Paper:

Using Granular Representation of Time Series to Spread Risk of Portfolio Selection

Haoran Ren[1,⋆], Benyu Zhang[1], and Xuemei Yang[2]

[1] State Grid Maintenance co.of SEPC, Taiyuan, 030032, China
[2] Software Development Center of ICBC, Beijing, 100094, China
⋆ E-mail: cgyrenhaoran@163.com

Abstract. In the process of portfolio selection decision-making, investors always want to achieve a portfolio selection result that is as evenly distributed as possible to spread the risk from financial market. In this paper, a granulation method for time series data is proposed to improve the quality of portfolio decision results. In the method, the time series data corresponding to the investment objects are first constructed to information granules which are represented by fuzzy numbers as their mathematical representation, and then the portfolio decision model is employed to give a corresponding decision result after a comprehensive analysis of these information granules. In the process of constructing information granules, in order to make information granules which associate with the corresponding investment objects as informative as possible, the principle of justifiable granularity is used to achieve a compromise between justifiability and specificity of the information granules. The experiments using data from China Shanghai Stock Exchange clearly show that the proposed method improves the dispersion of portfolio investment.

Keywords: Data Analysis, Portfolio Selection, Time Series, Information Granule, L-R Fuzzy Number

1. Introduction

Modern portfolio theory refers to a portfolio composed of several kinds of investment objects, whose return is the weighted average of the returns of these investment objects, but its risk is not the weighted average risk of these investment objects. Investors, investment advisers, and fund managers, though, would also take into account risk factors before the "mean-variance" portfolio framework has been proposed by Markowitz [1], [8]. However, due to the lack of effective measurement of risk, they can only take focus on the return of an investment. Markowitz puts forward the assumption that typical investors are risk averse, and they prefer to avoid risks as much as possible while pursuing high expected returns. And then, in order to solve the problem of measuring the risk of assets, Markowitz uses expected value to represent the investment return (rate) and variance (or standard deviation) to represent the risk of return, which is known as the mean-variance model. Markowitz indicates that portfolio can reduce the unsystematic risk of a particular investment, and the core of the portfolio is the dispersion of capital. Compared with the previous investment theory, diversification investment introduces an important change to the principle of risk and return equivalence. That is, diversification investment can reduce risk without reducing return. Since Markowitz put forward the idea that the most efficient portfolio selection result can be determined by means of mean-variance analysis, numerous mean-variance models have been established under the framework of probability theory, and the return of each investment object is described as a random variable.

However, Hasuike *et al.* [2] pointed out that the risk in an actual financial market is not limited to the probability factors, and many non-probabilistic factors, such as vagueness and ambiguity, generally exist in the investment environment. Because of the existence of these fuzzy factors, researchers gradually realize that the portfolio model based on probability theory is limited in the analysis of fuzzy uncertainty. Therefore, using fuzzy sets theory to integrate historical data, expert knowledge and subjective opinions given by investors into the portfolio selection model can more effectively quantify the changes of returns and risks in financial markets. Different from mean-variance models built under the framework of probability theory, the return of each investment object should be described as fuzzy variables in advance when fuzzy sets theory is employed to describe the process of portfolio selection. And in most cases the fuzzy statistical method is used to transform the time series data into a fuzzy variable [3], [4]. Using fuzzy statistical method to construct fuzzy variables can objectively reflect the membership degree of elements in the universe relative to fuzzy concepts, but it requires high accuracy of historical statistical data and is greatly affected by outliers. Therefore, the small changes of time series data lead to the obvious change of the constructed fuzzy variables, which makes a drastic change of the portfolio weight given by the fuzzy mean-variance model. In order to obtain a fuzzy variable with stable structure and good semantics [5], we use granulation method [6], [7] instead of fuzzy statistical method to capture the essence of time series data. Here, these fuzzy variables obtained by the granulation method are called information granules and given a deeper meaning. These fuzzy variables (information granules) are the abstract sets of some elements with indistinguishability, similarity and correlation, and can be used as the key components of knowledge representation and processing.

The structure of this paper is organized as follows. In Section 2, we make a brief introduction to the principle of justifiable granularity and genetic algorithm (GA). In Section 3, a granulation method and a granular-based portfolio selection model are elaborated. Two optimization processes for the entire portfolio selection approach are given in Section 4. The comparative experiment using data sets from China Shanghai Stock Exchange is reported in Section 5 to demonstrate the effectiveness of the proposed method. Finally, conclusions are given in Section 6.

2. Preliminaries

In this section, we give a brief review of some preparatory knowledge including the principle of justifiable granularity and Genetic algorithm (GA), which is needed for the subsequent analysis.

2.1. Principle of Justifiable Granularity

The principle of justifiable granularity [6] is one of the most useful concepts in information granulation, which can provide guidance for building information granules with stable structure and good semantics [5]. The principle of justifiable granularity emphasizes that a theoretical variable is a meaningful information granule formed on the basis of available experimental evidence. Therefore, there are two intuitively compelling requirements, which are experimental evidence and semantic meaning, need to be met. As an example, we take the information granule Ω as an example to illustrate experimental evidence and semantic meaning. Experimental evidence indicates that for the information granule Ω, the higher the sum of membership degrees of the data belonging to Ω, the higher the justifiability of the information granule. And semantic meaning requires that this information granule Ω should come with a well-defined semantics (meaning) at the same time, which means that Ω should be as specific as possible. Obviously, these two demands are in conflict. The increase of justifiability inevitably means the decrease of specificity of the information granule. Therefore, the key point of a granulation method is how to balance these two requirements.

2.2. Genetic Algorithm

Genetic algorithm (GA), which is designed according to the law of biological evolution in the nature, is first proposed by Holland [8]. The essence of genetic algorithm is to simulate the natural selection and genetic mechanism in the process of biological evolution by using Darwin's theory of biological evolution, so its core idea is to search for the optimal solution by simulating the natural evolution process. The algorithm transforms the solving process of the problem into a process similar to the crossover and mutation of chromosome genes in biological evolution by means of mathematical method and computer simulation. When solving complex combinatorial optimization problems, compared with some conventional optimization algorithms, it usually can get better optimization results quickly. The basic calculation process of GA includes six main steps: initialization, individual evaluation, selection operation, crossover operation, mutation operation, and termination condition judgment. It is worth mentioning that selection, crossover and mutation are the three basic operators used to simulate the genetic operation, which directly determine the quality of the optimization results.

3. Portfolio Selection Driven by Information Granules

For a specific investment environment, suppose there are N investment objects waiting to be invested. Then we can know from the historical records that a time series can be constituted by the corresponding historical return data of each investment object such that $S_i = \{s_{i1}, s_{i2}, \ldots, s_{in}\}$, and $i = 1, 2, \ldots, N$.

3.1. Build Information Granule

According to the principle of justifiable granularity[6], [7] both experimental evidences (Legitimacy) and semantic meaning (Meaning)[5] are two indispensable factors when the time series data is considered to be converted to an information granule. Experimental evidences require that the constructed information granule should be supported by the observed data (time series of each investment object) as much as possible. In other words, the more data points the information granule Ω covers in the original time series S_i, the better the experimental evidence that the information granule reflects the observation data, and the more legal this information granule is. In the process of granulation, experimental evidence can be described by a coverage function as follows

$$Cov(\Omega_i) = \sum_{j=1}^{n} \mu_i(s_{ij}),\ i = 1, 2, \ldots, N \quad (1)$$

where Ω_i is represented by a L-R fuzzy number, and μ_i is its corresponding membership function. However, semantic meaning emphasizes that information granules should have the ability to explain the main characteristics of time series, which requires that the constructed information granule $\Omega_i = \mu_i(x)$ should be as specific as possible. Semantic meaning reveals that in order to make the information granule have a high quality, the granulation process does not need to cover all the data coming from the original time series S_i. Similarly, semantic meaning can be quantified by a specific function such that

$$Spe(\Omega_i) = \int_{-\infty}^{+\infty} \mu_i(x)\,dx,\ i = 1, 2, \ldots, N \quad (2)$$

It can be readily checked that there is a contradiction between experimental evidence and semantic meaning.

For a specific information granule Ω_i, since $Cov(\Omega_i)$ and $Spe(\Omega_i)$ have a similar variation trend, the cost of one index getting better is another index getting worse. In order to form a sound compromise between these two requirements, we introduce an auxiliary function such that

$$I(x) = 1 - \frac{x}{\max_j \{s_{ij}\} - \min_j \{s_{ij}\}} \quad \ldots\ldots \quad (3)$$

and then let

$$\overline{Spe}(\Omega_i) = I(Spe(\Omega_i)),\ i = 1,2,\ldots,N \quad \ldots \quad (4)$$

It is easy to verify that $Cov(\Omega_i)$ and $\overline{Spe}(\Omega_i)$ are positive indicators in the same time but have a different variation trend, which means that through a game between $Cov(\Omega_i)$ and $\overline{Spe}(\Omega_i)$, an optimal equilibrium point can be found as the optimal information granule.

Based on the above analysis, the following expression can be used as a comprehensive performance index

$$Q(\Omega_i) = Cov(\Omega_i) \cdot \overline{Spe}(\Omega_i),\ i = 1,2,\ldots,N \quad . \quad (5)$$

And with the aid of Eq. (5), the problem of building fuzzy variables can be transformed into the following optimization problem, which can be solved by using optimization algorithms (such as genetic algorithm (GA))

$$\begin{cases} \max & Q(\Omega_i) \\ \text{s.t.} & \min \mu(x) = 0 \\ & \max \mu(x) = 1 \end{cases},\ i = 1,2,\ldots,N \quad .\ . \quad (6)$$

Consider that the information granular Ω_i can be represented by a L-R fuzzy number such that

$$\Omega_i = \begin{cases} 1, & a_i \le x \le b_i \\ L(x;p_i,a_i), & x < a_i \\ R(x;q_i,b_i), & x > b_i \end{cases},\ i = 1,2,\ldots,N \quad (7)$$

where

$$\begin{aligned} p_i &= \min\left\{ \mathop{\arg\min}_{x > \max \mathrm{Ker}\ \Omega_i} \mu_i(x) \right\} \\ q_i &= \max\left\{ \mathop{\arg\min}_{x < \min \mathrm{Ker}\ \Omega_i} \mu_i(x) \right\} \end{aligned} \quad \ldots\ldots\ldots \quad (8)$$

It is evident that Eq. (6) can be transformed into an equivalent form as follows

$$\begin{cases} \max & Q(p_i,a_i,b_i,q_i) \\ \text{s.t.} & p_i \le a_i \le b_i \le q_i \end{cases},\ i = 1,2,\ldots,N \quad .\ . \quad (9)$$

Example 1: If the L-R fuzzy number used to represent the information granule is a triangular fuzzy number, the i-th optimal information granule can be solved by using the following the following formula

$$\begin{cases} \max\ Cov(\Omega_i)\left(1 - \frac{1}{2} \cdot \frac{q_i - p_i}{\max_j \{s_{ij}\} - \min_j \{s_{ij}\}}\right) \\ \text{s.t.} \quad p_i \le a_i \le b_i \le q_i \\ \qquad\quad a_i = b_i \end{cases} \quad (10)$$

Example 2: If the L-R fuzzy number used to represent the information granule is a trapezoidal fuzzy number, the i-th optimal information granule can be solved by using the following the following formula

$$\begin{cases} \max\ Cov(\Omega_i)\left(1 - \frac{1}{2} \cdot \frac{b_i - a_i + q_i - p_i}{\max_j \{s_{ij}\} - \min_j \{s_{ij}\}}\right) \\ \text{s.t.} \quad p_i \le a_i \le b_i \le q_i \end{cases} \quad (11)$$

3.2. Granular-based Portfolio Selection Model

Since the introduction of the mean variance framework, variance has been widely accepted as a risk measure in portfolio selection analysis. As a measure of fuzzy events, Liu and Liu [9] have defined a self dual credibility measure, and given the expected value and variance of fuzzy numbers on this basis, which made it possible to conduct portfolio analysis in fuzzy environment. With the aid of the credibility measure, Huang has quantified the return and risk of portfolio based on the expected value and variance of fuzzy numbers, and proposed a fuzzy mean-variance portfolio selection model [10]. Here, in order to match the information granule as the input of portfolio selection model, we take Huang's model as the target model and extend it to a granular-based fuzzy mean-variance portfolio selection model.

According to the previous analysis we can know that for a specific investment market, if there are N investment objects waiting to be invested, the time series data corresponding to each investment object can be granulated into an information granule by using the proposed method. Let Ω_i stand for the information granule of the i-th investment object, and x_i denote the investment proportion of the investment object i. Then the fuzzy expected value can be used to measure the investment return and the fuzzy variance can be employed to estimate the investment risk.

Aggressive investors have strong risk tolerance. This type of investors prefer to pursue higher returns, and can bear the volatility risk of asset prices. In other words, aggressive investors want to be able to pursue higher returns while minimizing risk. Therefore, the aggressive portfolio selection model can be described as follows

$$\begin{cases} \min\ V[x_1\Omega_1 + x_2\Omega_2 + \cdots + x_m\Omega_m] \\ \text{s.t.} \quad E[x_1\Omega_1 + x_2\Omega_2 + \cdots + x_m\Omega_m] \ge \beta \\ \qquad\quad x_1 + x_2 + \cdots + x_m = 1 \\ \qquad\quad x_i \ge 0,\ i = 1,2,\ \cdots,m \end{cases} \quad . \ (12)$$

where β denotes the predetermined minimum return level.

Different from the aggressive investors, conservative investors pay more attention to risk than to return, and hope to obtain stable return under lower risk. Conservative investors are typical risk averse. They pay attention to obtain relatively certain return on investment, but they do not pursue high returns, and can not tolerate large fluctuations of assets in the short term. That is, conservative investors take reducing or avoiding risk as the first important meaning to guide investment behavior. Consequently, the conservative portfolio selection model can be formulated as follows

$$\begin{cases} \max\ E[x_1\Omega_1 + x_2\Omega_2 + \cdots + x_m\Omega_m] \\ \text{s.t.}\quad V[x_1\Omega_1 + x_2\Omega_2 + \cdots + x_m\Omega_m] \le \gamma \\ \qquad x_1 + x_2 + \cdots + x_m = 1 \\ \qquad x_i \ge 0,\ i = 1, 2,\ \cdots, m \end{cases} . \quad (13)$$

where γ denotes the maximum risk level that the conservative investors can tolerate.

4. Optimization Process Based on Genetic Algorithm

Through the analysis in Section 3, it can be seen that both the granulation process and the portfolio decision-making process need to be solved with the aid of optimization algorithm. Genetic algorithm (GA) is a stochastic, population-based evolution strategy optimization algorithm, which can be used as a general search algorithm.

4.1. Genetic Algorithm for Granulation

In general, the performance index function $Q : \mathbf{X} \subset \mathbb{R}^n \to \mathbb{R}$ in Eq. (9) may be a nonlinear, nondifferential or multimodal objective function. The goal is to find an optimal $\mathbf{x}_o = (p_i, a_i, b_i, q_i) \in \mathbf{X}$ such that $\mathbf{x}_o = \arg\max Q(\mathbf{x})$. Taking Eq. (9) as an example, a GA for granulation is designed as follows:

Step 1: In the optimization process, take $\mathbf{C} = (p_i, a_i, b_i, q_i)$ as a chromosome directly, where the values of four genes p_i, a_i, b_i, and q_i are restricted to the interval $\left(\min_j \{s_{ij}\}, \max_j \{s_{ij}\}\right)$.

Randomly generate four numbers x_1, x_2, x_3, x_4 and then sort them from small to large such that $x_{1'} \le x_{2'} \le x_{3'} \le x_{4'}$. Next, $\mathbf{C} = (x_{1'}, x_{2'}, x_{3'}, x_{4'})$ can be selected as an initial chromosome. After having repeated the above process *popsize* times, a set of initial chromosomes $\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_{popsize}$ can be obtained.

Step 2: Calculate the performance index Q of each chromosome, then, rank these chromosomes in accordance with the value of Q such that $Q(\mathbf{C}_1) \ge Q(\mathbf{C}_2) \ge \ldots \ge Q(\mathbf{C}_{popsize})$. We give a parameter $\tau \in (0,1)$ and construct the *rank-based evaluation function* as follows:

$$Eval(\mathbf{C}_i) = \tau(1-\tau)^{i-1},\ i = 1, 2,\ \cdots, popsize,$$

Step 3: Here, we use the *roulette wheel* method to select the high quality chromosomes. In order to make the chromosomes with high quality produce more offspring, we design the following procedure to complete the selection operation:

1) Compute the reproduction probability such that

$$P_k = \sum_{i=0}^{k} Eval(\mathbf{C}_i),\ k = 1, 2,\ \cdots, popsize,$$

where $P_0 = 0$.

2) Generate a real number r from $(0, P_{popsize}]$ randomly.

3) Select the chromosome $\mathbf{C}_i$ as an offspring when $P_{i-1} < r \le P_i$, $i = 1, 2,\ \cdots, popsize$.

4) repeat step 2) and step 3) *popsize* times until produce a complete population.

Step 4: Carry out crossover operation according to the following formula

$$\mathbf{C}_i' = r_c * \mathbf{C}_i + (1 - r_c) * \mathbf{C}_j,$$

$$\mathbf{C}_j' = (1 - r_c) * \mathbf{C}_i + r_c * \mathbf{C}_j,$$

where r_c is a random number.

Step 5: Carry out mutation operation according to the following formula

$$\mathbf{C}' = \mathbf{C} * \mathbf{P}(1 \pm \delta),$$

where $\mathbf{P}(1 \pm \delta)$ is the elementary matrix and $\delta \in (0, 0.1)$.

Step 6: Repeat above steps *generation* times.

Step 7: Select the best chromosome as a solution.

4.2. Genetic Algorithm for Portfolio Selection

The process of GA for portfolio selection is similar to that for granulation. Taking the aggressive portfolio selection model (12) as an example, a GA for portfolio selection is designed as follows:

Step 1: For a specific solution $\mathbf{x} = (x_1, x_2, \ldots, x_N)$, it can be represented by a chromosome $\mathbf{C} = (c_1, c_2, \ldots, c_N)$ such that

$$x_i = \frac{c_i}{c_1 + c_2 + \cdots + c_N},\ i = 1, 2,\ \cdots, N.$$

Randomly generate a chromosome $\mathbf{C}$ satisfying $c_i \in (0,1)$ for all $i = 1, 2,\ \cdots, N$. Then calculate the corresponding fuzzy expected value E. If $E \ge \beta$, this chromosome $\mathbf{C}$ can be selected as an initial chromosome. After having repeated the above process *popsize* times, a set of initial chromosomes $\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_{popsize}$ can be obtained.

Step 2: Calculate the corresponding fuzzy variance V of each chromosome, then, rank these chromosomes in accordance with the value of V such that $V(\mathbf{C}_1) \le V(\mathbf{C}_2) \le \ldots \le V(\mathbf{C}_{popsize})$. We give a parameter $\tau \in (0,1)$ and construct the *rank-based evaluation function* as follows:

$$Eval(\mathbf{C}_i) = \tau(1-\tau)^{i-1},\ i = 1, 2,\ \cdots, popsize,$$

Step 3: The selection operation is the same as the GA for granulation.

Step 4: The crossover operation is the same as the GA for granulation.

Step 5: Carry out mutation operation according to the following formula

$$\mathbf{C}' = \mathbf{C} * \mathbf{P}(i, j),$$

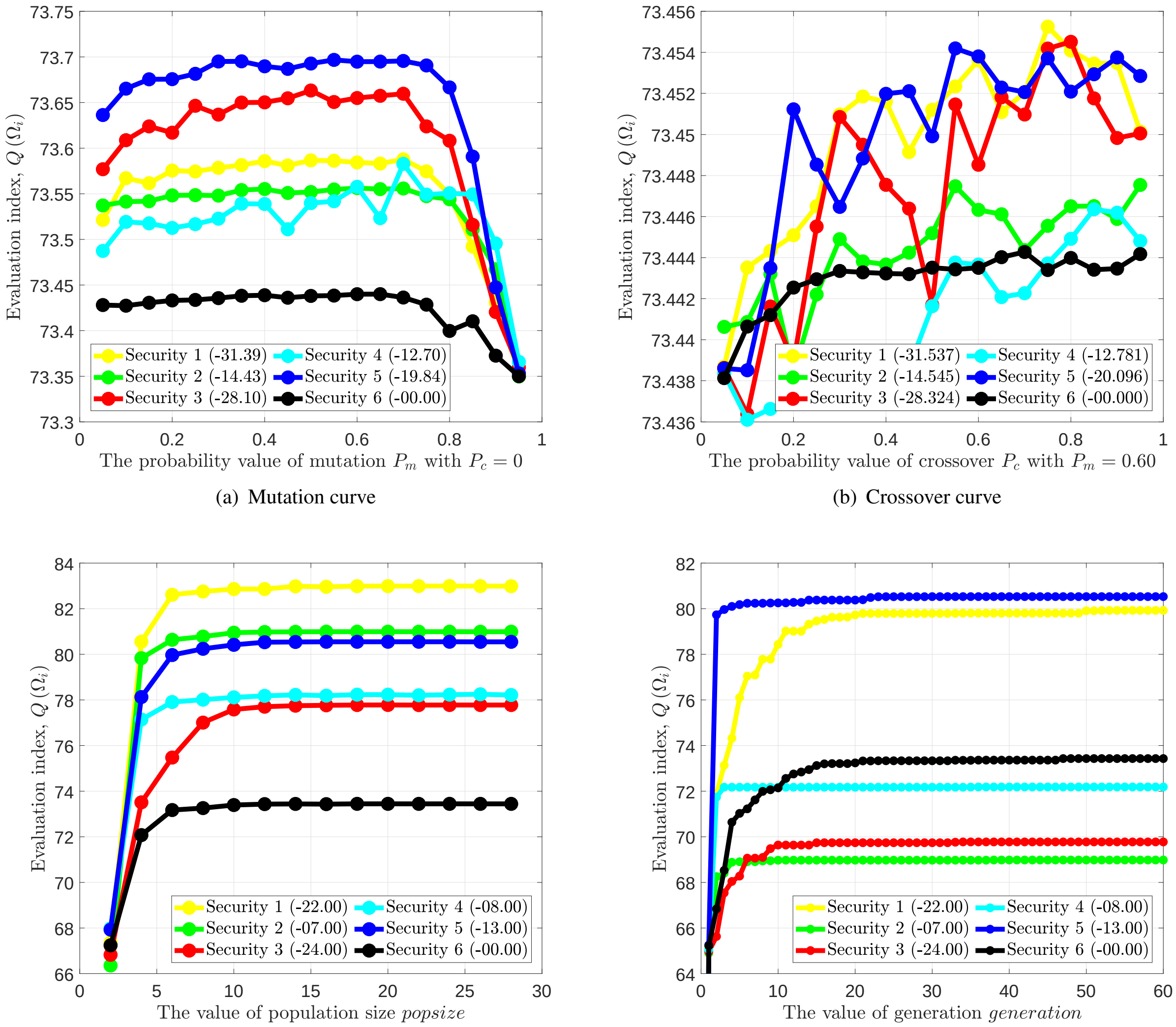
where $\mathbf{P}(i, j)$ is the elementary matrix.

Step 6: Repeat above steps *generation* times.

Step 7: Select the best chromosome as a solution.

5. Numerical Experiment

In this section, a real world data set from China Shanghai Stock Exchange is employed to demonstrate the effec-



(a) Mutation curve

(b) Crossover curve

(c) Population size curve

(d) Generation curve

Fig. 1. Relationship between objective value and mutation probability (P_m), crossover probability (P_c), population size (*popsize*), or generation (*generation*): (a) Mutation curve; (b) Crossover curve; (c) Population size curve; and (d) Generation curve.

Table 1. Information granules represented by triangular fuzzy numbers.

No.	Security code	Fuzzy return	No.	Security code	Fuzzy return
1	600168	$(-2.4365, 0.0000, 3.0454)$	**4**	600448	$(-3.8216, 0.6644, 3.1370)$
2	600201	$(-2.1304, 0.6201, 3.2735)$	**5**	600525	$(-4.0357, 0.3397, 3.5740)$
3	600368	$(-3.1571, 0.8805, 3.0316)$	**6**	600606	$(-1.8037, 0.0086, 2.2893)$

tiveness of the proposed approach.

When genetic algorithm (GA) is used to construct information granules, the mutation probability (P_m), crossover probability (P_c), population size (*popsize*), and generation (*generation*) should satisfy the inequalities such that $0.5 \leq P_m \leq 0.7$, $0.7 \leq P_c \leq 0.9$, $popsize \geq 20$, and $generation \geq 50$ according to Fig. 1. Therefore, we set the parameters of GA as follows, $P_m = 0.60$, $P_c = 0.80$, $popsize = 30$, and $generation = 150$. After having run the GA with the above parameters, the information granules of each investment object represented by fuzzy sets are shown in Table 1. For aggressive investors, the portfolio selection results under different return levels by using the fuzzy statistical and proposed granulation method are shown in Table. 2. And for conservative investors, the detailed portfolio selection results are shown in Table. 3. It can be seen from Table. 2 and Table. 3 that no matter which model is used, conservative investors tend to invest

Table 2. Comparisons of optimal investment (%) between fuzzy statistical and granulation method under different return level β.

No.	Fuzzy Statistical Method							Granulation Method						
	0.18	0.22	0.26	0.30	0.34	0.38	0.42	0.18	0.22	0.26	0.30	0.34	0.38	0.42
1	07.43	31.09	04.54	00.28	10.05	15.74	04.23	05.71	14.16	12.64	05.52	13.87	08.10	09.40
2	17.69	31.02	18.48	38.72	36.74	42.63	50.13	17.36	20.99	29.55	34.87	34.12	37.06	40.94
3	06.34	12.18	36.43	28.31	14.97	28.05	22.50	07.55	22.34	07.45	07.86	20.94	28.52	35.08
4	32.46	04.91	22.98	19.84	03.00	01.45	13.95	21.90	05.93	15.00	15.03	19.77	16.63	09.45
5	14.27	20.54	04.09	12.19	05.91	04.52	03.64	22.84	25.94	08.15	16.18	02.70	08.44	00.30
6	21.80	00.26	13.48	00.65	29.33	07.61	05.55	24.64	10.65	27.22	20.55	08.59	01.24	04.84
$D(\mathbf{x})$	**0.95**	**1.71**	**1.50**	**2.35**	**1.83**	**2.53**	**3.22**	**0.66**	**0.59**	**0.91**	**1.10**	**1.20**	**1.87**	**2.88**

Table 3. Comparisons of optimal investment (%) between fuzzy statistical and granulation method under different risk level γ.

No.	Fuzzy Statistical Method							Granulation Method						
	2.40	2.70	3.00	3.30	3.60	3.90	4.20	2.40	2.70	3.00	3.30	3.60	3.90	4.20
1	03.81	31.21	30.83	28.33	15.67	29.66	24.24	14.68	26.62	29.19	25.56	14.16	23.15	17.81
2	31.08	18.37	28.73	22.25	36.66	05.14	23.97	26.52	22.34	06.09	18.15	24.46	14.25	23.55
3	11.75	00.53	03.10	18.92	19.86	26.92	19.08	13.05	22.82	11.06	15.80	21.36	08.44	22.38
4	03.13	01.09	09.36	07.97	02.94	17.86	22.02	10.41	01.75	09.26	22.50	16.04	20.13	16.87
5	11.79	11.21	10.15	08.41	19.67	06.19	06.58	03.93	05.17	24.08	07.81	15.95	15.67	11.05
6	38.44	37.59	17.83	14.12	05.20	14.23	04.12	31.40	21.29	20.31	10.18	08.02	18.36	08.34
$D(\mathbf{x})$	**2.16**	**2.37**	**1.25**	**0.65**	**1.48**	**1.05**	**0.81**	**1.07**	**1.09**	**0.85**	**0.47**	**0.33**	**0.26**	**0.36**

in more reliable targets with low profits, while conservative investors tend to invest their money in the products with high-risk and high-reward. However, by comparing the variance of portfolio results given by both the fuzzy statistical and proposed granulation method, it can be seen that whether in the process of aggressive or conservative investment, the proposed method always lead more distributed results. Therefore, compared with fuzzy statistical, the proposed method can disperse risks more effectively by forming information granules with stable structure and good semantics through time series, which indicate that it is more in line with the investment principle that "never put all your eggs in one basket".

6. Conclusions

In this paper, a granulation method for time series data is proposed, and two portfolio selection models driven by information granules are given with the aid of the proposed granulation method. Instead of using the fuzzy variables obtained by fuzzy statistical method as the input of a portfolio model, since information granules have stable structure and good semantics, the portfolio selection model driven by information granules can achieve more distributed investments. In addition, through comparative experiment using a real world data set, the correctness and feasibility of the proposed method is also confirmed. The method proposed in this paper has strong generality and can be extended to the portfolio selection framework which employ other numerical characteristics (such as entropy, higher-order moments) as a risk measure.

References:

[1] Markowitz and M. Harry, "How to Represent Mark-to-Market Possibilities with the General Portfolio Selection Model.," J. Portf. Manag., 2013.

[2] T. Hasuike, H. Katagiri, and H. Ishii, "Portfolio selection problems with random fuzzy variable returns," Fuzzy Sets Syst., vol. 160, no. 18, pp. 2579–2596, 2007.

[3] R. Coppi, M. A. Gil, and H. A. L. Kiers, "The fuzzy approach to statistical analysis," Comput. Stat. Data Anal., vol. 51, no. 1, pp. 1–14, 2016.

[4] R. Viertl, Statistical methods for fuzzy data. Wiley, 2011.

[5] X. Zhu, W. Pedrycz, and Z. Li, "Granular Data Description: Designing Ellipsoidal Information Granules," IEEE Trans. Cybern., vol. 47, no. 12, pp. 4475–4484, 2017.

[6] W. Pedrycz and W. Homenda, "Building the fundamentals of granular computing: A principle of justifiable granularity," Appl. Soft Comput., vol. 13, no. 10, pp. 4209–4218, 2013.

[7] W. Pedrycz and X. Wang, "Designing Fuzzy Sets With the Use of the Parametric Principle of Justifiable Granularity," IEEE Trans. Fuzzy Syst., vol. 24, no. 2, pp. 489–496, 2016.

[8] J. Zhou, X. Li, and W. Pedrycz, "Mean-Semi-Entropy Models of Fuzzy Portfolio Selection," IEEE Trans. Fuzzy Syst., pp. 1–1, 2016.

[9] B. Liu and Y. K. Liu, "Expected value of fuzzy variable and fuzzy expected value models," IEEE Trans. Fuzzy Syst., vol. 10, no. 4, pp. 445–450, 2002.

[10] X. Huang, "Portfolio selection with fuzzy returns," J. Intell. Fuzzy Syst., vol. 18, no. 4, pp. 383–390, 2007.