For instance, in each week (or month), we can collect realistic return data for individual securities, in which the week-specific or month-specific data can be regarded as having correlations among each other, and each week or month can be regarded as a considered stochastic sample with a specified probability. On the basis of these historical data, it is desirable for decision-makers to produce the reliable portfolio strategies to effectively reduce the investigation risk incurred by various uncertain factors. This research will particularly address this issue. To the best of our knowledge, no existing researches paid much attention to the data representation with inherent correlations.

As addressed above, the majority of existing studies always focus on the mean-variance (or semivariance) based approach to decrease the decision risks, in which the objective is often assumed to find the maximal expected return within the given risk threshold. However, in uncertain portfolio selection, the definition of an optimal portfolio may vary, since there are a large number of optimality criteria to measure the existing risks. Through adopting the sample-based random return to capture the randomness in the investigation process, this paper focuses on introducing three risk criteria into the portfolio selection problem to produce the risk-aversion planning, including the min-max criterion, percentile criterion, and expected disutility criterion, following the classical von Neumann and Morgenstern paradigm of decision under risk in economics [20]. In particular, these risk-aversion criteria have been successfully applied to some real-world applications, for instance, management, economics, and transportation (see [21–24]). In particular, with the sample-base random data representation and these risk measures, it is possible for us to transform the formulated models into linear (or mix-integer) programming, which can be expected to solve through either commercial optimization software or the variants of current exiting algorithms. In addition, to effectively solve the proposed models, we in particular design a genetic algorithm with the polishing strategy to search for the near-optimal solutions. Numerical experiments show that the solution quality can be improved greatly in comparison to the traditional genetic algorithm without the polishing strategy, demonstrating the effectiveness of the proposed approaches.

The remainder of this paper is organized as follows. In Section 2, we formulate the problem of interest with three risk evaluation criteria, and the equivalent linear or mixed-integer programming models are also deduced with specific proofs. In Section 3, a genetic algorithm is designed based on the polishing strategy to solve the proposed models when the model cannot be transformed into linear form or solved by commercial optimization software. Section 4 implements different experiments to test performance of the proposed models and algorithms. In Section 5, a conclusion is made finally.

2. Formulation

Portfolio selection problem involves how to allocate the capital to a given number of securities so that the involved return can be maximized. Because of the uncertainty of the real-world investment environments, the returns of different securities are always set as independent uncertain variables over the entire decision process. This paper aims to handle the portfolio selection problem with a different input data representation, which is termed as sample-based random returns for different securities. Each sample corresponds to a distinct value (vector of values) that random return vector can take. By this method, it can be regarded that the elements in each sample vector are correlated with each other. Particularly, the historical return data over different periods (e.g., month, week) can be used as the input data in producing the least risk investment strategy. In the following, some relevant parameters and notations will be firstly given below.

(i) *Notations and Decision Variables.* Consider the following:

$i = \{1, 2, \dots, I\}$: the index of securities.

$\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$: the index of samples.

p_ω : the probability of sample $\omega \in \Omega$.

ξ_i^ω : the return of the i th security over the sample ω .

ξ_i : the return of the i th security, which is denoted by a random variable.

x_i : the proportion invested for security i .

(ii) *Constraints.* Since x_i denotes the proportion invested for security i , it is required that the sum of all the proportion should be a unity. Thus we need to consider the following constraints; that is,

$$x_1 + x_2 + \dots + x_I = 1. \quad (1)$$

(iii) *Objective Functions.* In the following discussion, we will specify different types of objective functions in formulating the problem of interest. We have a total of N sets of historical data about the returns of different securities due to the portfolio activities in the past. In each set of historical data, we assume that all the returns of individual securities are correlated with each other. With the current set of historical data which may have a lot of uncertainties among different data set, we aim to find the most reliable decision-making plans to decrease the decision-making risks.

Typically, there are a lot of methods to handle the risks of the securities. Currently, we propose three types of methods to clearly handle the inherent uncertainties in the portfolio process. The first model is referred to as the max-min reliability model, detailed below.

2.1. Max-Min Reliable Model of the Portfolio Selection Problem.

According to the a priori information, we have a total of N sample data about the returns of different securities. With this concern, we can actually produce N types of total returns over

different samples for each given portfolio selection strategy. Max-min model aims to find the most reliable portfolio plan across different samples such that the risk of decision-making can be decreased as much as possible. In detail, we denote the each sample-based total return by $R(X, \omega_n), n \in \{1, 2, \dots, N\}$, given below:

$$R(X, \omega_n) = \xi_1^{\omega_n} x_1 + \xi_2^{\omega_n} x_2 + \dots + \xi_I^{\omega_n} x_I, \quad (2)$$

$$n = 1, 2, \dots, N.$$

To integrate these N returns, we first calculate the minimal value over different returns as follows:

$$R(X) = \min_{1 \leq n \leq N} \{R(X, \omega_n)\}. \quad (3)$$

Then the max-min reliable model can be formulated by

$$\begin{aligned} \max \quad & \min_{1 \leq n \leq N} \{\xi_1^{\omega_n} x_1 + \xi_2^{\omega_n} x_2 + \dots + \xi_I^{\omega_n} x_I\} \\ \text{s.t.} \quad & x_1 + x_2 + \dots + x_I = 1 \\ & x_i \geq 0, \quad i = 1, 2, \dots, I. \end{aligned} \quad (4)$$

In this model, it is easy to see that the objective function aims to optimize the lower bound of the N returns over different samples. Typically, this is also a conservative decision-making model for the real-world applications.

Figure 1 gives an illustration for the random return for different solutions. As shown, different solutions might correspond to individual return values for different samples. According to the max-min criterion, the least realization value will be used as the evaluation of this solution. In Figure 1, we consider two solutions X and X' . Typically, since $R(X') \leq R(X)$, solution X is better than X' in this criterion.

Clearly, model (4) is not a linear programming form for the portfolio selection problem. To handle this model effectively through existing commercial optimization software, we can reformulate this model through introducing an auxiliary variable f_X . Then, the max-min reliable model can be transformed into the following linear programming model.

Proposition 1. *The max-min reliable model is equivalent to the following linear programming model:*

$$\begin{aligned} \max \quad & f_X \\ \text{s.t.} \quad & \xi_1^{\omega_n} x_1 + \xi_2^{\omega_n} x_2 + \dots + \xi_I^{\omega_n} x_I \geq f_X, \\ & n = 1, 2, \dots, N \end{aligned} \quad (5)$$

$$x_1 + x_2 + \dots + x_I = 1$$

$$x_i \geq 0, \quad i = 1, 2, \dots, I, \quad f_X \geq 0.$$

Proof. Suppose that $\bar{X} = \{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_I\}$ is the optimal solution for model (4). Then, let

$$\bar{\mu} = \min_{1 \leq n \leq N} \{\xi_1^{\omega_n} \bar{x}_1 + \xi_2^{\omega_n} \bar{x}_2 + \dots + \xi_I^{\omega_n} \bar{x}_I\}. \quad (6)$$

We thus have

$$\xi_1^{\omega_n} \bar{x}_1 + \xi_2^{\omega_n} \bar{x}_2 + \dots + \xi_I^{\omega_n} \bar{x}_I \geq \bar{\mu}, \quad n = 1, 2, \dots, N, \quad (7)$$

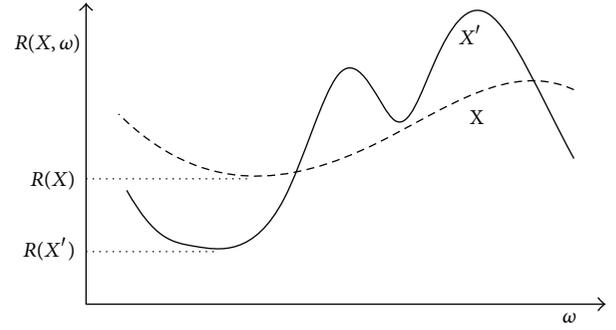


FIGURE 1: An illustration of random return and solution comparison.

which implies $(\bar{X}, \bar{\mu})$ is a feasible solution for model (5). On the other hand, let (X, f_X) be a feasible solution for model (5), and we then have

$$\begin{aligned} \bar{\mu} &= \min_{1 \leq n \leq N} \{\xi_1^{\omega_n} \bar{x}_1 + \xi_2^{\omega_n} \bar{x}_2 + \dots + \xi_I^{\omega_n} \bar{x}_I\} \\ &\geq \min_{1 \leq n \leq N} \{\xi_1^{\omega_n} x_1 + \xi_2^{\omega_n} x_2 + \dots + \xi_I^{\omega_n} x_I\} \geq f_X, \end{aligned} \quad (8)$$

which implies $(\bar{X}, \bar{\mu})$ is the optimal solution for model (5). On the other hand, if $(\bar{X}, \bar{\mu})$ is the optimal solution for model (5), \bar{X} is typically a feasible solution for model (4). Let X be a feasible solution for model (4) and set

$$\mu = \min_{1 \leq n \leq N} \{\xi_1^{\omega_n} x_1 + \xi_2^{\omega_n} x_2 + \dots + \xi_I^{\omega_n} x_I\}. \quad (9)$$

Thus, (X, μ) is a feasible solution for model (5). We have

$$\begin{aligned} \min_{1 \leq n \leq N} \{\xi_1^{\omega_n} \bar{x}_1 + \xi_2^{\omega_n} \bar{x}_2 + \dots + \xi_I^{\omega_n} \bar{x}_I\} &\geq \bar{\mu} \geq \mu \\ &= \min_{1 \leq n \leq N} \{\xi_1^{\omega_n} x_1 + \xi_2^{\omega_n} x_2 + \dots + \xi_I^{\omega_n} x_I\}. \end{aligned} \quad (10)$$

Then \bar{X} is the optimal solution for model (4). The proof is completed. \square

Remark 2. It is easy to see from the max-min model that all the sample-based total returns have been handled equally, since the objective function has no relationship with the probability of each sample. In other words, this model is suitable for decision-makers with extreme risk-aversion.

2.2. Percentile Reliability Model of the Portfolio Selection Problem. Next, we will introduce a percentile reliability model for the portfolio selection problem, which aims to generate portfolio strategies with varying probability confidence levels. We firstly give a definition of the critical value of random variables as follows.

Definition 3. Let η be a random variable and a probability confidence level $\alpha \in (0, 1]$. Then α critical value of η is defined by

$$\eta_{\text{sup}}(\alpha) = \sup \{f \mid \Pr \{\eta \geq f\} \geq \alpha\}. \quad (11)$$

Remark 4. Assume that η is a discrete random variable with realizations $\bar{\eta}_1, \bar{\eta}_2, \dots, \bar{\eta}_L$ and corresponding probabilities $p_{\omega_1}, p_{\omega_2}, \dots, p_{\omega_L}$. If $\bar{\eta}_1 \geq \bar{\eta}_2 \geq \dots \geq \bar{\eta}_L$, we have $\eta_{\text{sup}}(\alpha) = \bar{\eta}_{k'}$, where $k' = \min \{k \mid \sum_{l=1}^k p_{\omega_l} \geq \alpha\}$.

As the return function $R(X, \omega)$ is essentially a random variable for each given solution X . We can use α critical value of $R(X, \omega)$ as the evaluation index of the portfolio selection plan. In detail, we firstly formulate the following model:

$$\begin{aligned} \max \quad & f_X \\ \text{s.t.} \quad & \Pr \{ \xi_1^\omega x_1 + \xi_2^\omega x_2 + \dots + \xi_I^\omega x_I \geq f_X \} \geq \alpha \\ & x_1 + x_2 + \dots + x_I = 1 \\ & x_i \geq 0, \quad i = 1, 2, \dots, I. \end{aligned} \quad (12)$$

In this model, we aim to maximize the α critical value of the random return $R(X, \omega)$. In practice, the optimal objective implies that the sample-based returns are larger than the optimal objective at least with a probability confidence level α . In particular, parameter α is used to denote the risk-aversion level in the practice. A risk-aversion decision-maker would like to set a relatively large parameter α , while small α leads to the high-risk portfolio selection plans.

Figure 2 is given to show the illustration of random return. In detail, for any $\alpha \in (0, 1]$, we have the corresponding critical values of solution X and X' , that is, $R(X, \omega)_{\text{sup}}(\alpha)$ and $R(X', \omega)_{\text{sup}}(\alpha)$, respectively. Since $R(X, \omega)_{\text{sup}}(\alpha) \leq R(X', \omega)_{\text{sup}}(\alpha)$, solution X' is better than X in this reliability criterion.

As shown in the percentile reliability model, the form of the first constraint is typically not a linear constraint, which will probably lead to the increase of complexity in the solution process. Next, we aim to transform this constraint into a linear form through introducing two types of decision variables.

Proposition 5. *The percentile reliability model is equivalent to the following linear programming model (M is a sufficiently large number):*

$$\begin{aligned} \max \quad & g_X \\ \text{s.t.} \quad & \xi_1^{\omega_n} x_1 + \xi_2^{\omega_n} x_2 + \dots + \xi_I^{\omega_n} x_I \geq g_X - y_{\omega_n} M, \\ & n = 1, 2, \dots, N \\ & \sum_n p_{\omega_n} y_{\omega_n} \leq 1 - \alpha \\ & x_1 + x_2 + \dots + x_I = 1 \\ & x_i \geq 0, \quad i = 1, 2, \dots, I \\ & g_X \in \mathbb{R}, \quad y_{\omega_n} \in \{0, 1\}. \end{aligned} \quad (13)$$

Proof. Let X' be the optimal solution for models (12). Then we have

$$\max f_{X'} \geq \max f_X, \quad \forall X. \quad (14)$$

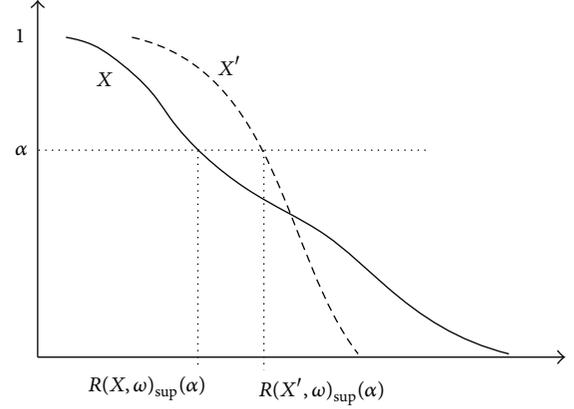


FIGURE 2: Critical value of random return and solution comparison.

Denote $R(X', \omega_1) \geq R(X', \omega_2) \geq \dots \geq R(X', \omega_N)$. Then $\max f_{X'} = R(X', \omega_{\bar{n}})$, where $\bar{n} = \min \{\bar{n} \mid \sum_{n=1}^{\bar{n}} p_n \geq \alpha\}$. Next, let $y_{\omega_n} = 1$ for $n \geq \bar{n}$ and $y_{\omega_n} = 0$ for $n \leq \bar{n}$, and then $(\max f_{X'}, Y, X')$ is a feasible solution for model (13). Denote the optimal solution for model (13) by $(\max g_{X''}, Y'', X'')$. Thus we have

$$\max f_{X'} \leq \max g_{X''}. \quad (15)$$

On the other hand, since $(\max g_{X''}, Y'', X'')$ is a feasible solution for (13), we here denote $A = \{\omega \mid y_{\omega} = 0\}$ and $A^c = \{\omega \mid y_{\omega} = 1\}$. Then

$$\Pr \{A^c\} = \sum_n p_{\omega_n} y_{\omega_n} \leq 1 - \alpha, \quad (16)$$

$$\Pr \{A\} = 1 - \Pr \{A^c\} \geq \alpha.$$

Thus

$$\begin{aligned} \Pr \{ \xi_1^\omega x_1 + \xi_2^\omega x_2 + \dots + \xi_I^\omega x_I \geq (\max g_{X''}) \} &\geq \Pr \{A\} \\ &\geq \alpha, \end{aligned} \quad (17)$$

which implies X'' is a feasible solution for model (12). We then have

$$\max f_{X'} \geq \max g_{X''}. \quad (18)$$

Equations (15) and (18) prove the equivalence of models (12) and (13). \square

Typically, model (13) is a mix-integer programming model, which can be easily solved by some existing commercial optimization software, such as LINGO and CPLEX.

Proposition 6. *Let $\alpha_1, \alpha_2 \in (0, 1]$ be two parameters with $\alpha_1 \leq \alpha_2$ and $\bar{f}_{X'}(\alpha_1), \bar{f}_{X'}(\alpha_2)$ the corresponding optimal objectives. One then has $\bar{f}_{X'}(\alpha_1) \geq \bar{f}_{X'}(\alpha_2)$.*

Proof. Since a large parameter α corresponds to a small feasible solution region, this result is obvious. \square

Proposition 7. *When $\alpha = 1$, the percentile reliable model degenerates to the max-min model.*

Proof. In model (13), if $\alpha = 1$, the second constraint leads to $y_{\omega_n} = 0, n = 1, 2, \dots, N$. Then model (13) will be degenerated to model (5) trivially. The proof is thus completed. \square

2.3. Expected Disutility Model of the Portfolio Selection Problem. In the following, we focus on minimizing the expected disutility function associated with the total profits to produce the least-risk portfolio strategy, following the classical von Neumann and Morgenstern paradigm of decision under risk in economics [20]. For statement simplicity, we next introduce the detailed formulation process of the disutility function in the modeling process. We first give a return target T to denote the upper bound of the historical return data. Then we calculate the following gap functions for the returns of different samples:

$$G(X, \omega_n) = T - \xi_1^{\omega_n} x_1 - \xi_2^{\omega_n} x_2 - \dots - \xi_I^{\omega_n} x_I, \quad (19)$$

$$n = 1, 2, \dots, N.$$

Here, target T is required to be suitably chosen so as to ensure that the gap functions are all positive, which is the discussion focus of this research.

Disutility function will be defined as an increasing (or nondecreasing at least) function over the return gaps that can be either linear or nonlinear. When we take identity mapping function, the return gap itself can be viewed as a special case of the disutility function. With this concern, decision-makers are required to minimize expected disutility when they select the investment proportion of different securities. In this study, we will use an exponential disutility function $D(x) = e^{\alpha x}$ to measure the risk-aversion levels. Then we have the following disutility function for each gap function

$$D(G(X, \omega_n)) = \exp(\alpha \cdot G(X, \omega_n)). \quad (20)$$

Thus, the expected disutility function can be written as follows:

$$E[D(G(X, \omega))] = \sum_{n=1}^N p_n \cdot \exp(\alpha \cdot G(X, \omega_n)). \quad (21)$$

Practically, parameter α in the disutility function represents the level of risk-aversion in the decision-making process. Specifically, if one would like to choose a large parameter α , then he/she should be a risk-aversion decision-maker. If parameter α is close to zero, the decision-maker will be risk-compromise. To clearly state the above idea in the portfolio selection process, we give the following illustrations.

Suppose we have two sample-based random returns, that is, 10 and 8 (unit: thousand dollars), with the probabilities 0.6 and 0.4, respectively. Supposing that return target is $T = 12$, we can obtain the following random gap distribution function (2, 0.6; 4, 0.4). Clearly, this distribution typically corresponds to the expected gap 2.8. We give Table 1 to show the corresponding mapping relationships between parameter α and its corresponding expected disutility.

TABLE 1: Parameter α and its expected disutility value.

α	0.1	0.2	0.5	0.8	1.0	1.5	2	2.2
$E[D(G(X, \omega))]$	1.33	1.79	4.59	12.78	26.27	173.42	1225.14	2702.57
$E[G(X, \omega)]$	2.8							

TABLE 2: Parameter α and value g .

α	0.1	0.2	0.5	0.8	1.0	1.5	2	2.2
g	2.85	2.90	3.05	3.19	3.27	3.44	3.56	3.59

In Table 1, it is obvious that, with the increase of parameter α , the corresponding expected disutility increases drastically. Thus, through minimizing the expected disutility, we can obtain the optimal portfolio selection plan with different risk levels. To further show the relationship between parameter α and risk-preference, we here consider a certain gap ($g, 1.0$), which implies that the gap is g with the probability 1.0. We consider the same relationship between the random gap (2, 0.6; 4, 0.4) and certain gap ($g, 1.0$) with respect to different parameter α (i.e., the gaps represented by (2, 0.6; 4, 0.4) and ($g, 1.0$) correspond to the same evaluation value with each parameter α).

Obviously, as shown in Table 2, a larger parameter α will correspond to a poor value g in the deterministic case. In this sense, the risk-aversion decision-makers usually avoid setting the small values of parameter α to optimize a less risky decision plan.

In the following, the expected disutility model for the portfolio selection problem can be formulated as follows:

$$\begin{aligned} \min \quad & E[D(G(X, \omega))] \\ \text{s.t.} \quad & x_1 + x_2 + \dots + x_I = 1 \\ & x_i \geq 0, \quad i = 1, 2, \dots, I. \end{aligned} \quad (22)$$

Note that the expected disutility function is essentially a nonlinear form in the corresponding optimization model, which might cause potential difficulties by traditional analytical methods in the solution process. In designing the heuristic, we can produce a lower bound model for model (22) by Jensen's inequality. Specifically, if function $f(x)$ is a convex function and ξ is a random variable, we then have $E[f(\xi)] \geq f(E[\xi])$, which leads to the following proposition.

Proposition 8. *According to Jensen inequality, one has the following relationship:*

$$\begin{aligned} \exp\left(\alpha \cdot \sum_{n=1}^N p_n \cdot G(X, \omega_n)\right) &= D(E[G(X, \omega)]) \\ &\leq E[D(G(X, \omega))]. \end{aligned} \quad (23)$$

With this proposition, one can easily deduce an optimization model for providing an effective lower bound to the

original expected disutility model. That is, one introduces the following model:

$$\begin{aligned} \min \quad & D(E[G(X, \omega)]) \\ \text{s.t.} \quad & x_1 + x_2 + \cdots + x_I = 1 \\ & x_i \geq 0, \quad i = 1, 2, \dots, I. \end{aligned} \quad (24)$$

For this model, some existing algorithms can be employed to solve an approximate optimal solution. However, it is difficult for the existing commercial optimization software to solve this model due to the nonlinear form of the objective function. With this concern, we here equivalently transform the nonlinear objective function into a linear form-based equation by introducing the logarithmic function $\ln(\cdot)$. That is, we have

$$\begin{aligned} \ln(D(E[G(X, \omega)])) &= \alpha \cdot E[G(X, \omega)] \\ &= \alpha \cdot \sum_{n=1}^N p_n \cdot (T - \xi_1^{\omega_n} x_1 - \xi_2^{\omega_n} x_2 - \cdots - \xi_I^{\omega_n} x_I). \end{aligned} \quad (25)$$

Note that minimizing $D(E[G(X, \omega)])$ is in essence equivalent to minimizing $\ln(D(E[G(X, \omega)]))$ over the region of feasible solutions. Then, we can use $\ln(D(E[G(X, \omega)]))$ as the alternative objective function in the solution process, and the exponential function can be imposed on the optimal objective value to obtain the optimal objective of model (24).

Remark 9. Typically, the expected disutility model is a generalization of the expected value model, in which the expected gap can be regarded as an evaluation index of each portfolio selection strategy if we adopt the linear disutility function in formulating the objective function.

3. Genetic Algorithm

In particular, for the max-min and percentile reliability models, the commercial optimization software (e.g., Lingo and CPLEX) can be used to produce the exact (or near) optimal solutions for these models. For the expected disutility model, we can adopt the genetic algorithm-base heuristics to generate an approximate optimal solution because of the nonlinear objective function. Genetic algorithm is a kind of evolutionary algorithm proposed by Holland [25] in 1975, which can be used for seeking high-quality solutions of mathematical programming models. With the various technical details, the genetic algorithm has been applied to solving a variety of real-world problems, such as vehicle routing problem, transportation problem, and operations management (For more details, please see Aytug et al. [26], Chang and Sim [27], Yang et al. [24], Chung et al. [28], Xu et al. [29], etc.). In the following, we will introduce the detailed techniques in designing the genetic algorithm to solve the problem considered in this paper.

Solution Representation. In the solution algorithm, we use an I -dimensional array to denote the chromosome of the proposed model, in which each element is randomly generated in

a prespecified positive interval, denoted by $[0, a]$. For clarity, we denote the chromosome by the following form:

$$Y = (y_1, y_2, \dots, y_I), \quad (26)$$

where $y_i, i = 1, 2, \dots, I$ are randomly extracted from interval $[0, a]$. Note that this form is typically infeasible for the portfolio selection process. We then set up the following mapping relationship from a chromosome to a feasible solution:

$$\begin{array}{ccc} Y & & X \\ (y_1, y_2, \dots, y_I) & \longrightarrow & (x_1, x_2, \dots, x_I) \end{array}, \quad (27)$$

in which $x_i = y_i / \sum_{k=1}^I y_k, i = 1, 2, \dots, I$.

By this method, it is clear that each chromosome with form Y corresponds to its unique feasible solution X . Thus, in the following operations, each element of chromosome Y should be controlled in interval $[0, a]$, corresponding to a feasible solution. For instance, we have a total of 5 securities, and Y is produced in the interval $[0, 2]$ as $Y = (1, 2, 1, 2, 2)$. Then the corresponding feasible solution X is

$$X = (0.125, 0.25, 0.125, 0.25, 0.125). \quad (28)$$

In the beginning of the algorithm, we need to produce a total of Pop_size chromosomes in the population as the initial input data, denoted by $Y_1, Y_2, \dots, Y_{\text{Pop_size}}$.

Selection Operations. In the solution process, it is required to generate a new population for the selection operations. In detail, we first rank the chromosomes in a good-to-bad sequence, which is also denoted by $Y_1, Y_2, \dots, Y_{\text{Pop_size}}$ for notation simplicity. Then the evaluation value of each chromosome will be given by

$$\text{Eval}(Y_k) = \rho(1 - \rho)^{k-1}, \quad k = 1, 2, \dots, \text{Pop_size}, \quad (29)$$

where $\rho \in (0, 1)$ is a prespecified parameter in determining evaluation value of each chromosome. With these evaluation values, we select the chromosomes by using the roulette wheel. Specifically, we first produce a sequence by setting a sequence $\{q_k\}_{k=0}^{\text{Pop_size}}$, where

$$\begin{aligned} q_0 &= 0, \\ q_k &= \sum_{l=1}^k \text{Eval}(Y_l), \quad k = 1, 2, \dots, \text{Pop_size}. \end{aligned} \quad (30)$$

Then, we implement the following procedure for a total of Pop_size times: randomly generate a number r in interval $(0, q_{\text{Pop_size}})$; if there exists an index k such that $q_{k-1} \leq r \leq q_k$, then chromosome Y_k will be put into the new population for the following crossover and mutation operations. Thus, the newly generated population also contains a total of Pop_size chromosomes even if some chromosomes are selected repeatedly.

Crossover Operations. Crossover operations aim to produce the new chromosomes for the population in order to find the

approximate optimal solution as soon as possible. For this purpose, we firstly specify the chromosomes that take part in crossover operations. Note that this operation is performed on the basis of a predetermined crossover probability P_c . That is, the following procedure is employed to determine the crossover chromosomes: for each Y_k , randomly generate a number r in interval $(0, 1)$; if $r \leq P_c$, then Y_k will be selected for crossover operations. Typically, a total of $P_c \cdot \text{Pop_size}$ chromosomes can be expectedly chosen for the crossover operations. We perform the crossover operations based on each pair of selected chromosomes. Let Y' and Y'' be two selected chromosomes. We firstly generate a crossover parameter $\lambda \in (0, 1)$, and the newly generated chromosomes \bar{Y}' and \bar{Y}'' are given as follows:

$$\begin{aligned}\bar{Y}' &= \lambda Y' + (1 - \lambda) Y'', \\ \bar{Y}'' &= (1 - \lambda) Y' + \lambda Y''.\end{aligned}\quad (31)$$

It is easy to see that since Y' and Y'' are generated in interval $[0, a]$, \bar{Y}' and \bar{Y}'' should also lie in this interval, which leads to the feasibility of \bar{Y}' and \bar{Y}'' . We then replace Y' and Y'' by \bar{Y}' and \bar{Y}'' , respectively, in the initial population.

Mutation Operations. Mutation operations intend to increase the diversity of the chromosomes in the population so as to avoid premature convergence. This operation is also performed under the consideration of the mutation probability P_m . Specifically, for each Y_k , randomly generate a number r in interval $(0, 1)$; if $r \leq P_m$, then Y_k will be selected for mutation operations. In this process, a total of $P_m \cdot \text{Pop_size}$ chromosomes will be expectedly chosen. For each chosen chromosome Y , we implement the mutation operations according to the following procedure: randomly generate a mutation vector d in interval $[-1, 1]$; choose a suitable step size M such that the newly produced chromosome Y' (i.e., $Y' = Md + Y$) is feasible; replace Y by Y' in the population.

Polishing Chromosomes. Note that, in this solution representation, each security is allowed to own an investment ratio for any feasible solution. Accordingly, in the near-optimal solution, it is possible that some securities have relatively small ratios (probably close to zero), which is typically undesirable in the real-world applications. To avoid this case, we herein give a ratio threshold for the each security in order to improve the solution quality. If the ratio of some security is less than this threshold, we will not consider investing this security in the solution. With this treatment, we can select an approximate optimal solution with guaranteed quality and desirability.

Procedure of Genetic Algorithm. Next, we will give the detailed procedure of the designed algorithm.

Step 1. Initialize Pop_size chromosomes in interval $[0, a]$ for the population.

Step 2. Compute the objective value of each chromosome in the population.

Step 3. Compute the evaluation value according to the good-to-bad sequence of chromosomes.

Step 4. Perform the selection operations over the population.

Step 5. Perform the crossover operations.

Step 6. Perform the mutation operations.

Step 7. Repeat Step 2 to Step 6 for a given number of times.

Step 8. Output the best solution encountered as the approximate optimal solution.

4. Numerical Examples

To test the performance of the proposed models, we will implement different sets of numerical experiments in the following discussion. All the experiments are performed on a personal computer with 4 GB memory and 1.60 GHz processors.

Example 1. In this set of experiments, we assume there are 10 securities. Ten samples are given in Table 3 to show the randomness of the security returns in the proposed models, listed in Table 3.

In this problem, we need to determine the optimal investment ratio of each security. Since the max-min reliability model is a special case of the percentile reliability model, only percentile model will be employed to test the performance of the proposed approaches. Note that the percentile model is equivalent to a linear programming model. In implementations, we can use the CPLEX solver in GAMS commercial software to solve the equivalent mix-integer linear programming models.

In this set of numerical experiments, we test the performance of the optimal objective function with respect to different probability confidence levels, that is, $\alpha = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0$. In the GAMS software, we set the allowable relative gap parameter as $\text{OPTCR} = 0.001$. That is, when a solution is found within the relative gap 0.1% from the estimated optimal objective value, it will be outputted as a near-optimal solution for the model. With this parameter, all the experiments can output the corresponding optimal solution with zero relative gaps. The variation tendency of the optimal objectives is given in Figure 3, where the optimal objective takes decreasing tendency with the increase of the probability confidence level α . This result can be explained suitably because a large probability confident level will lead to a small region of feasible solutions, corresponding to a smaller optimal objective, which also coincides with the results in Proposition 6. Note that when we set $\alpha = 1.0$, the percentile reliable model outputs the same optimal objective to the max-min reliability model with the same optimal solution.

In the following, we aim to demonstrate the variation of different solutions with respect to varied probability confidence levels, which are given in Figure 4. It is easy to see that, for all the considered probability confidence

TABLE 3: Sample data for returns of different securities.

Sample	p_ω	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$	$i = 6$	$i = 7$	$i = 8$	$i = 9$	$i = 10$
ω_1	0.1	5.3	8.4	7.3	6.8	5.6	3.5	2.6	4.9	4.3	3.5
ω_2	0.1	4.4	6.5	5.6	7.6	5.2	7.6	8.5	5.6	5.3	4.8
ω_3	0.1	7.8	6.8	8.2	9.4	3.5	3.2	3.7	8.2	3.6	3.5
ω_4	0.1	5.8	7.4	7.4	6.4	5.6	8.5	2.6	7.9	4.6	4.5
ω_5	0.1	4.3	7.8	9.4	5.7	8.2	7.6	3.5	3.6	5.7	6.8
ω_6	0.1	5.6	3.5	2.6	7.9	9.3	3.6	4.2	6.4	8.8	8.6
ω_7	0.1	8.2	7.6	3.5	8.6	8.8	3.5	6.4	7.6	8.9	7.5
ω_8	0.1	3.5	9.2	4.7	8.2	8.4	4.8	7.5	8.9	8.2	8.3
ω_9	0.1	6.5	5.5	5.4	7.6	8.6	4.6	7.8	9.4	7.8	6.4
ω_{10}	0.1	5.8	4.7	3.9	6.5	8.7	8.2	8.3	6.4	7.9	4.8

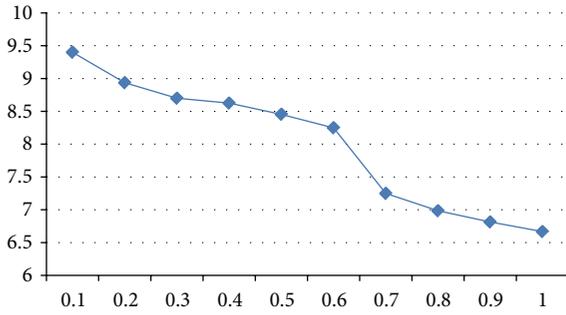


FIGURE 3: Critical value of random return and solution comparison.

levels, the involved nonzero optimal decision variables are $x_2, x_3, x_4, x_5, x_6, x_7, x_8$, respectively. In this figure, the x -axis represents the probability confidence levels and y -axis denotes the returned optimal solutions for each decision variable. Clearly, in this set of numerical experiments, the most active decision variables are x_4, x_5 , which have a total of five and eight times for investment among these ten experiments. When we set a relative small probability confidence level (e.g., $\alpha \leq 0.6$), at most two securities can be selected for investment. For instance, for $\alpha = 0.1$ or 0.3 , only one security is adopted in the optimal solution, corresponding to x_4 or x_5 , respectively. When much larger probability confidence levels are adopted, more than three securities need to be invested. For example, when $\alpha = 0.9$, securities 3, 4, 5, and 6 are needed to be invested with different optimal ratios.

Example 2. Next, the second set of experiments will be performed on the expected disutility model to show the performance of the proposed approaches. Since the expected disutility model is a nonlinear model, it is difficult to use commercial optimization software to solve it. With this concern, we adopt the proposed genetic algorithm to solve the approximate optimal solutions.

In this experiment, we take 20 securities into consideration for possible investment. The sample-based random returns are randomly generated in different intervals given in Table 4.

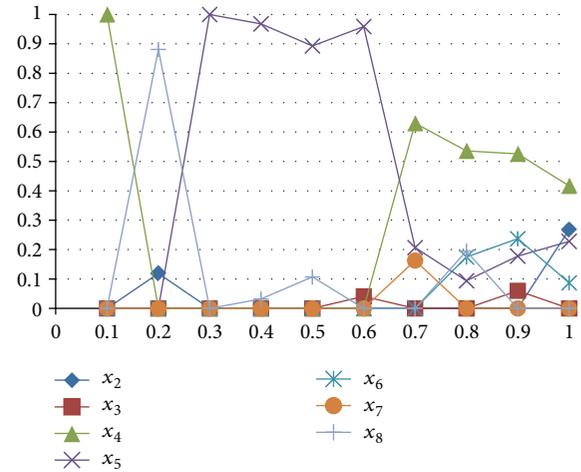


FIGURE 4: Variation of different decision variables in the experiments.

As we need to produce a total of N samples in the problem, we also use the randomly generated sample probabilities as the occurrence chance of each sample. The following procedure is used to determine the probability of each sample.

Step 1. Randomly generate a number in interval $[0, 1]$ for each sample n , denoted by q_n .

Step 2. Compute the probability of each sample by $p_{\omega_n} = q_n / \sum_{n=1}^N q_n$, $n = 1, 2, \dots, N$.

In this set of experiments, two solution strategies are adopted to generate approximated optimal solutions, given below.

Strategy 1. Use the chromosomes representation without the polishing strategy.

Strategy 2. Use the chromosomes representation with the polishing strategy.

In this example, we randomly generate ten samples from the given intervals. For these two chromosome representation strategies, different experiments are performed with

TABLE 4: The intervals for generating sample-based random return data.

Security	Interval
1	[2, 5]
2	[6, 8]
3	[4, 5]
4	[1, 3]
5	[8, 10]
6	[5, 7]
7	[8, 9]
8	[6, 8]
9	[4, 10]
10	[2, 8]
11	[6, 9]
12	[6, 8]
13	[4, 5]
14	[3, 5]
15	[4, 6]
16	[2, 4]
17	[3, 6]
18	[4, 7]
19	[6, 8]
20	[8, 9]

different parameter settings, which are listed in Table 5. We test these two solution strategies with randomly generated crossover probability and mutation probability. To give a straightforward understanding of the returned best objectives, the corresponding values for these two strategies are listed, where the gaps between two results are also given to demonstrate the algorithmic characteristics. Specifically, Gap 1 represents the difference of results between Strategies 1 and 2, which essentially denotes the improvement degree of the result by Strategy 2 in comparison to that by Strategy 1. On the other hand, Gap 2 is calculated based on Strategy 2, which also shows the solution quality improvement of Strategy 2 when compared to Strategy 1. These two gaps are computed according to the following equations:

$$\begin{aligned}
 \text{Gap 1} &= \frac{(\text{Result of Strategy 1} - \text{Result of Strategy 2})}{\text{Result of Strategy 1}} \\
 &\quad \times 100\%, \\
 \text{Gap 2} &= \frac{(\text{Result of Strategy 1} - \text{Result of Strategy 2})}{\text{Result of Strategy 2}} \\
 &\quad \times 100\%.
 \end{aligned}
 \tag{32}$$

In these results, we calculate different gaps for individual parameters and strategies. Clearly, the first strategy in genetic algorithm has relatively robust characteristics as the returned best objective varies slightly in interval [1500, 1600]. However, this strategy might return a practically undesirable solution since all the securities correspond to their nonzero ratios even with a very small value. This case is practically

TABLE 5: The computational results for different parameters.

Test index	P_c	P_m	Gen.	Strategy 1	Strategy 2	Gap 1	Gap 2
1	0.7	0.8	500	1570.89	324.86	79.32%	383.56%
2	0.5	0.5	500	1587.18	324.86	79.53%	388.57%
3	0.6	0.7	500	1541.26	284.92	81.51%	440.94%
4	0.6	0.8	500	1541.26	324.86	78.92%	374.44%
5	0.7	0.5	500	1570.51	261.08	83.38%	501.54%
6	0.7	0.6	500	1587.18	242.91	84.70%	553.40%
7	0.4	0.6	500	1573.65	324.86	79.36%	384.41%
8	0.7	0.9	500	1570.89	286.06	81.79%	449.15%
9	0.7	0.4	500	1587.18	309.62	80.49%	412.62%
10	0.8	0.5	500	1570.81	306.47	80.49%	412.55%

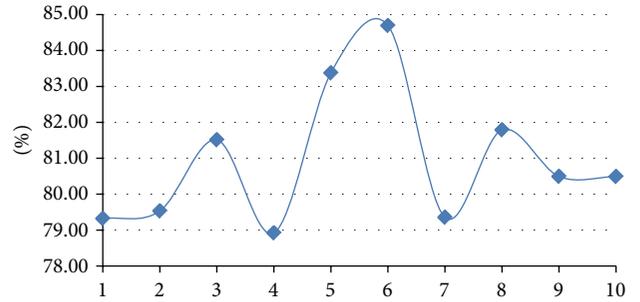


FIGURE 5: The variation of Gap 1 in different tests.

undesirable. However, we can improve this situation through the polishing strategy, and the outputted results can be improved to a great extent. As shown, in comparison to Strategy 1, Strategy 2 can reduce the returned objective up to almost 80%. In other words, the returned near-optimal objectives in Strategy 1 are about 4-5 times of the results in Strategy 2, where the best near-optimal objective turns out to be 242.91. These computational results show that the polishing strategy is more effective than the solution Strategy 1 for each set of parameters. For comparison convenience, we also give Figure 5 to show the variation of different Gap 1, in which the largest gap (i.e., 84.70%) occurs in the sixth experiment (the x -axis denotes the test indexes and y -axis denotes the gaps).

5. Conclusions

Using the sample-based random data to capture uncertainties of the decision parameters, we developed three different models for the portfolio selection problem with stochastic characteristics of each security return, including max-min reliability model, percentile reliability model and expected disutility model. With our random data representation methods, the max-min and percentile reliability models could be transformed into their linear forms through introducing different auxiliary variables, which could be easily solved by the commercial optimization software. The expected disutility model was formulated based on the disutility function. A lower bound linear programming problem was also deduced for this model based on our random data representation.

To effectively solve the proposed models, we proposed a polishing strategy-based genetic algorithm to produce the approximated optimal solutions. The numerical examples are implemented to specify the detailed characteristics of the proposed models and algorithm.

It is worth mentioning that we propose three models with different decision-making criteria. As for the percentile reliability mode, if the decision-maker is risk-appetite, he/she can take a small parameter α ; otherwise, a larger parameter should be considered. In extreme case, if we set $\alpha = 1$, the percentile reliability model will degenerate to the max-min model reliability model. In this sense, the max-min reliability model is a special case of the percentile reliability model, and it is more suitable for decision-makers with extreme risk-aversion. This situation also holds for the expected disutility model with the exponential disutility function. In general, we cannot determine which model is the best in the real-world applications, and the use of different models is closely related to the preferences of the decision-makers.

Further research can focus on the following two aspects.

- (1) Different uncertainties may occur for the real-world applications with either enough samples or no enough samples. Thus, the reliable portfolio selection model with other uncertainties can be a new topic in the further research.
- (2) The proposed models can be easily generalized to the more complicated situations with variance or semivariance threshold constraints. The characteristics of these problems can also be investigated in future study.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] H. Markowitz, "Portfolio selection," *The Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.
- [2] H. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*, John Wiley & Sons, New York, NY, USA, 1959.
- [3] H. Markowitz, P. Todd, G.-L. Xu, and Y. Yamane, "Computation of mean-semivariance efficient sets by the critical line algorithm," *Annals of Operations Research*, vol. 45, no. 1–4, pp. 307–317, 1993.
- [4] J. Gao, D. Li, X. Cui, and S. Wang, "Time cardinality constrained mean-variance dynamic portfolio selection and market timing: a stochastic control approach," *Automatica*, vol. 54, pp. 91–99, 2015.
- [5] L. Yi, X. Wu, X. Li, and X. Cui, "A mean-field formulation for optimal multi-period mean-variance portfolio selection with an uncertain exit time," *Operations Research Letters*, vol. 42, no. 8, pp. 489–494, 2014.
- [6] X. Xing, J. Hu, and Y. Yang, "Robust minimum variance portfolio with L-infinity constraints," *Journal of Banking & Finance*, vol. 46, no. 1, pp. 107–117, 2014.
- [7] H. Levy and M. Levy, "The benefits of differential variance-based constraints in portfolio optimization," *European Journal of Operational Research*, vol. 234, no. 2, pp. 372–381, 2014.
- [8] M. C. Chiu and H. Y. Wong, "Mean-variance portfolio selection with correlation risk," *Journal of Computational and Applied Mathematics*, vol. 263, pp. 432–444, 2014.
- [9] A. Palczewski and J. Palczewski, "Theoretical and empirical estimates of mean-variance portfolio sensitivity," *European Journal of Operational Research*, vol. 234, no. 2, pp. 402–410, 2014.
- [10] R. Castellano and R. Cerqueti, "Mean-variance portfolio selection in presence of infrequently traded stocks," *European Journal of Operational Research*, vol. 234, no. 2, pp. 442–449, 2014.
- [11] C. Fu, A. Lari-Lavassani, and X. Li, "Dynamic mean-variance portfolio selection with borrowing constraint," *European Journal of Operational Research*, vol. 200, no. 1, pp. 312–319, 2010.
- [12] W.-G. Zhang, X.-L. Zhang, and W.-L. Xiao, "Portfolio selection under possibilistic mean-variance utility and a SMO algorithm," *European Journal of Operational Research*, vol. 197, no. 2, pp. 693–700, 2009.
- [13] W.-G. Zhang, Y.-J. Liu, and W.-J. Xu, "A possibilistic mean-semivariance-entropy model for multi-period portfolio selection with transaction costs," *European Journal of Operational Research*, vol. 222, no. 2, pp. 341–349, 2012.
- [14] X. Huang, "Portfolio selection with a new definition of risk," *European Journal of Operational Research*, vol. 186, no. 1, pp. 351–357, 2008.
- [15] A. A. Najafi and S. Mushakhian, "Multi-stage stochastic mean-semivariance-CVaR portfolio optimization under transaction costs," *Applied Mathematics and Computation*, vol. 256, pp. 445–458, 2015.
- [16] W. Yan, R. Miao, and S. Li, "Multi-period semi-variance portfolio selection: model and numerical solution," *Applied Mathematics and Computation*, vol. 194, no. 1, pp. 128–134, 2007.
- [17] S.-C. Yang, T.-L. Lin, T.-J. Chang, and K.-J. Chang, "A semi-variance portfolio selection model for military investment assets," *Expert Systems with Applications*, vol. 38, no. 3, pp. 2292–2301, 2011.
- [18] X. Huang and T. Zhao, "Mean-chance model for portfolio selection based on uncertain measure," *Insurance: Mathematics and Economics*, vol. 59, pp. 243–250, 2014.
- [19] X. Li, Z. Qin, and L. Yang, "A chance-constrained portfolio selection model with risk constraints," *Applied Mathematics and Computation*, vol. 217, no. 2, pp. 949–951, 2010.
- [20] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, USA, 1944.
- [21] T. Xing and X. Zhou, "Reformulation and solution algorithms for absolute and percentile robust shortest path problems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 2, pp. 943–954, 2013.
- [22] H. Huang and S. Gao, "Optimal paths in dynamic networks with dependent random link travel times," *Transportation Research Part B: Methodological*, vol. 46, no. 5, pp. 579–598, 2012.
- [23] L. Yang and X. Zhou, "Constraint reformulation and a Lagrangian relaxation-based solution algorithm for a least expected time path problem," *Transportation Research Part B: Methodological*, vol. 59, pp. 22–44, 2014.
- [24] L. Yang, Y. Zhang, S. Li, and Y. Gao, "A two-stage stochastic optimization model for the transfer activity choice in metro networks," *Transportation Research Part B: Methodological*, vol. 83, pp. 271–297, 2016.
- [25] J. H. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, Mich, USA, 1975.

- [26] H. Aytug, M. Khouja, and F. E. Vergara, "Use of genetic algorithms to solve production and operations management problems: a review," *International Journal of Production Research*, vol. 41, no. 17, pp. 3955–4009, 2003.
- [27] C. S. Chang and S. S. Sim, "Optimising train movements through coast control using genetic algorithms," *IEE Proceedings—Electric Power Applications*, vol. 144, no. 1, pp. 65–73, 1997.
- [28] J.-W. Chung, S.-M. Oh, and I.-C. Choi, "A hybrid genetic algorithm for train sequencing in the Korean railway," *Omega*, vol. 37, no. 3, pp. 555–565, 2009.
- [29] X. Xu, K. Li, L. Yang, and J. Ye, "Balanced train timetabling on a single-line railway with optimized velocity," *Applied Mathematical Modelling*, vol. 38, no. 3, pp. 894–909, 2014.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

