Iranian Journal of Fuzzy Systems
Volume 23, Number 3, (2026), pp. 95-117
Original paper

A stochastic EMD-Choquet integral approach for multi-attribute decision-making

J. Zhou [1], Z. Gong [2], X. Xu [3], X. Luo [4] and G. Wei [5]

[1,2,3,4] *Research Center of Risk Management and Emergency Decision Making, School of Management Science and Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China*
[5] *Department of Mathematics and Computer Science, University of North Carolina at Pembroke, One University Drive, Pembroke, NC, 28372, USA*

202312630007@nuist.edu.cn, zwgong26@163.com, xiaoxia_xu1991@163.com, 1953517976@qq.com, guo.wei@uncp.edu

Abstract

Extracting meaningful information from high-volatility data and uncovering multi-scale knowledge from stable-state data remain key challenges in complex multi-attribute decision-making (MADM) problems. To address these challenges, a novel methodology that integrates stochastic empirical mode decomposition (EMD) with the Choquet integral is proposed. The resulting three-stage framework first decomposes the original data into trend terms, reflecting objective laws, and deviation terms, capturing subjective cognition. These components are then aggregated using Choquet integrals with Shapley values to explicitly model interactions among attributes. Finally, the framework is extended to accommodate four decision scenarios involving known or unknown attribute sets and complete or incomplete attribute values, with regularization introduced to mitigate potential bias. Case studies in investment decision-making demonstrate the effectiveness of the proposed method in integrating objective trends with subjective deviations, highlighting its advantages in multi-attribute information fusion and adaptability to complex decision environments.

Keywords: Multi-attribute decision-making (MADM), empirical mode decomposition (EMD), Choquet integral, data mining.

1 Introduction

Multi-attribute decision-making (MADM) synthesizes multiple attributes to rank or classify alternatives, supporting optimal choices in complex scenarios [23, 26]. According to data characteristics, MADM problems can be categorized into high-volatility scenarios or stable-state scenarios. Solving by traditional methods, the former usually faces difficulties in decoupling motifs, separating noise or quantifying risks [16, 45]; whilst the latter struggles with strong similarity of solution sets, redundancy of attribute representations, or balanced weight distributions [32]. Extant studies [24, 41] indicate that applying multi-scale data mining to MADM problems can improve information utilization in the above scenarios and reveal valuable latent patterns, ultimately enhancing decision accuracy.Existing multi-scale data mining strategies usually follow one of three paths: (1) Attribute-based mining, which employs methods such as the Choquet integral to fully exploit implicit inter-attribute interactions; (2) Preference-based mining, which utilizes both direct and indirect preferences through preference learning methodologies to derive decision-makers' (DMs') authentic utility values; and (3) Feature-based mining, which applies data decomposition techniques to separate a fluctuation series into trend and deviation terms, extracting intrinsic data structures and evolutionary patterns. Overall, these strategies focus on extracting inter-data relationships and elucidating the internal structure of the data. However, prior mining techniques for MADM problems are frequently restricted to a single perspective, tending either to overemphasize inter-data relationships or to concentrate exclusively on the data's intrinsic features. To bridge this gap, this paper proposes a

Corresponding Author: Z. Gong
Received: September 2025; Revised: February 2026; Accepted: May 2026.
https://doi.org/10.22111/ijfs.2026.53137.9406

new data mining and fusion strategy, which employs empirical mode decomposition (EMD) to separate decision data into a trend term (reflecting objective laws) and a deviation term (capturing subjective cognition), thereby extracting their respective latent information. Subsequently, a preference learning model based on the Choquet integral is developed to analyze and fuse the decomposed components, improving decision accuracy.

1.1 Feature extraction via EMD

Within the category of mining approaches grounded in the connotative features of data, the EMD method performs decomposition entirely based on the inherent features of the data, without the need to predefine a basis function, thereby offering high flexibility and adaptability [29]. Pioneered by Huang et al.[22], the EMD method breaks complex signals into intrinsic mode functions (IMFs), representing distinct time-scale characteristics. Initially applied to nonlinear systems [21, 22], it now spans multiple fields [36], such as in the field of the MADM. For example, Zhou et al.[50] first applied it in this context, and subsequent studies integrated it with methods like normal cloud theory and Monte Carlo simulation [28, 46]. Wang et al.[43] also employed the complete ensemble EMD method to extract residual trend components, combining them with the UTA model to analyze online hotel review data. Currently, the research of EMD combined with MADM solely focuses on the trend terms, neglecting the deviation terms, which carry essential information entropy of the high volatility data, and thus their utilization directly impacting the reliability of the decision-making models. Taking high volatility investment decision-making as an example, it is necessary to probe both their objective law trends and abnormal behavior insights. Only through leveraging expert experience to identify the mental account effect and mine high volatility characteristics, can we achieve precise decisions that are both practical and personalized [25, 42].

1.2 Preference learning via the Choquet integral

To effectively aggregate and learn from the decomposed trend and deviation terms, a method capable of modeling complex attribute interactions is essential. Traditional MADM approaches typically aggregate multiple attributes using domain knowledge, experience, and subjective preferences. Key aggregation methods involve integration operators [20], weighted sum model (WSM) [15], simple multi-attribute rating technique (SMART) [13], gray correlation analysis method [31], VIšekriterijumska Optimizacija I Kompromisno Resenje (VIKOR) [35], or outranking methods such as Elimination et Choice Translating Reality (ELECTRE) [37] and Preference Ranking Organization Method for Enrichment Evaluations (PROMETHEE) [9]. All these mainstream methods assume attribute independence, although decision attributes actually exhibit varying degrees and types of interdependencies, such as synergy (i.e., cooperation) or redundancy (i.e., competition) [17]. Fuzzy measures and the Choquet integral address this limitation by jointly considering attribute weights and interactions [40]. The Choquet integral is particularly suited for preference contexts where DMs exhibit non-additive valuation patterns (e.g., sensitivity to specific attribute combinations), which linear models cannot capture. Recent theoretical advances further support its use, such as capacities satisfying the buoyancy property [8] and characterizations of delay-averse Choquet integrals [3], reinforcing its role in learning preferences from indirect information.

In parallel, accurately capturing preference information among alternatives and attributes remains a key research challenge. Preference learning emerged to address this by modeling decision data to infer interrelationships among various attributes [11], initially applied in goal programming [39]. In decision-making, DMs typically use an aggregation model to reflect their value system based on given preference information [5], which can be either direct or indirect. The former means that DMs explicitly provide parameter values (e.g., weights, thresholds), which is often difficult to articulate directly. In contrast, the latter such as the comparison of alternatives [14], the categories to which they belong [6] or the importance of attributes [4], along with their interactions [47], is easier to provide and less cognitively demanding for DMs [19]. Thus, the use of indirect preference information, combined with an aggregation model capable of capturing interaction effects, such as the Choquet integral, can reproduce attribute interdependencies more accurately and perform attribute learning more effectively.

Meanwhile, when employing the Choquet integral model to solve for inter-attribute relationship indices, one must consider the large number of parameters involved, which makes estimation extremely difficult. At this point, the class of k-additive measures (notably Sugeno-type fuzzy measures) provides a structured and interpretable way to handle this challenge, especially when dealing with large-scale preference data. In this paper, we adopt the widely used 2-additive fuzzy measure as the underlying capacity for the Choquet integral. This choice offers a balanced compromise between modeling flexibility and computational tractability: it captures all pairwise interactions among attributes while avoiding the exponential complexity associated with higher-order capacities. Recent studies have reinforced the practical relevance of these choices; for example, 2-additive capacities have been shown to effectively model compliance and intertemporal trade-offs in dynamic decision contexts [2], while Sugeno-type capacities offer a natural extension of exponential and hyperbolic discounting within fuzzy integrals [1]. These developments motivate our adoption of a 2-additive Choquet integral as the core aggregation tool for learning preferences from indirect comparisons.

Nevertheless, practitioners frequently encounter high computational complexity, model overfitting, or parameter polarization, when solving attribute interactions via Choquet integral. To address these issues, regularization, commonly employing L_1 and L_2 norms, is widely used to minimize output weights, proving especially effective for small-sample, high-dimensional datasets [51]. Thus, this paper will integrate both L_1 and L_2 regularization methods to construct an elastic net regularization term, which will be incorporated as an additional penalty term in the objective function, so as to address issues such as parameter extremization in the newly developed preference learning model.

1.3 Research contents and structure

This paper presents a stochastic EMD-Choquet integral method for MADM that integrates multi-scale feature analysis with attribute aggregation to capture DMs' true preferences, decompose high-volatility data, and analyze the deep structure of stable-state data. The main contributions are:

(1) Stochastic EMD method based on data feature mining: We develop a data enhancement technique that integrates random sampling with the EMD method, overcoming the limitations of traditional single-dimensional data mining. Through stochastic non-replacement sampling, multiple attribute sets are generated, and their trend and deviation terms are decomposed via the EMD. Such process enables data volume expansion, tacit knowledge discovery, and precise extraction of deep features, thereby providing a more robust data foundation for subsequent analysis.

(2) Interactive preference aggregation model based on Choquet integral: A nonlinear aggregation framework considering attribute correlation and indirect preference learning is used to model the complex dependencies among attributes. Namely, the trend and deviation terms are separately analyzed using Choquet integrals. Later, fuzzy measures (represented by weights and interactions) are iteratively learnt to enhance the model's adaptability and precision, by incorporating indirect preferences such as pairwise comparisons of alternatives or attributes.

(3) MADM model based on multi-scenario adaptation: A four-class adaptive learning mechanism is introduced to address incomplete and heterogeneous data, enabling flexible decision support. By classifying scenarios according to the known/unknown status of aggregated attribute sets or the completeness of attribute values, four distinct MADM models are built, achieving expected loss minimization, difference maximization, and attribute parameter optimization under extreme values, thereby comprehensively meeting complex decision requirements.

The rest is organized as follows: Section 2 presents preliminaries of this paper, including the decomposition method based on stochastic EMD and the aggregation method based on Choquet integral. Section 3 constructs novel stochastic EMD-Choquet integral models for MADM problems. Section 4 introduces a case study and Section 5 performs some comparative analysis. Finally, concluding remarks and several promising research directions are given in Section 6.

2 Preliminaries

Assume there exists a set of alternatives in a MADM problem, denoted as $X = \{x_1, x_2, \ldots, x_m\}$, where x_i is the $i-th$ $(i \in M)$ alternative; a set of attributes $G = \{g_1, g_2, \ldots, g_n\}$, where g_j means the $j-th$ $(j \in N)$ attribute. Let $g_j(x_i)$ denote the evaluation value of alternative x_i under attribute g_j. Without loss of generality, we assume that higher attribute values, the better performance. The decision score for each alternative is computed using the aggregation function $f(\cdot)$, yielding the set of aggregated scores $Y = \{y_1, y_2, \ldots, y_m\}$, where $y_i = f(g_1(x_i), \ldots, g_n(x_i))$ represents the aggregated value of the $i-th$ $(i \in M)$ alternative.

2.1 Decomposition method based on stochastic EMD

As a classical time-series analysis tool, the EMD is widely used in data forecasting. Its key advantage lies in effectively capturing objective trends and subjective deviations in data, thereby revealing intrinsic data structures more clearly and improving forecast accuracy and reliability. In MADM problems, implicit feature mining of attribute evaluations decomposes an alternative's performance on each attribute into objective trend and cognitive deviation terms, providing DMs with granular insights to optimize their decisions. This paper introduces the EMD technique into MADM for attribute feature mining. To address fundamental differences between decision alternatives and time series, we propose a stochastic multi-attribute EMD method that incorporates sequential relationships among alternatives. We begin by briefly summarizing the conventional EMD procedure for time series.

2.1.1 Traditional EMD methods for time series

The basic idea of EMD for time series [22] is to decompose a non-smooth signal into a series of IMFs by iterative means. Each IMF must satisfy two conditions: (1) The number of extrema (local maxima and minima) is equal or differs by at most one across the entire signal; and (2) The mean value of the envelope formed by local maxima and minima is zero at any point.

The steps of the traditional EMD method are as follows:

Step 1: Let $x(t)$ denote the original signal, where t represents time points. Identify all local maxima and minima of $x(t)$, then construct upper and lower envelopes by cubic spline interpolation of these extrema.

Step 2: Compute the mean envelope $m_1(t)$ as the average of the upper and lower envelopes.

Step 3: Subtract the mean envelope $m_1(t)$ from the original information sequence $x(t)$ to obtain the new sequence $h_1(t)$, and determine whether $h_1(t)$ satisfies the conditions of IMF as

$$h_1(t) - x(t) = m_1(t). \tag{1}$$

Step 4: If $h_1(t)$ fails to meet the IMF conditions, repeat Steps 1-3 by treating $h_1(t)$ as the original signal until the conditions of IMF are met. Assume that $h_k(t)$ is obtained after k times to meet the two conditions of IMF, then $h_1(t)$ is the first intrinsic modal sequence of the original signal $x(t)$, denoted as $c_1(t)$, and the decomposition is completed once, i.e.,

$$c_1(t) = h_k(t). \tag{2}$$

$$r_1(t) = x(t) - c_1(t). \tag{3}$$

Step 5: Define $r_1(t)$ as the residual trend sequence after one decomposition cycle (Steps 1-4). Repeat Steps 1-4 with $r_1(t)$ as the original signal. After n iterations, if the residual $r_n(t)$ meets termination criteria (e.g., becomes monotonic or satisfies energy thresholds), all IMFs are extracted. The original signal can then be expressed as

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) = C(t) + r_n(t). \tag{4}$$

2.1.2 Stochastic multi-attribute EMD methods

In time series forecasting, residual trends from decomposed signals typically exhibit stable characteristics, while IMFs capture fluctuations. Similarly, for MADM problems, we decompose attribute evaluations of alternatives to obtain: trend terms (reflecting objective laws) and deviation terms (capturing subjective cognition), thus enabling efficient processing of high-volatility data and deep mining of stable-state data.

Unlike time series, MADM alternatives lack temporal sequences. To minimize the impacts of ordering uncertainty, we generate scheme position sets via stochastic non-replacement sampling. Accordingly, we propose a stochastic multi-attribute EMD method that employs the decomposed residual trends and IMFs to derive two categories of feature information characterizing the attribute's internal structure: trend information and deviation information. The flow of the stochastic multi-attribute EMD method is illustrated as Figure 1.

Step 1: Based on the alternative set $X = \{x_1, x_2, \ldots, x_m\}$, assign indices from 1 to m to generate a position set $\{1, 2, \ldots, m\}$. The evaluation vector for the $j-th$ attribute is defined as $\{g_j(x_i)|i \in M,\ j \in N\}$.

Step 2: Perform s iterations of non-replacement random sampling on the alternative set X. For each iteration s, we generate a resampled set $X^{(s)} = \{x^{(s)}_{\sigma(1)}, x^{(s)}_{\sigma(2)}, \ldots, x^{(s)}_{\sigma(m)}\}$, where $\sigma(1), \ldots, \sigma(m)$ is a random permutation of the position indices $\{1, 2, \ldots, m\}$, and the attribute evaluation set $\{g_j(x^{(s)}_{\sigma(i)})|i \in M,\ j \in N\}$.

Step 3: Taking the $j-th$ attribute set $\{g_j(x^{(s)}_{\sigma(i)})|i \in M,\ j \in N\}$ as the input of the original signal, employ the traditional EMD method in Section 2.1.1 to decompose the sequence, and obtain the two types of features for the attribute g_j from the $s-th$ random sampling, i.e., the trend term set $\{C^{(s)}_j(x_i)\}$ and the deviation term set $\{r^{(s)}_j(x_i)\}$, respectively.

Step 4: Calculate the mean values of the attribute feature obtained from j times of random sampling decomposition, so as to derive the attribute feature sets representing different scenarios, i.e., the trend term set $\{C_j(x_i)\}$ and the deviation term set $\{r_j(x_i)\}$, with $C_j(x_i) = (1/s)\sum\limits_s C^{(s)}_j(x_i)$ and $r_j(x_i) = (1/s)\sum\limits_s r^{(s)}_j(x_i)$.

Here, s represents the number of random sampling iterations, with $1 \leq s \leq m!$ (where m is the number of alternatives). As the number of iterations increases, the ranking order of the options gradually stabilizes. Based on subsequent specific case analyses, this paper sets the number of iterations to $s = 5000$, which can meet the decision-making requirements while also taking computational efficiency into account [27, 28, 46].

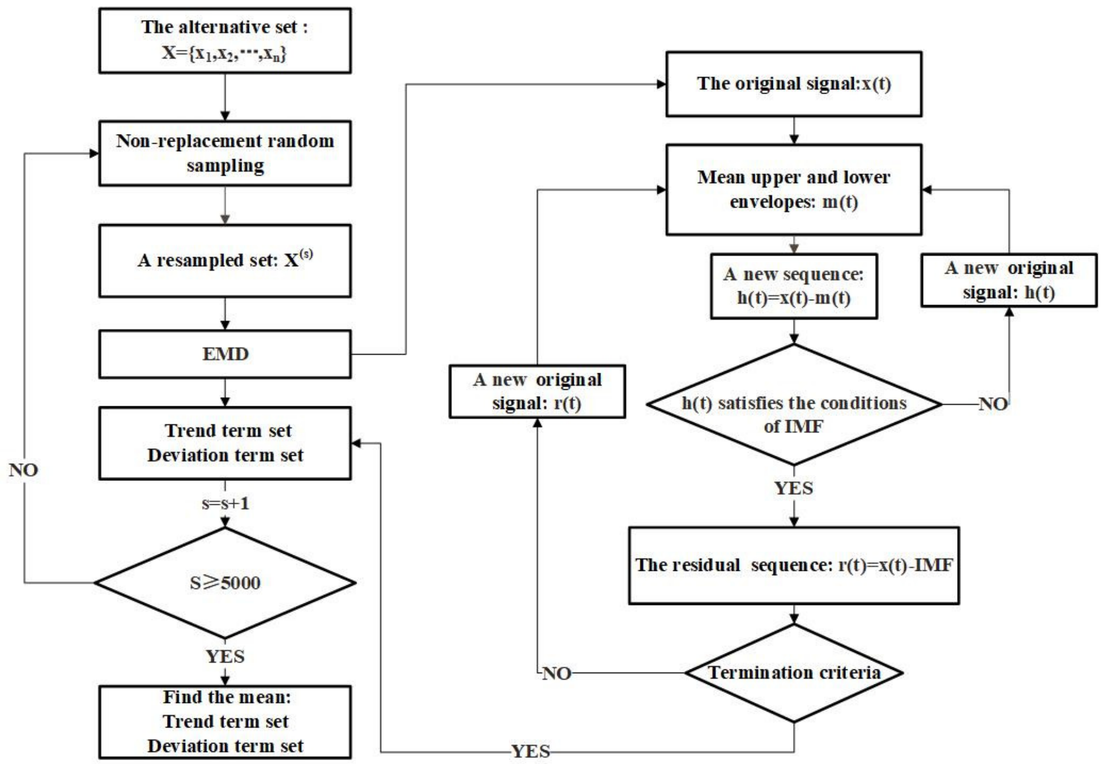


Figure 1: Flow of the stochastic multi-attribute EMD method

2.2 Aggregation method based on Choquet integral

The Choquet integral is a prominent attribute aggregation method for handling non-additive interactions, whose advantage lies in fully capturing associative relationships among attributes [17, 33]. This section describes a preference learning method using indirect pairwise comparisons to derive attribute weights and interactions, and the Choquet integral for aggregation, to achieve more accurate rankings and decisions.

2.2.1 Choquet integral aggregation method

Definition 2.1. *A capacity (fuzzy measure) on N is a real-valued set function $\mu : 2^N \to [0,1]$ defined on the power set 2^N of N (set of all sub-sets of N), satisfying boundedness (i.e., normalization condition) and monotonicity (i.e., monotonicity condition) below:*

$$E_1^C \begin{cases} (1a) \textit{normalization condition}: \ \mu(\varnothing) = 0, \ \mu(N) = 1 \\ (1b) \textit{monotonicity condition}: \ B \subseteq C \subseteq N, \ \mu(B) \leq \mu(C) \end{cases}$$

The Möbius transformation of μ is a signed-value set function $m : 2^N \to \mathbb{R}$ defined by Grabisch[18]:

$$m(C) = \sum_{B \subseteq C} (-1)^{|C \setminus B|} \mu(B). \tag{5}$$

Accordingly, the inverse of the Möbius transform is

$$\mu(C) = \sum_{B \subseteq C} m(B), \tag{6}$$

satisfying the property [12] as

(1) $m(\varnothing) = 0, \ \sum\limits_{B \subseteq N} m(B) = 1.$

(2) $\forall C \subseteq N, \ \sum\limits_{B \subseteq C} m(B) \geq 0.$

The Choquet integral of a set of alternatives $X = \{x_1, x_2, \dots, x_m\}$ with respect to the capacity μ is defined as

$$C_\mu(x_i) = \sum_{j=1}^{n} (g_{\sigma(j)}(x_i) - g_{\sigma(j-1)}(x_i))\mu(B_{\sigma(j)}), \tag{7}$$

or equivalently as

$$C_\mu(x_i) = \sum_{j=1}^{n} g_{\sigma(j)}(x_i)(\mu(B_{\sigma(j)}) - \mu(B_{\sigma(j+1)})), \tag{8}$$

where the subscript $(\cdot)$ denotes an ordering on N such that $0 = g_{\sigma(0)} \leq g_{\sigma(1)} \leq g_{\sigma(2)} \leq \dots \leq g_{\sigma(n)}$, $B_{\sigma(j)} = \{\sigma(j), \dots, \sigma(n)\}$, and $B_{\sigma(n+1)} = \varnothing$.

Based on the Möbius transformation, the discrete Choquet integral [18] is evaluated as

$$C_\mu(x_i) = \sum_{C \subseteq N} m(C) \min_{j \in C} \, g_j(x_i). \tag{9}$$

Decision problems using Choquet integral aggregating n attributes require a consideration of $2^n - 2$ parameters (except for $\mu(B) = \mu(\varnothing) = 0,\ \mu(B) = \mu(N) = 1$), which are difficult to determine in reality. Therefore, to reduce the complexity of parameter computation, scholars often use k-additive measures [18]. In this paper, we adopt the most concise from the 2-additive fuzzy measure, as described below.

$$\begin{aligned} &\mu(\{j\}) = m(\{j\}),\ j \in N, \\ &\mu(\{j,k\}) = m(\{j\}) + m(\{k\}) + m(\{j,k\}),\ \{j,k\} \subseteq N, \\ &\mu(B) = \sum_{j \in B} m(\{j\}) + \sum_{\{j,k\} \subseteq B} m(\{j,k\}),\ B \subseteq N,\ |B| \geq 2. \end{aligned} \tag{10}$$

Where the following conditions are satisfied as

$$E_2^C \begin{cases} (2a)\text{normalization condition: } m(\varnothing) = 0,\ \sum_{j \in N} m(\{j\}) + \sum_{\{j,k\} \subseteq N} m(\{j,k\}) = 1 \\ (2b)\text{monotonicity condition} \begin{cases} m(\{j\}) \geq 0,\ j \in N \\ m(\{j\}) + \sum_{k \in B} m(\{j,k\}) \geq 0,\ j \in N,\ B \subseteq N \backslash \{j\},\ B \neq \varnothing \end{cases} \end{cases}$$

Under the 2-additive fuzzy measure, the Choquet integral can be converted to the representation with Möbius transform as

$$C_\mu(x_i) = \sum_{j \in N} m(\{j\}) g_j(x_i) + \sum_{\{j,k\} \subseteq N} m(\{j,k\}) \min \{g_j(x_i), g_k(x_i)\}. \tag{11}$$

In MADM problems, the importance of each attribute is affected by both its intrinsic characteristics and its interactions with other attributes within the dataset. The Shapley value, a game-theoretic concept, is commonly applied to quantify attribute importance and interaction effects [34, 38]. Specifically,, the Shapley importance index φ_j for an attribute $j \in N$ is given by

$$\varphi_j = m(\{j\}) + \sum_{k \in N \backslash \{j\}} \frac{m(\{j,k\})}{2}. \tag{12}$$

While the Shapley interaction index $\varphi_{j,k}$ for any pair of attributes $j, k \in N$ simplifies to

$$\varphi_{j,k} = m(\{j,k\}). \tag{13}$$

2.2.2 Preference aggregation based on Choquet integral

In MADM, DMs' preferences can be divided into direct and indirect types, where the latter reduces cognitive load, as they impose lower demands on DMs. This study integrates pairwise comparisons among alternatives or attributes into the Choquet integral aggregation framework. The resulting preference-embedded Choquet model reveals non-additive interactions among attributes, enhancing decision robustness.

For alternatives $x_i, x_l \in X$ and attributes $g_j, g_k \in G$, the indirect preference information derived from the Choquet integral is mathematically quantified as follows:

· DMs believe that alternative x_i is preferred to $x_l : x_i \succ x_l \Leftrightarrow C_\mu(x_i) > C_\mu(x_l)$.
· DMs believe that alternative x_i is indifferent to $x_l : x_i \sim x_l \Leftrightarrow C_\mu(x_i) = C_\mu(x_l)$.
· DMs believe that attribute g_j is more important than $g_k : g_j \succ g_k \Leftrightarrow \varphi_j > \varphi_k$.
· DMs believe that attribute g_j is equally important as $g_k : g_j \sim g_k \Leftrightarrow \varphi_j = \varphi_k$.
· DMs believe there is a synergistic interaction between g_j and $g_k : \varphi_{j,k} > 0$.
· DMs believe there is a redundant interaction between g_j and $g_k : \varphi_{j,k} < 0$.

Then, the above indirect preference information is represented by the set E^{DM} as

$$E^{DM}\begin{cases} C_\mu(x_i) \geq C_\mu(x_l) + \varepsilon, \text{ if } x_i \succ x_l, \ \ x_i, x_l \in X \\ C_\mu(x_i) = C_\mu(x_l), \text{ if } x_i \sim x_l, \ x_i, x_l \in X \\ \varphi_j \geq \varphi_k + \varepsilon, \text{ if } g_j \succ g_k, \ j,k \in N \\ \varphi_j = \varphi_k, \text{ if } g_j \sim g_k, \ j,k \in N \\ \varphi_{j,k} \geq \varepsilon, \text{ if } \varphi_{j,k} > 0, \ j,k \in N \\ \varphi_{j,k} \leq -\varepsilon, \text{ if } \varphi_{j,k} < 0, \ j,k \in N \end{cases} \tag{14}$$

where ε is an auxiliary variable that transforms a strict inequality into a weak inequality [33].

This paper employs a 2-additive fuzzy measure-based difference-maximizing model [33] which systematically infers DMs' implicit preferences and calculates interactive weights among attributes to comprehensively characterize their inherent relationships. The resulting Choquet integral aggregation model incorporating indirect preferences is formulated as

$$\begin{aligned} &\max \ \varepsilon \\ &\text{s.t.} \begin{cases} E^{DM} \\ E^{C} \end{cases} \end{aligned} \tag{15}$$

where E^{DM} represents the indirect preference information (i.e., Model (14)), and E^C is the basic constraints that need to be satisfied introduced in Section 2.2.1. All the constraints are adapted based on fuzzy measure selection criteria and the Möbius transformation under the subsets E_1^C and E_2^C.

3 Construction of the novel MADM models

The MADM datasets can be categorized into four scenarios: (1) Aggregated value set known, the attribute value set complete; (2) Aggregated value set known, the attribute value set incomplete; (3) Aggregated value set unknown, the attribute value set complete; and (4) Aggregated value set unknown, the attribute value set incomplete. Note that the aggregated value set $Y = \{y_1, y_2, \ldots, y_m\}$ denotes the set of attribute evaluations obtained via specific aggregation methods for each alternative, while the attribute value set represents collections of attribute evaluations reflecting inter-attribute relationships under each alternative.

Given the above four scenarios, this section constructs corresponding MADM models to fully explore attribute information and improve decision accuracy and reliability. The modeling framework operates as follows:

(1) Attribute decomposition: The stochastic multi-attribute EMD method decomposes attribute data to generate two new feature sets, i.e., the trend term set and the deviation term set.

(2) Preference modeling: Using indirect preferences originated from pairwise comparisons over both feature sets, a Choquet integral aggregation model is built to minimize expected loss and maximize decision divergence, by learning attribute weight and inter-attribute interaction parameters, while conducting in-depth analysis of interactions among alternatives and attributes.

(3) Attribute aggregation: Given the above derived parameters, both feature sets are aggregated with Choquet integrals and indirect preferences, yielding the trend feature and the deviation feature aggregation sets.

(4) Decision ranking: A linear weighted approach is to re-aggregate both attribute sets, generating the final decision scheme and determining the optimal rankings.

Due to incomplete attribute information in MADM problems, the targeted model faces fewer constraints and risks of attribute parameter polarization. To mitigate such effects, this section introduces regularization terms [51], where elastic net regularization combines L_1 and L_2 regularization properties to prevent parameter polarization and enable feature selection. Therefore, in the preference learning model constructed in this paper, it is incorporated into the objective function as a penalty term to constrain the values of fuzzy measures. The elastic net regularization term [44] is defined as

$$L_{1,2} = \lambda_1 \sum_{j=1}^{n} |w_j| + \lambda_2 \sum_{j=1}^{n} w_j^2, \tag{16}$$

where $\lambda_1 \sum_{j=1}^{n} |w_j|$ is the L_1 regularization term, $\lambda_2 \sum_{j=1}^{n} w_j^2$ is the L_2 regularization term, λ_1 and λ_2 are the regularization parameters, and w_j is the attribute weights.

In this paper, w_j denotes the fuzzy measure $m(\{j\})$ or $m(\{j,k\})$, $j,k \in N$, $j \neq k$. At the same time, according to the actual situation of the case in this article, λ_1 and λ_2 are determined using a grid search combined with 5-fold cross-validation. Define $\lambda_1, \lambda_2 \in \{0.001,\ 0.01,\ 0.1,\ 1,\ 10\}$, and then traverse all parameter combinations. For each set (λ_1, λ_2), the average performance of the model is evaluated using 5-fold cross-validation, and the parameter combination that performs best on the validation set is ultimately chosen [7, 30, 48, 49].

3.1 MADM models with known aggregated value sets

When the aggregated value set of a decision dataset is known, the attribute evaluation values for each alternative and their aggregated values can be presented in tabular form, shown as Table 1. Aiming to minimize the error between the aggregated value of alternative x_i (i.e., $f(g_1(x_i), \ldots, g_n(x_i))$) and its true value y_i, we define the expected loss function as $\eta = \sum_{i=1}^{m} \left(f(g_1(x_i), \ldots, g_n(x_i)) - y_i\right)^2$. In this paper, the Choquet integral is employed as the aggregation function (i.e., $f(g_1(x_i), \ldots, g_n(x_i)) = C_\mu(x_i)$) in all models, accounting for the interrelationships between attributes. Its calculation is given by Formula (11). From this, we derive the aggregated values for all alternatives and define their collection as the constraint set E^{AR}.

$$E^{AR}:\ C_\mu(x_i) = \sum_{j \in N} m(\{j\}) g_j(x_i) + \sum_{\{j,k\} \subseteq N} m(\{j,k\}) \min\{g_j(x_i), g_k(x_i)\},\ i \in M. \tag{17}$$

Therefore, for cases with known aggregated value sets, we establish an expected loss minimization MADM optimization model, that pursues Pareto optimality in decision accuracy within deterministic environments, formulated as follows

$$\begin{aligned} &\min\ \eta \\ &\text{s.t.} \begin{cases} E^{AR} \\ E_2^C \end{cases} \end{aligned} \tag{18}$$

where E^{AR} denotes the constraint set derived from the original decision information, and E_2^C signifies the fundamental normalization and monotonicity conditions.

Table 1: Decision dataset with known aggregated value sets

Program(x_i)	g_1	$\cdots$	g_n	y_i
1	$g_1(x_1)$	$\cdots$	$g_n(x_1)$	y_1
2	$g_1(x_2)$	$\cdots$	$g_n(x_2)$	y_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
m	$g_1(x_m)$	$\cdots$	$g_n(x_m)$	y_m

The Choquet integral-based aggregation method reveals that different alternative rankings support distinct values of interactive attribute fuzzy measures. For a system with n interactive attributes, the alternative ranking contains $n!$ distinct records, meaning that there exist $n!$ permutations of fuzzy measures for n attributes. Only when all $n!$ attribute orderings exist can each fuzzy measure be data-supported to obtain accurate values. Such scenario is referred to as complete attribute value information; otherwise, it is incomplete. Consequently, when the aggregated value set is known, MADM datasets can be further categorized into complete and incomplete attribute value classifications.

When the aggregated value set is known and attribute value is complete, the constructed expected loss minimization

MADM optimization model is formulated as follows.

$$
\begin{aligned}
&\min\ \eta^h \\
&s.t. \begin{cases}
\eta^h = \sum_{i=1}^{m} \left(C_\mu^h(x_i) - y_i^h\right)^2 \\
E^{AR}:\ C_\mu^h(x_i) = \sum_{j\in N} m^h(\{j\}) g_j^h(x_i) + \sum_{\{j,k\}\subseteq N} m^h(\{j,k\}) \min\{g_j^h(x_i), g_k^h(x_i)\},\ i \in M \\
E_2^C \begin{cases}
m^h(\varnothing) = 0,\ \sum_{j\in N} m^h(\{j\}) + \sum_{\{j,k\}\subseteq N} m^h(\{j,k\}) = 1 \\
m^h(\{j\}) + \sum_{k\in B} m^h(\{j,k\}) \geq 0, m^h(\{j\}) \geq 0,\ j \in N,\ B \subseteq N\backslash\{j\}
\end{cases}
\end{cases}
\end{aligned}
\tag{19}
$$

The model aims to determine the optimal fuzzy measure by minimizing the error. The constraints consist of two parts: the set E^{AR}, which defines the alternative aggregation rules based on the 2-additive fuzzy measure and the Choquet integral, and the set E_2^C, which specifies the normalization and monotonicity conditions on the fuzzy measure. Here, the superscript $h = \{1, 2\}$ respectively represents trend features and deviation features, and $\{g_j^1(x_i)\} = \{C_j(x_i)\}$ denotes the objective trend feature set, while $\{g_j^2(x_i)\} = \{r_j(x_i)\}$ represents the cognitive deviation feature set. Unless otherwise stated, the same symbols retain these meanings in the following text.

When the aggregated value set is known but attribute value is incomplete, introducing an elastic net regularization term effectively prevents parameter polarization. The extreme-case expected loss minimization MADM optimization model is formulated below.

$$
\begin{aligned}
&\min\ \eta^h + L_{1,2} \\
&s.t. \begin{cases}
\eta^h + L_{1,2} = \sum_{i=1}^{m} \left(C_\mu^h(x_i) - y_i^h\right)^2 + \lambda_1 \sum_{j=1}^{n} \left|w_j^h\right| + \lambda_2 \sum_{j=1}^{n} \left(w_j^h\right)^2 \\
E^{AR}:\ C_\mu^h(x_i) = \sum_{j\in N} m^h(\{j\}) g_j^h(x_i) + \sum_{\{j,k\}\subseteq N} m^h(\{j,k\}) \min\{g_j^h(x_i), g_k^h(x_i)\},\ i \in M \\
E_2^C \begin{cases}
m^h(\varnothing) = 0,\ \sum_{j\in N} m^h(\{j\}) + \sum_{\{j,k\}\subseteq N} m^h(\{j,k\}) = 1 \\
m^h(\{j\}) + \sum_{k\in B} m^h(\{j,k\}) \geq 0, m^h(\{j\}) \geq 0,\ j \in N,\ B \subseteq N\backslash\{j\}
\end{cases}
\end{cases}
\end{aligned}
\tag{20}
$$

This model introduces an elastic net regularization term on the basis of minimizing the error to determine the optimal fuzzy measure. Here, the weight w_j^h corresponds to either the single-attribute weight $m^h(\{j\})$ or the pairwise interaction weight $m^h(\{j,k\})$ $(j, k \in N,\ j \neq k)$ in the fuzzy measure m^h. The meanings of all other variables and parameters in the model remain consistent with those defined in Model (19).

3.2 MADM models with unknown aggregated value sets

When the aggregated value set is unknown, decision information can be characterized by an attribute evaluation matrix, shown as Table 2. The optimization models previously constructed are no longer applicable. However, indirect preference information among alternatives and attributes can be readily provided, thus, given the known pairwise preferences, the attribute evaluations per alternative are aggregated with the Choquet integral. As a result, it leads to a difference-maximizing MADM preference learning model that achieves robustness under uncertainty via the maximum entropy criterion, formulated as follows

$$
\begin{aligned}
&\max\ \varepsilon \\
&\text{s.t.} \begin{cases}
E^{AR} \\
E^{DM} \\
E_2^C
\end{cases}
\end{aligned}
\tag{21}
$$

where E^{AR} denotes the dataset constraint from the original information, E^{DM} represents the indirect preference over alternatives and attributes, and E_2^C signifies the normalization and monotonicity conditions.

In fact, preference relationships among alternatives and attributes may not be fully known due to various factors, with only partial preference information available. Therefore, if the aggregated value set is unknown, we assess the

Table 2: Decision dataset with unknown aggregated value sets

Program(x_i)	g_1	g_2	$\cdots$	g_n
1	$g_1(x_1)$	$g_2(x_1)$	$\cdots$	$g_n(x_1)$
2	$g_1(x_2)$	$g_2(x_2)$	$\cdots$	$g_n(x_m)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
m	$g_1(x_m)$	$g_2(x_m)$	$\cdots$	$g_n(x_m)$

completeness of attribute values based on indirect preferences, enabling the stratification of decision datasets into complete and incomplete categories.

When the aggregated value set is unknown but attribute value is complete, pairwise preference relationships among alternatives (or attributes) can be fully determined. By ranking alternatives in descending order, we build the difference-maximizing MADM preference learning model as follows.

$$\max\ \varepsilon^h$$
$$s.t.\begin{cases} E^{AR}:\ C_\mu^h(x_i)=\sum\limits_{j\in N} m^h(\{j\})g_j^h(x_i)+\sum\limits_{\{j,k\}\subseteq N} m^h(\{j,k\})\min\{g_j^h(x_i),g_k^h(x_i)\},\ i\in M \\ E^{DM}\begin{cases} C_\mu^h(x_i)-C_\mu^h(x_{i+1})\geq\delta+\varepsilon^h,\ \text{if } x_i\succ x_{i+1},\ i\in M,\ i\neq m \\ -\delta\leq C_\mu^h(x_i)-C_\mu^h(x_{i+1})\leq\delta,\ \text{if } x_i\sim x_{i+1},\ i\in M,\ i\neq m \\ \varphi_j^h-\varphi_k^h\geq\delta+\varepsilon^h,\ \text{if } j\succ k,\ j,k\in N \\ -\delta\leq\varphi_j^h-\varphi_k^h\leq\delta,\ \text{if } j\sim k,\ j,k\in N \end{cases} \\ E_2^C\begin{cases} m^h(\varnothing)=0,\ \sum\limits_{j\in N} m^h(\{j\})+\sum\limits_{\{j,k\}\subseteq N} m^h(\{j,k\})=1 \\ m^h(\{j\})+\sum\limits_{k\in B} m^h(\{j,k\})\geq 0, m^h(\{j\})\geq 0,\ j\in N,\quad B\subseteq N\backslash\{j\} \end{cases} \end{cases} \tag{22}$$

This model aims to determine a fuzzy measure that accurately reflects the DM's cognitive structure while remaining consistent with their preference information by maximizing preference consistency. In addition to the alternative aggregation rules E^{AR} and the fuzzy measure constraints E_2^C, the constraints also include the set E^{DM}, which transforms the DM's indirect preference information regarding alternatives and attributes into mathematical constraints. Furthermore, δ is a small positive number. By adjusting δ, we can modulate the sensitivity of partial rankings, thereby better reflecting the underlying preferences or priority levels in the decision-making process [33].

When the aggregated value set is unknown and attribute value is incomplete, i.e., with only partial preference relationships of alternatives and attributes available, an elastic net regularization term is introduced to prevent parameter polarization. Thus, we develop an extreme-case difference-maximizing MADM preference learning model below.

$$\max\ \varepsilon^h+L_{1,2}$$
$$s.t.\begin{cases} \varepsilon^h+L_{1,2}=\varepsilon^h+\lambda_1\sum\limits_{j=1}^{n}\left|w_j^h\right|+\lambda_2\sum\limits_{j=1}^{n}(w_j^h)^2 \\ E^{AR}:\ C_\mu^h(x_i)=\sum\limits_{j\in N} m^h(\{j\})g_j^h(x_i)+\sum\limits_{\{j,k\}\subseteq N} m^h(\{j,k\})\min\{g_j^h(x_i),g_k^h(x_i)\},\ i\in M \\ E^{DM}\begin{cases} C_\mu^h(x_i)-C_\mu^h(x_l)\geq\delta+\varepsilon^h,\ \text{if } x_i\succ x_l,\ i,l\in M \\ -\delta\leq C_\mu^h(x_i)-C_\mu^h(x_l)\leq\delta,\ \text{if } x_i\sim x_l,\ i,l\in M \\ \varphi_j^h-\varphi_k^h\geq\delta+\varepsilon^h,\ \text{if } j\succ k,\ j,k\in N \\ -\delta\leq\varphi_j^h-\varphi_k^h\leq\delta,\ \text{if } j\sim k,\ j,k\in N \end{cases} \\ E_2^C\begin{cases} m^h(\varnothing)=0,\ \sum\limits_{j\in N} m^h(\{j\})+\sum\limits_{\{j,k\}\subseteq N} m^h(\{j,k\})=1 \\ m^h(\{j\})+\sum\limits_{k\in B} m^h(\{j,k\})\geq 0, m^h(\{j\})\geq 0,\ j\in N,\ B\subseteq N\backslash\{j\} \end{cases} \end{cases} \tag{23}$$

Based on maximizing preference consistency, this model introduces an elastic net regularization term to determine the optimal fuzzy measure, thereby seeking a balance between fitting the DM's preferences and ensuring the model's generalization capability. The definitions of all other parameters and variables in the model remain consistent with those in the preceding text and will not be reiterated.

Finally, based on the actual form of the data, the corresponding model (19)/ (20)/ (22)/ (23) is selected to derive the attribute weights and interaction parameters between attributes for the two feature sets. Subsequently, these parameters are aggregated to obtain the final combined set and the optimal ranking of the decision alternatives.

4 Case study

In investment environments, decision-making data presents two typical patterns: high-volatility and stable-state. High–volatility data faces three core issues: motivation decoupling, noise separation, and risk quantification. Stable-state data exhibits three salient characteristics: high scenario similarity, redundant attribute representation, and balanced weight distribution. Conventional methods often fail to effectively distinguish long-term trends from short-term fluctuations. Consequently, the method proposed in this study demonstrates unique advantages when processing such investment data. This approach employs stochastic multi-attribute EMD to decompose the data into trend and deviation terms, followed by the Choquet integral to model investor preferences and achieve final decision-making.

Consider 20 investment alternatives $X = \{x_1, x_2, \ldots, x_{20}\}$, each evaluated across three financial attributes $G = \{g_1, g_2, g_3\}$ (risk exposure, return potential, and liquidity). Let $Y = \{y_1, y_2, \ldots, y_{20}\}$ denote the aggregated evaluation values for each alternative. Based on the method proposed above, the optimal ranking of investment alternatives and the interrelationship between attributes are explored. First, the stochastic multi-attribute EMD method decomposes the decision dataset into trend terms and deviation terms, forming two attribute feature sets. Then, based on the dataset structure, appropriate aggregation models are selected for subsequent analysis (Table 3, Table 4, Figure 2 - 5). As observed from the decomposed dataset, this case falls under the known aggregated value set with incomplete attribute value information category.

Table 3: Decision datasets known for attribute aggregation sets

Program(x_i)	$g_1(x_i)$	$g_2(x_i)$	$g_3(x_i)$	y_i	Optimal ordering
1	9.4	7.8	6.2	6.888	3
2	2.2	4.1	.6	2.272	19
3	9	8.5	7.6	7.959	1
4	5.2	4.4	3.3	3.752	12
5	9.6	8.9	6.5	7.566	2
6	2.8	4.5	6.4	2.908	17
7	4.6	6.9	4.4	4.976	10
8	7.3	7.5	5.8	6.42	4
9	4.2	7.3	1.3	3.584	16
10	4.2	2.2	6.4	2.486	18
11	3.1	5.6	5.2	3.323	15
12	4.2	7.7	5.6	5.082	9
13	3.3	5.3	4.7	3.582	14
14	5.8	9.1	5.3	6.217	5
15	2.9	1.7	3.3	1.844	20
16	6.1	5.9	8.5	5.994	7
17	5.7	7.1	7.5	5.754	6
18	3.7	4.3	5	3.739	13
19	6.2	5.4	4.3	4.752	11
20	7.2	9.2	3.2	5.44	8

Based on the previous models, we can derive interaction relationships and attribute weights for both trend and deviation terms, shown as in Table 5 and Table 6.

Attribute importance and interaction indexes for the trend term attribute set:

$$\varphi_1^1 = 0.3339;\ \varphi_2^1 = 0.4119;\ \varphi_3^1 = 0.2542;\ \varphi_{1,2}^1 = 0.3455;\ \varphi_{1,3}^1 = 0.3055;\ \varphi_{2,3}^1 = 0.0807.$$

Table 4: Decision attribute sets of the trend and deviation terms

Program(x_i)	$g_1(x_i)$		$g_2(x_i)$		$g_3(x_i)$		y_i	
	$C_1(x_i)$	$r_1(x_i)$	$C_2(x_i)$	$r_2(x_i)$	$C_3(x_i)$	$r_3(x_i)$	y_i^1	y_i^2
1	8.4814	0.9186	7.4031	0.3969	5.6832	0.5168	6.5112	0.3768
2	1.9850	0.2150	3.8914	0.2086	4.2166	0.3834	2.1477	0.1243
3	8.1205	0.8795	8.0675	0.4325	6.9666	0.6334	7.5236	0.4354
4	4.6918	0.5082	4.1761	0.2239	3.0249	0.2751	3.5467	0.2053
5	8.6618	0.9382	8.4471	0.4529	5.9582	0.5418	7.1521	0.4139
6	2.5264	0.2736	4.2710	0.2290	5.8666	0.5334	2.7489	0.1591
7	4.1505	0.4495	6.5489	0.3511	4.0333	0.3667	4.7038	0.2722
8	6.5866	0.7134	7.1183	0.3817	5.3166	0.4834	6.0688	0.3512
9	3.7896	0.4104	6.9285	0.3715	1.1916	0.1084	3.3879	0.1961
10	3.7896	0.4104	2.0880	0.1120	5.8666	0.5334	2.3500	0.1360
11	2.7971	0.3029	5.3150	0.2850	4.7666	0.4334	3.1412	0.1818
12	3.7896	0.4104	7.3082	0.3918	5.1332	0.4668	4.8040	0.2780
13	2.9775	0.3225	5.0303	0.2697	4.3083	0.3917	3.3860	0.1960
14	5.2332	0.5668	8.6369	0.4631	4.8583	0.4417	5.8769	0.3401
15	2.6166	0.2834	1.6135	0.0865	3.0249	0.2751	1.7431	0.1009
16	5.5039	0.5961	5.5998	0.3002	7.7915	0.7085	5.6661	0.3279
17	5.1430	0.5570	6.7387	0.3613	6.8749	0.6251	5.4392	0.3148
18	3.3384	0.3616	4.0812	0.2188	4.5833	0.4167	3.5345	0.2045
19	5.5941	0.6059	5.1252	0.2748	3.9416	0.3584	4.4920	0.2600
20	6.4964	0.7036	8.7318	0.4682	2.9333	0.2667	5.1424	0.2976

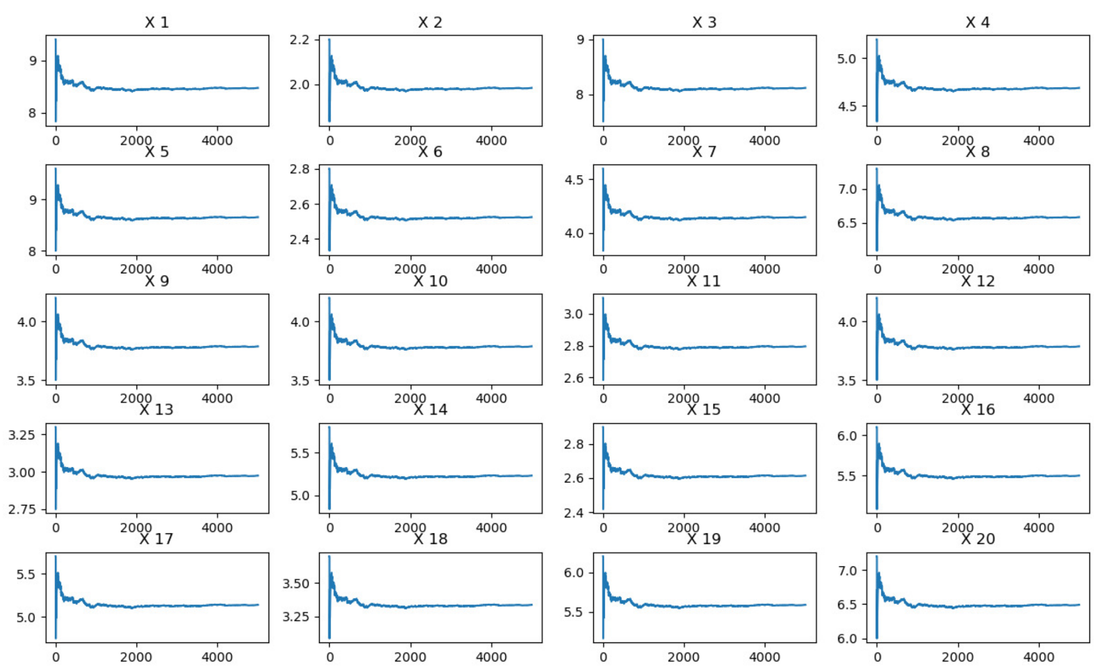


Figure 2: Attribute trend feature set $g_1^1(x_i) = C_1(x_i)$

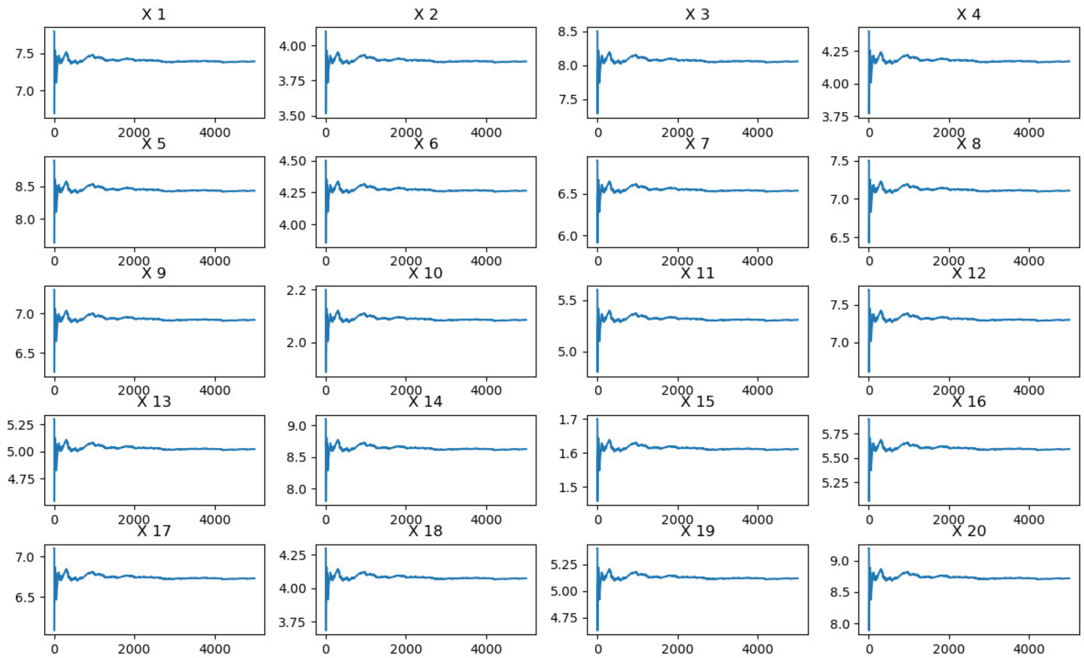


Figure 3: Attribute trend feature set $g_2^1(x_i) = C_2(x_i)$

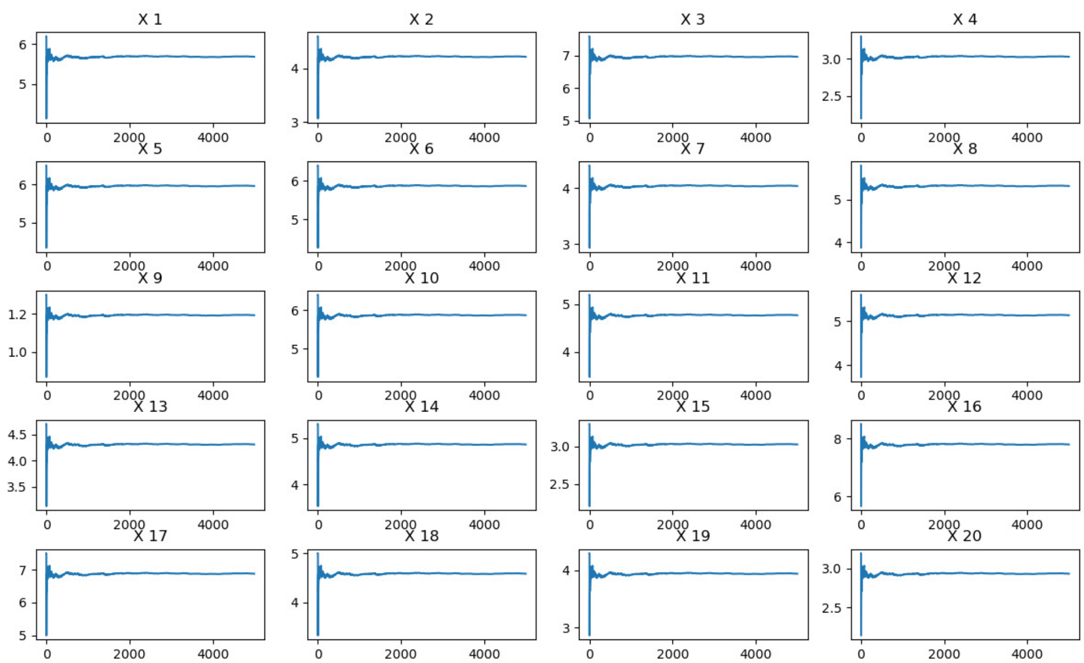


Figure 4: Attribute trend feature set $g_3^1(x_i) = C_3(x_i)$

Table 5: Attribute fuzzy measures for trend term attribute sets

Fuzzy measure	Not regularized	After regularization
$m^1(\{g_1\})$	0.0000	0.0084
$m^1(\{g_2\})$	0.2049	0.1988
$m^1(\{g_3\})$	0.0000	0.0611
$m^1(\{g_1, g_2\})$	0.3372	0.3455
$m^1(\{g_1, g_3\})$	0.4305	0.3055
$m^1(\{g_2, g_3\})$	0.0273	0.0807

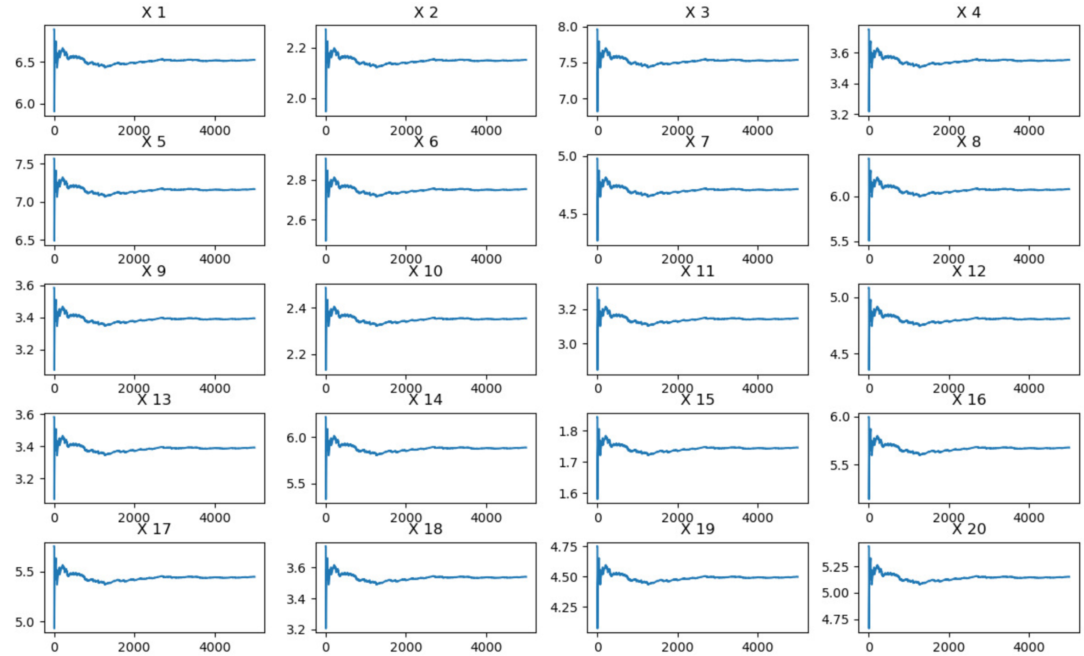


Figure 5: Attribute trend feature set y_i^1

Attribute importance and interaction indexes for the deviation term attribute set:

$$\varphi_1^2 = 0.2112;\ \varphi_2^2 = 0.503;\ \varphi_3^2 = 0.2859;\ \varphi_{1,2}^2 = 0.2445;\ \varphi_{1,3}^2 = 0.1613;\ \varphi_{2,3}^2 = 0.2727.$$

Table 6: Attribute fuzzy measures for deviation term attribute sets

Fuzzy measure	Not regularized	After regularization
$m^2(\{g_1\})$	0.0000	0.0083
$m^2(\{g_2\})$	0.1372	0.2444
$m^2(\{g_3\})$	0.0000	0.0689
$m^2(\{g_1, g_2\})$	0.1125	0.2445
$m^2(\{g_1, g_3\})$	0.0447	0.1613
$m^2(\{g_2, g_3\})$	0.7056	0.2727

The final ranking was obtained by integrating both the objective trends and investors' behavioral biases in this paper, with the third alternative identified as the optimal one. As evidenced by the data in Table 5 and Table 6, it shows successful optimization of the fuzzy measure system through regularization. Specifically, within the trend term attribute set, originally zero-weight measures $m^1(\{g_1\})$ and $m^1(\{g_3\})$ respectively increased to 0.0084 and 0.0611, while corresponding weights in the deviation term rose to 0.0083 and 0.0689, thereby completely eliminating attribute invalidation issues caused by data sparsity in the conventional methods. Concurrently, regularization effectively mitigated parameter polarization phenomena. The interaction set $m^1(\{g_1, g_3\})$ in the trend term decreased by 29% (from 0.4305 to 0.3055), while $m^2(\{g_2, g_3\})$ in the deviation term plummeted by 61% (from 0.7056 to 0.2727), substantially enhancing equilibrium in attribute weight distribution across both components.

The results of Shapley value further reveal the effectiveness of the dual-dimensional decision mechanism. Within the trend term attribute set, the core position of $\varphi_2^1 = 0.4119$ demonstrates that market risk factor g_2 serves as the primary driver of long-term trends, while the strong synergistic effect of $\varphi_{1,2}^1 = 0.3455$ validates the stable correlation between enterprise value g_1 and market risk g_2. In the deviation term attribute set, the exceptionally high weight of $\varphi_2^2 = 0.503$ (exceeding 50% of the total) quantitatively confirms investors' short-term overemphasis on risk attributes, whereas the key interaction value $\varphi_{2,3}^2 = 0.2727$ unveils the dynamic hedging mechanism through which liquidity g_3 counterbalances risk g_2.

In summary, this methodology constructs a decision-making paradigm integrating data decomposition with behavioral preferences, achieving precise separation between objective laws and subjective cognition. It thereby establishes a novel framework for complex investment decisions that concurrently embodies mathematical rigor and behavioral adaptability.

5 Comparative analysis

In order to demonstrate the theoretical and practical advantages of our new method, this section uses relevant data from m Brunelli and Corrente [10] to conduct a comparative analysis with the Choquet integral aggregation method used in their paper, while also further comparing it with several traditional multi-attribute decision-making methods (WSM, TOPSIS, VIKOR). Assume an investment project is evaluated based on four attributes (g_1 : expected return, g_2 : probability of success, g_3 : strategic impact, and g_4 : operational impact), then each project's attribute data is observed by using our stochastic multi-attribute EMD-Choquet integral aggregation method. The results in Table 7 reveal: (1) Trend term rankings exhibit significant divergence from the Choquet integral method. For example, Project 9 ranks 2nd in the Choquet integral method but drops to 9th in trend term rankings. (2) Deviation term rankings show high consistency with the Choquet integral method.

Table 7: Comparison between the Choquet integral method and this paper

Project(x_i)	Choquet	Trend term	Deviation term	Overall ranking
1	9	8	9	9
2	10	6	10	10
3	3	2	4	2
4	1	1	1	1
5	8	3	8	5
6	7	4	6	3
7	6	7	7	6
8	4	5	2	4
9	2	9	3	7
10	5	10	5	8

Further observation shown in Table 8 regarding the attribute importance indices reveals that:

(1) Significant differences exist between the trend term attribute importance indices and those in the Choquet integral method, both in ranking and numerical values. The trend term attribute g_3 surges to the top position with an importance index of 0.5223, while substantial numerical disparities emerge among attributes, validating that the EMD successfully strips away short-term noise, highlighting long-term value orientation, as evidenced by Project 6 rising from 7th to 4th place due to its g_3 advantage.

(2) While the ranking of deviant term attribute importance indices shows high consistency with the Choquet integral method, the numerical gaps between indices have widened significantly. The weight of expected returns g_1 surges to 0.4241, explaining the root cause behind Project 9 maintaining its high position (3nd) in the deviant term while plummeting to 9th place in the trend term, investors' excessive focus on short-term gains.

Table 8: Comparison of attribute importance index between the Choquet integral method and this paper

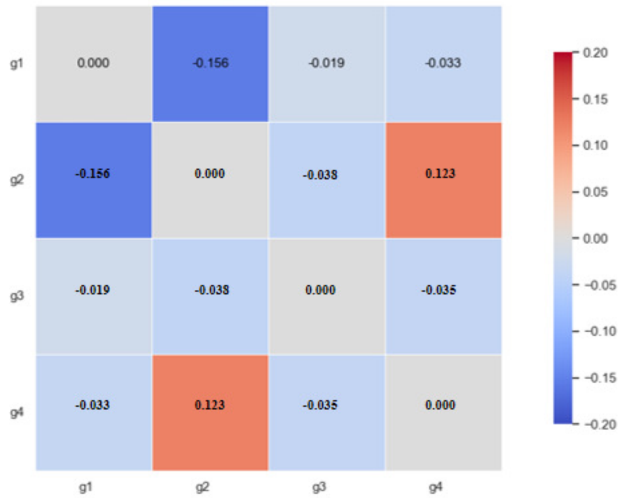
Importance index	Choquet	Trend term	Deviation term
φ_1	0.3015	0.3209	0.4241
φ_2	0.3007	0.1005	0.3448
φ_3	0.2246	0.5223	0.1187
φ_4	0.1741	0.0564	0.1118

The Results of attribute interaction indices are further exhibited, as shown in Figure 6.

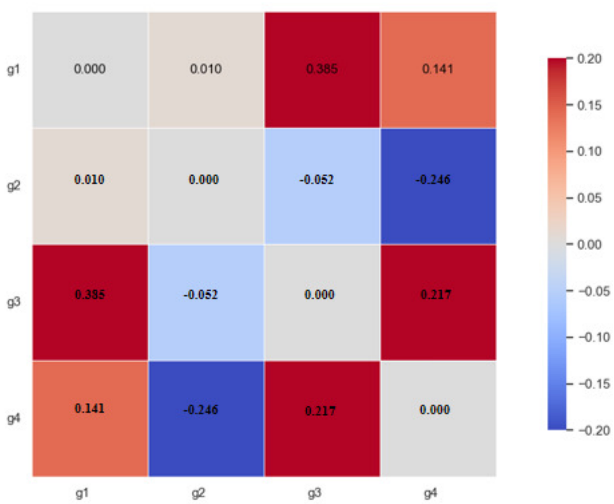
(1) The Choquet integral method potentially contains a critical misjudgment: the negative interaction (-0.156) between g_1 (expected return) and g_2 (probability of success) contradicts the fundamental principle of positive risk-return relationships in investment theory.

(2) Within the trend term, g_3 forms a core interaction hub that establishes a return-strategy transmission mechanism through strong synergy with g_1 (+0.385), while its positive interaction with g_4 (+0.217) enhances the long-term value

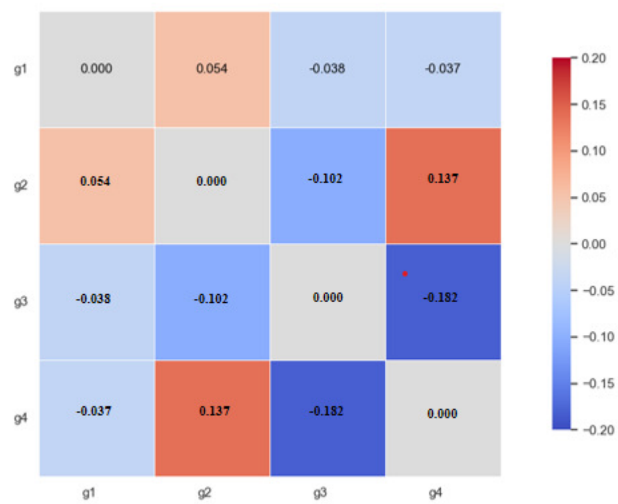
contribution of operational attributes.



(a) Attribute interaction index: Choquet



(b) Attribute interaction index: Trend term attribute



(c) Attribute interaction index: Deviant term attribute

Figure 6: Heatmap of attribute interaction index

Note: Color hue indicates interaction type as red for synergy (positive interaction) and blue for redundancy (negative interaction). Color darkness indicates the absolute strength of the interaction.

(3) Within the deviant term, heightened focus on risk hedging manifests through the strong synergy between g_2 and g_4 (+0.137), reflecting investors' tendency to bundle operational risks with success probability in assessments. Meanwhile, the negative interaction between g_1 and g_3 (-0.038), significantly lower than the trend term's +0.385, empirically evidences cognitive simplification in short-term decision-making, where strategic linkages are systematically neglected.

Below are the specific attribute evaluation values for each project, along with the decomposed trend term and deviation term values (Tables 9, 10, and 11).

To facilitate a more in-depth comparative analysis, we also employed several traditional multi-attribute decision-making methods to calculate and rank the rank the alternatives, shown as in Figure 7. The proposed EMD-based method demonstrates significant and comprehensive advantages in systematic comparison with other multi-attribute decision-making approaches (including WSM, TOPSIS, VIKOR, and the basic Choquet integral). The core advantage of the method proposed in this paper comes from the two-dimensional decomposition of investment data. Specifically, the original data is decoupled via the EMD method into trend terms, which reflects objective laws, and deviation terms, which captures subjective cognition and short-term fluctuations. These components are then aggregated separately

Table 9: Attribute evaluation values

Project(x_i)	$g_1(x_i)$	$g_2(x_i)$	$g_3(x_i)$	$g_4(x_i)$
1	0.25	0.2	0.2	0.6
2	0.15	0.2	0.1	0.8
3	0.7	0.15	0.7	0.4
4	0.9	0.05	0.9	0.9
5	0.4	0.5	0.35	0.3
6	0.4	0.6	0.3	0.4
7	0.3	0.7	0.1	0.7
8	0.25	0.8	0.2	0.6
9	0.15	0.8	0.1	0.8
10	0.05	0.9	0.05	0.7

Table 10: Attribute evaluation values-trend term

Project(x_i)	$g_1^1(x_i) = C_1(x_i)$	$g_2^1(x_i) = C_2(x_i)$	$g_3^1(x_i) = C_3(x_i)$	$g_4^1(x_i) = C_4(x_i)$
1	0.1908	0.1026	0.0683	0.2611
2	0.4054	0.1026	0.0341	0.3481
3	0.7631	0.0770	0.2390	0.1741
4	0.6200	0.0257	0.3073	0.3916
5	0.3816	0.2566	0.1195	0.1305
6	0.3339	0.3079	0.1024	0.1741
7	0.2623	0.3592	0.0341	0.3046
8	0.1908	0.4105	0.0683	0.2611
9	0.0954	0.4105	0.0341	0.3481
10	0.0238	0.4618	0.0171	0.3046

Table 11: Attribute evaluation values-deviation term

Project(x_i)	$g_1^2(x_i) = r_1(x_i)$	$g_2^2(x_i) = r_2(x_i)$	$g_3^2(x_i) = r_3(x_i)$	$g_4^2(x_i) = r_4(x_i)$
1	0.0592	0.0974	0.1317	0.3389
2	-0.2554	0.0974	0.0659	0.4519
3	-0.0631	0.0730	0.4610	0.2259
4	0.2800	0.0243	0.5927	0.5084
5	0.0184	0.2434	0.2305	0.1695
6	0.0661	0.2921	0.1976	0.2259
7	0.0377	0.3408	0.0659	0.3954
8	0.0592	0.3895	0.1317	0.3389
9	0.0546	0.3895	0.0659	0.4519
10	0.0262	0.4382	0.0329	0.3954

using the Choquet integral for learning.

In the correlation analysis of the methods (Figure 8 and 9), the underlying reason for its superiority is further revealed. The proposed method maintains a high correlation with TOPSIS, with Spearman's rank correlation coefficient of 0.988 and Kendall's tau correlation coefficient of 0.956, indicating that it effectively inherits the rationality of classical ideal point-based approximation methods. Simultaneously, it exhibits a moderate correlation with the basic Choquet integral, with Spearman's rank correlation coefficient of 0.636 and Kendall's tau correlation coefficient of 0.511, which shows that while retaining the core capability of capturing attribute interactions, the unique ranking logic is generated by the EMD preprocessing, which introduces a new and valuable dimension of information. The consistency analysis (Figure 10) reveals that the decision outcomes of the proposed method demonstrate moderate yet robust consensus with other methods, as indicated by an average Spearman's rank correlation coefficient of 0.741 and an average Kendall's tau coefficient of 0.599. Furthermore, due to the in-depth mining of data, its consistency level is significantly higher than that of VIKOR and is comparable to classical methods like WSM and TOPSIS. In contrast, the basic Choquet integral method, due to the lack of data decomposition, suffers from the defect of overemphasizing short-term fluctuations while neglecting long-term trends, whereas methods like VIKOR show greater inconsistency in their rankings.

Consequently, the method proposed in this study has successfully, to some extent, overcome the analytical limitations inherent in traditional approaches due to their reliance on undecomposed data. The trend term rectifies inflated rankings and anchors long-term development patterns by elevating the weight of strategic impact g_3 and restructuring interaction networks. The deviant term quantifies short-term behavioral biases by amplifying the weight of return attributes g_1 and capturing risk synergies. This methodology fundamentally addresses the inadequacy in decision analysis caused by undecomposed data in conventional approaches, thereby establishing a novel paradigm for complex investment decisions that integrates mathematical rigor with behavioral adaptability.

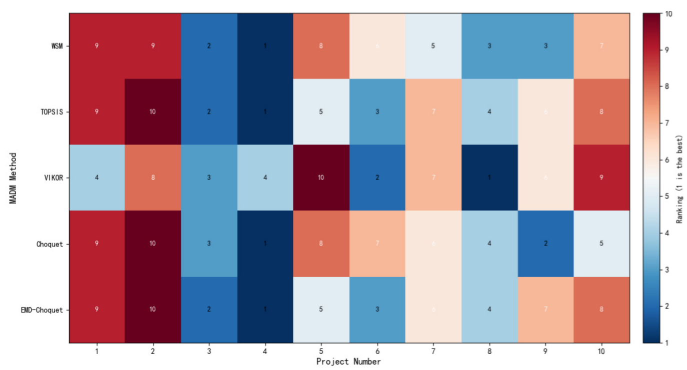


Figure 7: Heatmap comparing rankings of five MADM methods

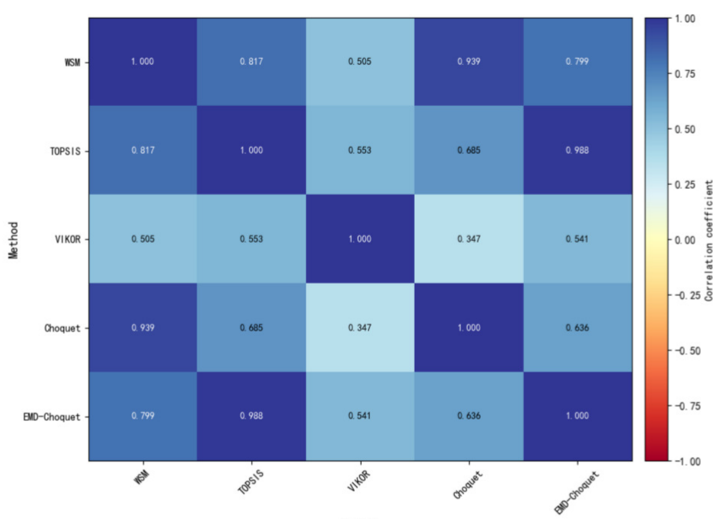


Figure 8: Spearman correlation coefficient matrix

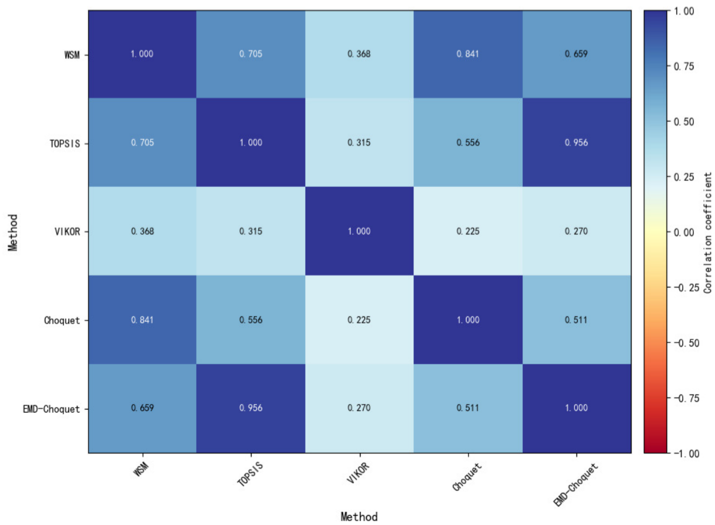


Figure 9: Kendall correlation coefficient matrix

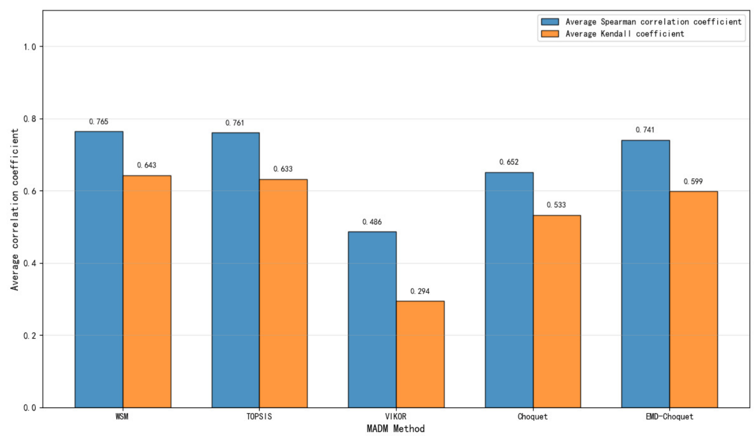


Figure 10: Consistency of each method with other methods

6 Conclusion

Addressing the critical research frontier of attribute information mining in MADM, this paper introduces a novel paradigm, the stochastic multi-attribute EMD-Choquet integral aggregation method. It is designed to capture the connotative features of high-volatility data, excavate the deep structure of stable-state data, and quantify nonlinear attribute interactions through fuzzy measures, allowing for the unified analysis of diverse data types in complex decision environments.

The primary contributions can be summarized in three key innovations: First, a stochastic multi-attribute EMD-based data feature mining method is proposed, which generates multiple attribute sets through stochastic non-repetitive sampling and decomposes them into trend and deviation terms, achieving data expansion and deep feature extraction. Second, an interactive preference aggregation model based on the 2-additive Choquet integral is constructed, which characterizes nonlinear interactions among attributes through fuzzy measure learning and enhances the model's adaptive capability by incorporating indirect preference information. Finally, a multi-scenario MADM framework is established to provide flexible decision support in complex environments. Comparative analysis results demonstrate that the proposed method achieves an average Spearman's coefficient of 0.741 and an average Kendall's coefficient of 0.599, showing broad robustness and consensus with classical methods. Furthermore, by leveraging deep data mining, it overcomes the analytical limitations of traditional methods caused by undecomposed data, thereby validating its effectiveness and practical utility.

While the proposed method offers advantages in terms of decision accuracy and interpretability, it still has certain limitations. On the one hand, as a data preprocessing technique, the EMD method relies on the assumption that the data can be represented by a finite number of IMFs. Its decomposition results are susceptible to the effects of the stopping criterion and boundary conditions, which may indirectly interfere with subsequent preference learning. On the other hand, although the model based on the 2-additive Choquet integral controls the parameter scale, it still faces a significant computational burden when handling high-dimensional attributes. Moreover, the model learning depends on indirect pairwise preference information provided by the DM. When such information is insufficient or contains substantial noise, it may compromise the robustness of parameter learning.

In response to the above limitations, future research can be further developed in the following directions. First, it is necessary to evaluate the parameter settings of EMD in a broader range of scenarios and develop adaptive methods to enhance its robustness. Second, the model theory itself can be extended by exploring more general fuzzy measures, such as the Sugeno-type measure [1], to capture higher-order complex interaction patterns that go beyond pairwise interactions. Third, at the application level, the model should be extended to complex systems such as medical resource allocation and energy system optimization, in order to build domain-adaptive dynamic decision-making frameworks. Finally, deep integration with advanced technologies such as deep learning and reinforcement learning could provide new pathways for systematically uncovering implicit knowledge in MADM.

Acknowledgement

This work was partly supported by the National Natural Science Foundation of China (Grant Nos. 72371137 and 72401141), the Major Project Plan of Philosophy and Social Sciences Research at Jiangsu University (Grant No. 2020SJZDA076), the Natural Science Foundation of Jiangsu Province (Grant No. BK20240691), the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant No. 24KJB630016), the China Postdoctoral Science Foundation (Grant No. 2024M761475), and the Startup Foundation for Introducing Talent of NUIST (Grant No. 2024r019).

References

[1] J. C. R. Alcantud, *Sugeno capacities for the extension of exponential and hyperbolic discounting to fuzzy integrals*, Punjab University Journal of Mathematics, **56**(4) (2024), 102-111. https://doi.org/10.52280/pujm.2024.56(3-4)04

[2] J. C. R. Alcantud, *The capacity compliance problem: Refinements of time discounting to Choquet integrals with 2-additive fuzzy measures*, Computational and Applied Mathematics, **44**(8) (2025), 424. https://doi.org/10.1007/s40314-025-03407-4

[3] J. C. R. Alcantud, *A characterization of delay averse Choquet integrals for intertemporal analysis*, JCR Alcantud', Theory and Decision, **99** (2025), 37-70. https://doi.org/10.1007/s11238-025-10047-x

[4] S. Angilella, S. Corrente, S. Greco, R. Słowiński, *Robust ordinal regression and stochastic multiobjective acceptability analysis in multiple criteria hierarchy process for the Choquet integral preference model*, Omega, **63** (2015), 154-169. https://doi.org/10.1016/j.omega.2015.10.010

[5] S. Angilella, S. Greco, B. Matarazzo, *Non-additive robust ordinal regression: A multiple criteria decision model based on the Choquet integral*, European Journal of Operational Research, **201**(1) (2010), 277-288. https://doi.org/10.1016/j.ejor.2009.02.023

[6] S. G. Arcidiacono, S. Corrente, S. Greco, *Robust stochastic sorting with interacting criteria hierarchically structured*, European Journal of Operational Research, **292**(2) (2021), 735-754. https://doi.org/10.1016/j.ejor.2020.11.024

[7] S. Arlot, A. Celisse, *A survey of cross-validation procedures for model selection*, Statistics Surveys, **4** (2010), 40-79. https://doi.org/10.1214/09-ss054

[8] G. Beliakov, S. James, J. Z. Wu, *Capacities satisfying the buoyancy property on average*, Theory and Decision, **99** (2025), 13-36. https://doi.org/10.1007/s11238-025-10039-x

[9] J. P. Brans, P. Vincke, B. Mareschal, *How to select and how to rank projects: The PROMETHEE method*, European Journal of Operational Research, **24**(2) (1986), 228-238. https://doi.org/10.1016/0377-2217(86)90044-5

[10] M. Brunelli, S. Corrente, *Modeling criteria and project interactions in portfolio decision analysis with the Choquet integral*, Omega, **126** (2024), 103076. https://doi.org/10.1016/j.omega.2024.103076

[11] R. Burke, *Hybrid recommender systems: Survey and experiments*, User Modeling and User-Adapted Interaction, **12**(4) (2002), 331-370. https://doi.org/10.1023/A:1021240730564

[12] A. Chateauneuf, J. Y. Jaffray, *Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion*, Mathematical Social Sciences, **17**(3) (1989), 263-283. https://doi.org/10.1016/0165-4896(89)90056-5

[13] W. Edwards, *How to use multiattribute utility measurement for social decisionmaking*, IEEE Transactions on Systems, Man, and Cybernetics, **7**(5) (1977), 326-340. https://doi.org/10.1109/TSMC.1977.4309720

[14] J. R. Figueira, S. Greco, R. Słowiński, *Building a set of additive value functions representing a reference preorder and intensities of preference: GRIP method*, European Journal of Operational Research, **195**(2) (2009), 460-486. https://doi.org/10.1016/j.ejor.2008.02.006

[15] P. C. Fishburn, *Additive utilities with incomplete product sets: Application to priorities and assignments*, Operations Research, **15**(3) (1967), 537-542. https://doi.org/10.1287/opre.15.3.537

[16] J. Fu, J. Lin, H. Hao, *Volatility analysis for the GARCH–Itô–Jumps model based on high-frequency and low-frequency financial data*, International Journal of Forecasting, **39**(4) (2023), 1698-1712. https://doi.org/10.1016/j.ijforecast.2022.08.006

[17] Z. Gong, W. Guo, R. Słowiński, *Transaction and interaction behavior-based consensus model and its application to optimal carbon emission reduction*, Omega, **104** (2021), 102491. https://doi.org/10.1016/j.omega.2021.102491

[18] M. Grabisch, *K-order additive discrete fuzzy measures and their representation*, Fuzzy Sets and Systems, **92**(2) (1997), 167-189. https://doi.org/10.1016/S0165-0114(97)00168-1

[19] S. Greco, V. Mousseau, R. Słowiński, *Robust ordinal regression for value functions handling interacting criteria*, European Journal of Operational Research, **239**(3) (2014), 711-730. https://doi.org/10.1016/j.ejor.2014.05.022

[20] J. C. Harsanyi, *Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility*, Journal of Political Economy, **63**(4) (1955), 309-321. https://doi.org/10.1086/257678

[21] N. E. Huang, Z. Shen, S. R. Long, *A new view of nonlinear water waves: The Hilbert spectrum*, Annual Review of Fluid Mechanics, **31**(1) (1999), 417-457. https://doi.org/10.1146/annurev.fluid.31.1.417

[22] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, H. H. Liu, *The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis*, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, **454** (1971), 903-995. https://doi.org/10.1098/rspa.1998.0193

[23] C. Hwang, K. Yoon, *Methods for multiple attribute decision making*, In: Multiple Attribute Decision Making. Lecture Notes in Economics and Mathematical Systems, Vol. **186**. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-48318-9_3

[24] Z. Jinghua, R. Haiying, *Multi-attribute decision-making based on data mining under a dynamic hybrid trust network*, Computers and Industrial Engineering, **185** (2023), 109672. https://doi.org/10.1016/j.cie.2023.109672

[25] B. Li, Y. Huang, *Portfolio optimization with mental accounts under uncertain random environment and butterfly optimization algorithm with adaptive strategies*, Applied Soft Computing, **161** (2024), 111720. https://doi.org/10.1016/j.asoc.2024.111720

[26] D. F. Li, J. Liu, *A parameterized nonlinear programming approach to solve matrix games with payoffs of I-fuzzy numbers*, IEEE Transactions on Fuzzy Systems, **23**(4) (2015), 885-896. https://doi.org/10.1109/TFUZZ.2014.2333065

[27] H. Li, D. Luo, J. Wang, *A stochastic EMD method of aggregating LGDM experts' information*, Statistics and Decision, **19** (2018), 46-50. https://doi.org/10.13546/j.cnki.tjyjc.2018.19.010.

[28] H. Li, D. Luo, B. Wei, *Method for large group decision-making with uncertain linguistic assessment information based on MC-EMD*, Chinese Journal of Management Science, **25**(4) (2017), 164-173. https://doi.org/10.16381/j.cnki.issn1003-207x.2017.04.020

[29] H. Li, Y. Mei, X. Hao, Z. Chen, *Out-of-sample equity premium predictability: An EMD-denoising based model*, Pacific-Basin Finance Journal, **88** (2024), 102536. https://doi.org/10.1016/j.pacfin.2024.102536

[30] Y. Liang, D. Dai, S. Jin, *On convergence of regularized covariance estimator based on modified Cholesky decomposition*, Journal of Multivariate Analysis, **212** (2026), 105553. https://doi.org/10.1016/j.jmva.2025.105553

[31] S. Liu, Y. Lin, *Introduction to grey system theory*, In: Grey Systems. Understanding Complex Systems, Vol. **68**. Springer, Berlin, Heidelberg, 2010. https://doi.org/10.1007/978-3-642-16158-2_1

[32] X. Ma, H. Liu, Y. Liu, J. Z. Zhang, *Multi-label feature selection considering label importance-weighted relevance and label-dependency redundancy*, European Journal of Operational Research, **322**(1) (2025), 215-236. https://doi.org/10.1016/j.ejor.2024.11.038

[33] J. L. Marichal, M. Roubens, *Determination of weights of interacting criteria from a reference set*, European Journal of Operational Research, **124**(3) (2000), 641-650. https://doi.org/10.1016/s0377-2217(99)00182-4

[34] T. Murofushi, S. Soneda, *Techniques for reading fuzzy measures (III): Interaction index*, in '9th Fuzzy System Symposium', Sapporo Japan, (1993), 693-696.

[35] S. Opricovic, G. H. Tzeng, *Multicriteria planning of post-earthquake sustainable reconstruction*, Computer-Aided Civil and Infrastructure Engineering, **17**(3) (2002), 211-220. https://doi.org/10.1111/1467-8667.00269

[36] A. Rojas, J. M. Górriz, J. Ramírez, I. A. Illán, F. J. Martínez-Murcia, A. Ortiz, M. M. Gómez Río, M. Moreno-Caballero, *Application of empirical mode decomposition (EMD) on DaTSCAN SPECT images to explore Parkinson disease*, Expert Systems with Applications, **40**(7) (2013), 2756-2766. https://doi.org/10.1016/j.eswa.2012.11.017

[37] B. Roy, *Classement et choix en présence de points de vue multiples*, Revue Française d'Informatique et de Recherche Opérationnelle, **2**(V1) (1968), 57-75. https://doi.org/10.1051/ro/196802v100571

[38] L. S. Shapley, *A value for n-person games*, 1952. https://scispace.com/papers/a-value-for-n-person-games-58w5vcviy3. https://doi.org/10.1017/CBO9780511528446.003

[39] J. Siskos, *Analyse de systèmes de décision multicritère en univers aléatoire*, Foundations of Control Engineering, **8**(3-4) (1983), 193-212.

[40] M. Sugeno, *Theory of fuzzy integrals and its applications*, Doctoral Thesis, Tokyo Institute of Technology, 1974.

[41] M. Tang, H. Liao, *Multi-attribute large-scale group decision making with data mining and subgroup leaders: An application to the development of the circular economy*, Technological Forecasting and Social Change, **167** (2021), 120719. https://doi.org/10.1016/j.techfore.2021.120719

[42] R. Thaler, *Toward a positive theory of consumer choice*, Journal of Economic Behavior Organization, **1**(1) (1980), 39-60. https://doi.org/10.1016/0167-2681(80)90051-7

[43] D. Wang, Z. Gong, G. Wei, M. Á. Martínez, E. Herrera-Viedma, *Using sentiment analysis and CEEMDAN to learn the preferences of consumer groups: A case study of online hotel reviews*, Applied Soft Computing, **191** (2026), 114625. https://doi.org/10.1016/j.asoc.2026.114625

[44] X. Wang, P. Wang, Y. Song, Q. Xiang, J. Li, *Recognition of high-resolution range profile sequence based on TCN with sequence length-adaptive algorithm and elastic net regularization*, Expert Systems with Applications, **248** (2024), 123417. https://doi.org/10.1016/j.eswa.2024.123417

[45] C. Wang, Y. Ye, L. Ma, D. Li, L. Zhuang, *Dual disentanglement of user–item interaction for recommendation with causal embedding*, Information Processing and Management, **60**(5) (2023), 103456. https://doi.org/10.1016/j.ipm.2023.103456

[46] B. Wei, N. Xie, *Large group decision-making method based on random simulation and filter analysis*, Control and Decision, **34**(8) (2019), 1761-1768. https://doi.org/1001-0920(2019)08-1761-08

[47] B. Zhang, Y. Dong, H. Zhang, W. Pedrycz, *Consensus mechanism with maximum-return modifications and minimum-cost feedback: A perspective of game theory*, European Journal of Operational Research, **287**(2) (2020), 546-559. https://doi.org/10.1016/j.ejor.2020.04.014

[48] X. Zhang, C. A. Liu, *Model averaging prediction by K-fold cross-validation*, Journal of Econometrics, **235**(1) (2023), 280-301. https://doi.org/10.1016/j.jeconom.2022.04.007

[49] Y. Zhao, W. Zhang, X. Liu, *Grid search with a weighted error function: Hyper-parameter optimization for financial time series forecasting*, Applied Soft Computing, **154** (2024), 111362. https://doi.org/10.1016/j.asoc.2024.111362

[50] R. Zhou, R. Chen, Y. Chen, J. Lu, S. Zhou, L. Liu, *The method of group decision making based on the EMD extracting experts linguistic assessment information*, Systems Engineering-Theory and Practice, **36**(3) (2016), 743-749. https://doi.org/10.12011/1000-6788(2016)03-0743-07

[51] K. Zhou, Z. Gong, G. Wei, R. Słowiński, *Preference disaggregation analysis with criteria selection in a regularization framework*, Omega, **133** (2025), 103252. https://doi.org/10.1016/j.omega.2024.103252