Open Journal of Statistics, 2026, 16(2), 121-173
https://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

A FCM-TODIM-STWD Method for Large-Group Multi-Attribute Decision Making with Probabilistic Linguistic Term Sets

Xiaoxia Ruan, Gaili Xu*

School of Mathematics and Statistics, Guilin University of Technology, Guilin, China
Email: *2007033@glut.edu.cn, 1790384424@qq.com

How to cite this paper: Ruan, X.X. and Xu, G.L. (2026) A FCM-TODIM-STWD Method for Large-Group Multi-Attribute Decision Making with Probabilistic Linguistic Term Sets. *Open Journal of Statistics*, **16**, 121-173.
https://doi.org/10.4236/ojs.2026.162007

Received: March 20, 2026
Accepted: April 17, 2026
Published: April 20, 2026

Abstract

Probabilistic linguistic term sets (PLTSs) can flexibly express uncertain evaluation information provided by decision makers (DMs). Large-group multi-attribute decision making (LGMADM) problems usually involve uncertainty, heterogeneity, and strong subjectivity, which makes it difficult for conventional methods to simultaneously achieve effective classification, behavioral modeling, and sequential information utilization. This study investigates LGMADM problems in the context of PLTSs and proposes a FCM-TODIM-STWD method. First, to deal with the structural heterogeneity of large-scale expert groups, FCM is introduced to cluster experts and generate cluster-level preference information. Second, a novel deviation-range entropy of PLTSs is developed to determine attribute weights objectively. Third, according to the obtained attribute weights, a sequential three-way decision framework is constructed, in which attributes are incorporated into the decision process one by one, and alternatives in the boundary region are progressively refined until all alternatives are classified. Meanwhile, TODIM is introduced to calculate dominance degrees and conditional probabilities, so that the psychological behavior of DMs can be incorporated into the decision process. Therefore, behavioral rationality and sequential discrimination can be considered in a unified framework. Finally, an air-quality case study based on the 2024 data of 20 cities is provided to demonstrate the feasibility of the proposed method. Comparative analysis, sensitivity analysis, and ablation experiments are further conducted to verify its effectiveness, interpretability, and superiority. The results show that the proposed method not only maintains high consistency with several classical methods in ranking results, but also provides additional classification information and better decision support. Hence, the proposed method provides an effective and interpretable framework for solving LGMADM problems in a PLTS environment.

Keywords

Probabilistic Linguistic Term Sets, Sequential Three-Way Decision, Large-Group Multi-Attribute Decision Making, TODIM, FCM-TODIM-STWD Method

1. Introduction

Against the backdrop of big data and crowd intelligence collaboration, real-world decision-making problems have increasingly exhibited such features as a large number of participants, multiple evaluation attributes, and uncertainty in information expression. As a result, large-scale group multi-attribute decision-making has gradually become an important topic in decision theory research. Compared with traditional small-scale group multi-attribute decision-making, large-scale group multi-attribute decision-making not only requires the integration of evaluation information provided by a large number of decision makers across multiple attributes, but also has to address practical challenges such as diverse expert sources, significant differences in cognitive patterns, and complex group structures [1] [2]. Since substantial differences in evaluations often exist among decision makers, directly aggregating all individual opinions in a unified manner may obscure the latent structural characteristics within the group, thereby hindering the effective extraction and utilization of evaluation information in subsequent stages [2] [3]. Therefore, how to properly organize large-scale expert groups in complex and uncertain environments, and on this basis achieve effective aggregation, analysis, and alternative screening of multi-attribute evaluation information, has become a key issue in current large-scale group decision-making research. Existing review studies have shown that research on large-scale group decision-making mainly focuses on expert clustering, preference representation, group structure identification, and large-scale information processing mechanisms, all of which are closely related to alternative evaluation, attribute utilization, and ranking and selection processes in multi-attribute decision-making [1]-[3].

Furthermore, in the context of large-scale group multi-attribute decision-making, as the number of participating experts increases and the complexity of evaluation tasks grows, decision information is often no longer suitable for direct representation by precise numerical values. Instead, experts tend to express their subjective judgments on alternatives with respect to different attributes in linguistic forms [2] [4]. This is mainly because large-scale expert groups are typically characterized by diverse sources, substantial differences in knowledge backgrounds, and inconsistent cognitive granularity. If all decision makers are forced to provide uniform and precise numerical evaluations, such a requirement would not only deviate from actual decision-making practices, but also easily lead to information distortion [1] [2]. To address this issue, existing studies have gradually introduced linguistic decision-making tools into large-scale group decision-making and have explored this topic from perspectives such as multi-granular linguistic information, probabilistic linguistic information, and linguistic aggregation in large-scale groups [2] [5]. Among these approaches, the probabilistic linguistic term set (PLTS), which can simultaneously represent multiple possible linguistic terms together with their associated probabilities, is more suitable than traditional single linguistic terms for characterizing experts' hesitation and uncertainty. In addition, it preserves the probability distribution characteristics of evaluation information, thereby demonstrating strong applicability in large-scale group multi-attribute decision-making [5] [6]. In recent years, scholars have conducted a series of studies on probabilistic linguistic information processing in large-scale groups, including decision-making with incomplete multi-granular probabilistic linguistic information, double-hierarchy large-scale linguistic decision-making, and dynamic probabilistic linguistic large-scale group multi-attribute decision-making. These studies indicate that employing PLTS as an information representation tool for large-scale group multi-attribute decision-making has a clear theoretical foundation and promising application prospects [2] [4] [6].

In complex and uncertain environments, three-way decision (TWD) can divide objects into the acceptance region, rejection region, and boundary region, thereby providing a theoretical basis for deferred judgment. For this reason, it has become an important tool for dealing with classification and uncertainty-related problems [7]. On this basis, sequential three-way decision (STWD) further characterizes the progressive discrimination process of objects from the perspectives of granular computing and multilevel cognition. Its basic idea is to first conduct an initial three-way classification under relatively coarse granularity or limited information conditions, and then continue the discrimination process for those objects that remain in an uncertain state by incorporating subsequent information, so as to gradually reduce the uncertain region and obtain more refined decision results [8] [9]. Relevant studies have shown that STWD does not rely on a single fixed mode of information utilization; rather, it can be integrated with multi-granularity structures, adaptive thresholds, and multi-stage information acquisition mechanisms. As a result, it demonstrates strong flexibility and adaptability in complex classification and decision-making problems [8] [10] [11]. In recent years, STWD has been gradually extended to multi-attribute group decision-making and heterogeneous information decision-making, showing promising application potential in problems such as multi-attribute evaluation, group discrimination, and large-scale group multi-attribute decision-making [12]-[14]. Meanwhile, with the continued development of linguistic decision-making and probabilistic linguistic information research, the idea of sequential three-way decision has also been introduced into hesitant fuzzy linguistic environments, probabilistic linguistic term sets, and other complex linguistic settings, providing new theoretical support for handling linguistic hesitation, evaluation uncertainty, and dynamic group judgment [15]-[18]. However, existing studies have focused more on the sequential discrimination framework itself, while there is still considerable room for further research on how to systematically integrate expert clustering, attribute weights, sequential discrimination mechanisms, and behavior-based conditional probability solving within the context of large-scale probabilistic linguistic multi-attribute decision-making.

From the existing studies on three-way decision and multi-attribute decision-making in the probabilistic linguistic term set (PLTS) environment, it can be seen that relevant methods generally focus on several key issues, including the construction of relative loss functions, the determination of attribute weights, and the derivation of conditional probabilities. Regarding the construction of relative loss functions, a relatively rich body of modeling approaches has already been developed. For example, Han *et al.* [19] constructed a relative loss matrix based on expectation information within the PLTS framework, whereas Wang *et al.* [20] further introduced prospect theory into probabilistic linguistic three-way decision-making and characterized the relative losses of different actions from the perspective of gain-loss perception. These studies indicate that the construction of relative loss functions in existing PLTS-TWD research has already been supported by a relatively solid theoretical foundation.

By contrast, there is still room for further improvement in the determination of attribute weights. Existing studies can generally be classified into three categories. The first category involves objective weighting methods, which mainly rely on information entropy, cross entropy, or related divergence measures to reflect the dispersion of attribute evaluation information. For instance, Han *et al.* [19] determined attribute weights by using the entropy differences of probabilistic linguistic information, while Xu *et al.* [21] proposed fuzzy entropy, hesitant entropy, and total entropy for PLTSs. The second category involves subjective weighting methods, in which attribute importance is determined directly according to the preferences of decision makers. For example, Wang *et al.* [22] and Li *et al.* [23] assigned attribute weights subjectively based on decision-makers' preferences. The third category consists of combined subjective-objective weighting methods, which integrate both decision-makers' preference information and the intrinsic differences in evaluation data to derive combined weights. For example, Chen and Yin [24] combined subjective weights obtained through the probabilistic linguistic best-worst method with objective weights derived from similarity minimization, while Wang *et al.* [25] further incorporated the importance of both experts and attributes within a probabilistic linguistic group decision-making framework to construct a more adaptive weighting scheme.

Overall, although existing weighting studies have evolved from purely subjective or purely objective methods to combined subjective-objective weighting approaches, research on developing dedicated weighting measures tailored to the intrinsic characteristics of probabilistic linguistic information remains relatively limited. A PLTS contains not only expectation information, but also features such as probability distributions, hesitancy, and potential incompleteness. If traditional weighting tools are directly applied, they often fail to fully reflect the actual role of attributes in distinguishing alternatives. Therefore, it is necessary to develop new entropy-based measures in the probabilistic linguistic environment so as to characterize attribute information differences and discriminative ability more accurately.

In addition, although existing studies have begun to pay attention to the application of sequential three-way decision-making in complex decision problems, they have generally focused more on the sequential discrimination framework itself, or applied it to specific issues such as group consensus and risk ranking. Sequential three-way decision emphasizes that, during multi-stage or multi-level information utilization, those objects that have not yet been clearly classified should continue to undergo subsequent three-way partitioning, thereby gradually reducing the uncertain region and yielding more refined discrimination results [8]. However, in the context of large-scale group multi-attribute decision-making, there is still a lack of systematic research on how to directly embed attribute importance into the sequential discrimination process, so that alternatives can be screened progressively as attribute information is invoked step by step. Especially in the probabilistic linguistic environment, although some scholars have extended sequential three-way decision-making to group decision-making and demonstrated its feasibility in dynamic judgment and uncertainty processing [15], discussion of mechanisms for progressively advancing discrimination according to the order of attribute weights remains insufficient. In other words, existing research on sequential three-way decision provides important theoretical inspiration for this study, but has not directly addressed the problem of attribute-driven sequential partitioning in large-scale probabilistic linguistic multi-attribute decision-making scenarios.

With regard to conditional probability derivation, existing studies mostly rely on intermediate measures such as dominance relations, similarity, closeness to ideal solutions, or behavioral comparison relations to realize three-way partitioning. For example, Han *et al.* [19] derived conditional probabilities based on dominance sets constructed by TODIM; Wang *et al.* [20] incorporated decision-makers' psychological and behavioral factors into three-way decision-making under the framework of prospect theory, thereby enhancing the behavioral interpretability of conditional probability construction; and Chen and Yin [24] further introduced behavior-based ranking ideas into the probabilistic linguistic decision-making process within the TODIM framework. Nevertheless, taken as a whole, existing methods still rely heavily on static comparison indicators and remain insufficient in providing a unified characterization of both decision-makers' psychological behavioral features and the non-compensatory relationships among attributes. Therefore, how to construct a more behaviorally rational conditional probability derivation method in the probabilistic linguistic environment remains a worthwhile issue for further investigation.

Based on the above analysis, this paper proposes a FCM-TODIM based sequential three-way decision method for large group multi-attribute decision-making in a probabilistic linguistic term environment. The proposed method has the following main features:

1) First, to address the problems of a large number of experts and strong heterogeneity in evaluation patterns in large group decision-making, the FCM is employed to cluster experts, thereby obtaining more structured cluster-level evaluation information. Compared with existing probabilistic linguistic three-way decision methods [19] [22] [23], the proposed method is better suited to large multi-expert decision-making scenarios.

2) Second, a new deviation-range entropy formula is proposed in the probabilistic linguistic environment to determine attribute weights, thereby providing a new objective basis for measuring attribute importance. Compared with existing probabilistic linguistic fuzzy entropy formulas [21] [26], the proposed formula has stronger discriminative power.

3) Third, alternatives are sequentially subjected to three-way partitioning according to the descending order of attribute weights, and only those falling into the boundary region under the current attribute are passed to the next attribute for further discrimination until the boundary region becomes empty. In this way, an attribute-driven sequential decision-making mechanism is constructed. Compared with existing three-way decision methods [19] [22] [23], the proposed method can simplify the decision-making process to a certain extent.

4) Finally, the TODIM method is introduced to derive decision conditional probabilities while taking the psychological behavioral characteristics of decision makers into account, thereby enhancing the behavioral rationality and practical applicability of the model. Compared with three-way decision methods that construct conditional probabilities based on dominance sets [19] [27] and methods that derive conditional probabilities using TOPSIS [28], the proposed method is less dependent on predefined dominance sets and is able to reflect the psychological factors of decision makers.

The remainder of this paper is organized as follows. Section 2 briefly reviews PLTSs, the STWD method, and the TODIM method. Section 3 introduces the deviation-range entropy of PLTSs. Section 4 develops the FCM-TODIM-STWD method under the PLTS environment. Section 5 presents the detailed procedure and algorithms of the proposed method. Section 6 provides a case study and comparative analyses to illustrate the feasibility and superiority of the proposed method. Section 7 conducts sensitivity analyses to evaluate the robustness of the proposed method. Finally, Section 8 concludes the paper. Moreover, the major notations and their meanings are listed in **Table 1** to enhance readability.

Table 1. Symbols and their corresponding explanation.

Notations	Corresponding meanings
X	The target state of concern
$\neg X$	The complementary state of X

Continued

$U(L_{ij}(P))$	The deviation-range entropy of alternative A_i under attribute C_j
$E^T(L_{ij}(p))$	The expectation value of alternative A_i under attribute C_j at granularity level T
$E^T_{\max}$	The maximum values of expectation with respect to attribute C_j
$E^T_{\min}$	The minimum values of expectation with respect to attribute C_j
$ac_\alpha(\alpha = P, B, N)$	The acceptance, delay, or rejection action
η	The risk aversion coefficient
$\mu^{T_i}_{\alpha\beta}(\beta = P, N)$	The relative loss functions of A_i at granularity level T
$POS^{GS^T}(X)$	The positive region at granularity level T
$BND^{GS^T}(X)$	The boundary region at granularity level T
$NEG^{GS^T}(X)$	The negative region at granularity level T
κ_{ho}	the membership degree of expert e_h belonging to cluster G_o ,
W_j	The weight of attribute C_j
W_{jr}	The relative weight of attribute C_j with respect to the reference attribute C_r
θ	The loss aversion coefficient
$dom^T(A^{T_i}, A^{T_k})$	The dominance degree of alternative A_i over alternative A_k at granularity level T
$\delta^T(A^{T_i})$	The incremental overall dominance degree of alternative A_i over all the other alternatives at granularity level T
$Pr^T(X\|[A^{T_i}])$	The probabilities of the object A_i belonging to X at granularity level T

2. Preliminaries

To facilitate the discussion in the sequel, this section reviews some theories on PLTSs, sequential three way decision and the TODIM.

2.1. Probabilistic Linguistic Term Sets

As mentioned in the Introduction, PLTSs can flexibly express the decision information by several possible linguistic terms associated with their probabilities, so it is a popular tool while evaluating alternatives.

Definition 1 [29] Let $S = \{s_t | t = -\tau, \cdots, -1, 0, 1, \cdots, \tau\}$ be a finite and ordered set of discrete linguistic terms (LTS), $2\tau + 1$ is the granularity of S, s_t indicates a linguistic term, s_0 indicates that the linguistic value is "undifferentiated", $s_{-\tau}$ and s_τ denotes the two boundaries of a linguistic term, symmetrically distributed on either side of s_0.

Definition 2 [30] Let $S = \{s_t \mid t = -\tau, \cdots, -1, 0, 1, \cdots, \tau\}$ be a LTS, the linguistic term s_t that expresses the equivalent information to the membership degree γ is obtained by a linguistic scale function f:

$$f : [-\tau, \tau] \to [0,1], f(s_t) = \frac{t + \tau}{2\tau} = \gamma. \tag{1}$$

Further, the membership degree γ that expresses the equivalent information to the linguistic variable s_t is obtained by the following function f^{-1}:

$$f^{-1}:[0,1]\to[s_{-\tau},s_\tau], f^{-1}(\gamma)=s_{(2\gamma-1)\tau}=s_t. \quad (2)$$

Definition 3 [5] Let $S=\{s_t \mid t=-\tau,\cdots,-1,0,1,\cdots,\tau\}$ be a LTS, then the probabilistic linguistic term sets (PLTSs) is defined as

$$L(p)=\left\{L^l(p^l)\mid L^l\in S, p^l\ge 0, l=1,\cdots,\#L(p), \sum_{l=1}^{\#L(p)} p^l\le 1\right\}, \quad (3)$$

where p^l represents the probability corresponding to the linguistic term L^l, and $\#L(p)$ is the number of linguistic terms in $L(p)$.

2.2. Probabilistic Linguistic Term Sets Expectation Score Function

Definition 4 [31] To facilitate comparisons between PLTSs, the expectation score function is defined as

$$E(L(p))=\sum_{l=1}^{\#L(p)}\left(f(I(L^l))p^l\Big/\sum_{l=1}^{\#L(p)}p^l\right), \quad (4)$$

where, $I(L^l)$ represents the subscript t of the l th linguistc L^l.

Furthermore, the variance function of PLTSs is defined as

$$\sigma(L(p))=\left\{\sum_{l=1}^{\#L(p)}\left(\left(f(I(L^l))-E(L(p))\right)^2\cdot p^l\right)\Big/\sum_{l=1}^{\#L(p)}p^l\right\}^{1/2}. \quad (5)$$

According to Equation (4) and Equation (5), the comparison between two PLTSs can be carried out based on the following criteria.

1) If $E(L_1(p))>E(L_2(p))$, then $L_1(p)>L_2(p)$;

2) If $E(L_1(p))<E(L_2(p))$, then $L_1(p)<L_2(p)$;

3) If, $E(L_1(p))=E(L_2(p))$ then

$$\begin{cases}\text{if } \sigma(L_1(p))<\sigma(L_2(p)), \text{then } L_1(p)>L_2(p),\\ \text{if } \sigma(L_1(p))>\sigma(L_2(p)), \text{then } L_1(p)<L_2(p),\\ \text{if } \sigma(L_1(p))=\sigma(L_2(p)), \text{then } L_1(p)=L_2(p).\end{cases} \quad (6)$$

2.3. Probabilistic Linguistic Term Sets Fuzzy Entropy

Definition 5 [21] Let $S=\{s_t \mid t=-\tau,\cdots,-1,0,1,\cdots,\tau\}$ be a linguistic term set, and let $L(p)=\left\{L^l(p^l)\mid L^l\in S, p^l\ge 0, l=1,\cdots,\#L(p),\sum_{l=1}^{\#L(p)}p^l\le 1\right\}$ be a PLTS defined on S. To evaluate the quality of decision information, the fuzzy entropy of a probabilistic linguistic term set is defined as

$$U^X=\frac{\ln\left(1+E(L(p))-\left(E(L(p))\right)^2\right)}{\ln 5-\ln 4}. \quad (7)$$

Subsequently, Wu *et al.* [26] pointed out that Equation (7) has insufficient discriminative power (see Section 3 for details), and proposed a fuzzy entropy formula that simultaneously considers the hesitation among linguistic terms and the characteristics of the probability distribution.

Definition 6 [26] Let $S=\{s_t|t=-\tau,\cdots,-1,0,1,\cdots,\tau\}$ be a linguistic term set, and let $L(p)=\left\{L^l(p^l)\middle|L^l\in S,p^l\geq 0,l=1,\cdots,\#L(p),\sum_{l=1}^{\#L(p)}p^l\leq 1\right\}$ be a PLTS defined on S. Its fuzzy entropy is defined as

$$U^W=\frac{\left(1+\frac{I(L^L)-I(L^1)}{2\tau}\right)\ln\left(1+\frac{I(L^L)-I(L^1)}{2\tau}\right)}{2\ln 2}\times\left(1-\sqrt{\sum_l\left(p^l-\frac{1}{I(L^L)-I(L^1)+1}\right)^2}\right),\tag{8}$$

where $I(L^L)$ and $I(L^1)$ denote the subscripts of the largest linguistic term and the smallest linguistic term in the probabilistic linguistic term set, respectively, and τ denotes the maximum subscript of the linguistic terms.

Although Equation (8) simultaneously considers the degree of hesitation among linguistic terms and the characteristics of the probability distribution, Wu *et al.* [26] only used the information of the maximum and minimum linguistic terms in the probabilistic linguistic term set, without taking the other included linguistic terms into account. As a result, this entropy formula still fails to effectively distinguish certain probabilistic linguistic term sets in some cases (see Section 3 for details).

2.4. The Aggregation Operators for Probabilistic Linguistic Term Sets

Definition 7 [5] Let $S=\{s_t\,|\,t=-\tau,\cdots,-1,0,1,\cdots,\tau\}$ be a linguistic term set, and let $L(p)$, $L_1(p)$, and $L_2(p)$ be three probabilistic linguistic term sets. Here, each probabilistic linguistic term $L^l(p^l)$ is arranged in ascending order according to the value of $I(L^l)\times p^l$ $(l=1,2,\cdots,\#L(p))$, where $I(L^l)$ denotes the subscript t^l of the linguistic term s_{t^l}. Then

1) $L_1(p)\oplus L_2(p)=\bigcup_{L_1^l\in L_1(p),L_2^l\in L_2(p)}\{L_1^lp_1^l\oplus L_1^lp_2^l\}$.
2) $L_1(p)\otimes L_2(p)=\bigcup_{L_1^l\in L_1(p),L_2^l\in L_2(p)}\{L_1^lp_1^l\otimes L_1^lp_2^l\}$.
3) $\lambda L(p)=\bigcup_{L^l\in L(p)}\{\lambda L^lp^l\},\lambda\geq 0$.
4) $(L(p))^\lambda=\bigcup_{L^l\in L(p)}\left\{(L^l)^{\lambda p^l}\right\},\lambda\geq 0$.

Definition 8 [5] Let $L_i(p)=\left\{L_i^l(p_i^l)\middle|l=1,\cdots,\#L_i(p)\right\}(i=1,2,\cdots,n)$ be n probabilistic linguistic term sets, where L_i^l and p_i^l denote the l th linguistic term in $L_i(p)$ and its associated probability, respectively. Then, the probabilistic linguistic averaging operator (*PLA*) is defined as

$$PLA\left(L_1(p),L_2(p),\cdots,L_n(p)\right)=\frac{1}{n}\left(L_1(p)\oplus L_2(p)\oplus\cdots\oplus L_n(p)\right)$$
$$=\frac{1}{n}\left(\bigcup_{L_1^l\in L_1(p),L_2^l\in L_2(p),\cdots,L_n^l\in L_n(p)}\left\{L_1^l p_1^l\oplus L_2^l p_2^l\oplus\cdots\oplus L_n^l p_n^l\right\}\right). \quad (9)$$

Definition 9 [5] Let $L_i(p)=\left\{L_i^l\left(p_i^l\right)|l=1,\cdots,\#L_i(p)\right\}(i=1,2,\cdots,n)$ be n probabilistic linguistic term sets, where L_i^l and p_i^l denote the l th linguistic term in $L_i(p)$ and its associated probability, respectively. Then, the probabilistic linguistic weighted averaging operator (*PLWA*) is defined as

$$PLWA\left(L_1(p),L_2(p),\cdots,L_n(p)\right)=w_1L_1(p)\oplus w_2L_2(p)\oplus\cdots\oplus w_nL_n(p)$$
$$=\bigcup_{L_1^l\in L_1(p)}\left\{w_1p_1^lL_1^l\right\}\oplus\bigcup_{L_2^l\in L_2(p)}\left\{w_2p_2^lL_2^l\right\}\oplus\cdots\oplus\bigcup_{L_n^l\in L_n(p)}\left\{w_np_n^lL_n^l\right\}, \quad (10)$$

where, $\omega=\left(\omega_1,\omega_2,\cdots,\omega_n\right)^{\mathrm{T}}$ is the weight vector of $L_i(p)(i=1,2,\cdots,n)$, satisfying $\omega_i\ge 0$ and $\sum_{i=1}^n\omega_i=1$. When $\omega_1=\omega_2=\cdots=\omega_n=\frac{1}{n}$, Equation (10) reduces to Equation (9).

The probabilistic linguistic aggregation operators proposed by Pang *et al.* [5] (Equations (9)-(10)) may lead to the loss of probability information when aggregating probabilistic linguistic term sets. Inspired by Zhang *et al.* [32], this paper further proposes a new probabilistic linguistic aggregation operator.

Definition 10 Let $L_i(p)=\left\{L_i^l\left(p_i^l\right)|l=1,\cdots,\#L_i(p)\right\}(i=1,2,\cdots,n)$ be n probabilistic linguistic term sets, where L_i^l and p_i^l denote the l th linguistic term in $L_i(p)$ and its associated probability, respectively. Then, the improved probabilistic linguistic averaging operator (*PLA*) is defined as

$$PLA\left(L_1(p),L_2(p),\cdots,L_n(p)\right)=\frac{1}{n}\left(L_1(p)\oplus L_2(p)\oplus\cdots\oplus L_n(p)\right)$$
$$=\frac{1}{n}\left(\bigcup_{L_1^l\in L_1(p),L_2^l\in L_2(p),\cdots,L_n^l\in L_n(p)}\left\{L_1^l p_1^l\oplus L_2^l p_2^l\oplus\cdots\oplus L_n^l p_n^l\left(p_1^l\cdot p_2^l\cdots p_n^l\right)\right\}\right), \quad (11)$$

where, $L^l p^l=I\left(L^l\right)p^l$, $I\left(L^l\right)$ denotes the subscript t^l of the linguistic term s_{t^l}, and p^l is the corresponding probability. Since *PLA*-type operators may generate linguistic terms with non-integer subscripts, while the linguistic term set $S=\left\{s_t\middle|t=-\tau,\cdots,-1,0,1,\cdots,\tau\right\}$ used in practice contains only integer subscripts, the subscripts of the aggregated linguistic terms should be rounded to the nearest integer.

Definition 11 Let $L_i(p)=\left\{L_i^l\left(p_i^l\right)|l=1,\cdots,\#L_i(p)\right\}(i=1,2,\cdots,n)$ be n probabilistic linguistic term sets, where L_i^l and p_i^l denote the l th linguistic term in $L_i(p)$ and its associated probability, respectively. Then, the improved probabilistic linguistic weighted averaging operator (*PLWA*) is defined as

$$PLWA\left(L_1(p),L_2(p),\cdots,L_n(p)\right)$$
$$=w_1L_1(p)\oplus w_2L_2(p)\oplus\cdots\oplus w_nL_n(p) \quad (12)$$
$$=\bigcup_{L_1^l\in L_1(p)}\left\{w_1p_1^lL_1^l\right\}\oplus\bigcup_{L_2^l\in L_2(p)}\left\{w_2p_2^lL_2^l\right\}\oplus\cdots\oplus\bigcup_{L_n^l\in L_n(p)}\left\{w_np_n^lL_n^l\right\}\left(\omega_1p_1\cdot\omega_2p_2\cdots\omega_np_n\right),$$

where, $\omega = (\omega_1, \omega_2, \cdots, \omega_n)^{\mathrm{T}}$ is the weight vector of $L_i(p)(i = 1, 2, \cdots, n)$, satisfying $\omega_i \geq 0$ and $\sum_{i=1}^{n} \omega_i = 1$. When $\omega_1 = \omega_2 = \cdots = \omega_n = \frac{1}{n}$, Equation (12) reduces to Equation (11).

Assume that two probabilistic linguistic term sets are given by $L_1(p) = \{s_1(0.5), s_2(0.5)\}$ and $L_2(p) = \{s_2(0.2), s_3(0.8)\}$. According to Equation (11), these two probabilistic linguistic term sets are aggregated. Since each set contains two linguistic terms, there are $2 \times 2 = 4$ possible combinations, that is,

$$PLA(L_1(p), L_2(p))$$

$$= \frac{1}{2}\{s_{1\times0.5+2\times0.2}(0.5\times0.2), s_{1\times0.5+3\times0.8}(0.5\times0.8), s_{2\times0.5+2\times0.2}(0.5\times0.2), s_{2\times0.5+3\times0.8}(0.5\times0.8)\}$$

$$= \frac{1}{2}\{s_{0.9}(0.1), s_{2.9}(0.4), s_{1.4}(0.1), s_{3.4}(0.4)\} = \{s_{0.45}(0.1), s_{1.45}(0.4), s_{0.7}(0.1), s_{1.7}(0.4)\}.$$

After rounding the subscripts of the linguistic terms in $PLA(L_1(p), L_2(p)) = \{s_{0.45}(0.1), s_{1.45}(0.4), s_{0.7}(0.1), s_{1.7}(0.4)\}$, obtain $PLA(L_1(p), L_2(p)) = \{s_1(0.1), s_2(0.4), s_1(0.1), s_2(0.4)\}$. By merging the linguistic terms with identical subscripts, the final result is $PLA(L_1(p), L_2(p)) = \{s_1(0.2), s_2(0.8)\}$.

2.5. Sequential Three-Way Decisions

Three way decision (TWD) originates from real-life cognitive and behavioral patterns. Different from two way decision, TWD classifies objects into three regions, including positive, negative and boundary domains. This cognitive pattern is common in the real decision making. For example, while evaluating a potential customer in business marketing, companies often adopt one of three strategies: converting this customer into a real one (positive), abandoning (negative) or continuously observing and tracking (boundary). Drawing on this cognitive pattern, Yao [7] introduced the TWD theory based on the rough set theory and Bayesian decision rules.

Let the alternative set $A = \{A_1, A_2, \cdots, A_n\}$, attribute set $C = \{C_1, C_2, \cdots, C_m\}$ and the attribute weight vector $W = \{W_1, W_2, \cdots, W_m\}^{\mathrm{T}}$ satisfying $\sum_{j=1}^{m} W_j = 1$ and $0 \leq W_j \leq 1$. Suppose the state set $\Theta = (X, \neg X)$ expressing that an object is in or not in the state X. The set of actions can be denoted as $ac_\alpha(\alpha = P, B, N)$, where the three actions, *i.e.*, ac_P, ac_B and ac_N, stand for the acceptance, delay and rejection actions, respectively. The loss functions with respect to action risks are represented as a 3×2 matrix. When the object is in state X, Q_{PP}, Q_{BP} and Q_{NP} represent the losses while taking actions ac_P, ac_B and ac_N. Similarly, when the object is in state $\neg X$, Q_{PN}, Q_{BN} and Q_{NN} represent the losses while taking actions ac_P, ac_B and ac_N. Given that an object A is described by its equivalence class $[A]$, $\Pr(X \mid [A])$ and $\Pr(\neg X \mid [A])$ indicate the probabilities of the object A belonging to X or $\neg X$, respectively. Furthermore, $\Pr(X \mid [A]) + \Pr(\neg X \mid [A]) = 1$.

According to the above statements, when we perform an action

$ac_\alpha\left(\alpha=P,B,N\right)$, the loss function for each object A with respective to each action $ac_\alpha\left(\alpha=P,B,N\right)$ can be expressed as **Table 2**.

Table 2. The loss functions.

	X	$\neg X$
ac_P	Q_{PP}	Q_{PN}
ac_B	Q_{BP}	Q_{BN}
ac_N	Q_{NP}	Q_{NN}

In the classical TWD models [7] [22], the loss functions are often assigned by DMs based on their experiences. Furthermore, these functions are assumed the loss for different objects when these objects take the same action in the same state. Obviously, this assumption may have certain limitations because the losses on different objects are usually distinct in real decision making. To overcome such limitations, some studies [20] [33] [34] introduced reference points and defined relative loss functions which are popular in many TWD models. By extending the frame of these relative loss functions into the PLTS environment, Han *et al.* [19] defined the relative loss functions shown in **Table 3** based on expectation scores of attribute values.

Table 3. The relative loss functions based the expectation score function of PLTSs.

	X	$\neg X$
ac_P	$\mu_{PP}=0$	$\mu_{PN}=E_j^{\max}-E\left(L_{ij}\left(p\right)\right)$
ac_B	$\mu_{BP}=\eta\left(E\left(L_{ij}\left(p\right)\right)-E_j^{\min}\right)$	$\mu_{BN}=\eta(E_j^{\max}-E\left(L_{ij}\left(p\right)\right)$
ac_N	$\mu_{NP}=E\left(L_{ij}\left(p\right)\right)-E_j^{\min}$	$\mu_{NN}=0$

Table 4. The aggregated relative loss functions based the expectation score function of PLTSs.

	X	$\neg X$
ac_P	$\mu_{PP}=0$	$\mu_{PN}=\sum_{j=1}^{m}W_j\left(E_j^{\max}-E\left(L_{ij}\left(p\right)\right)\right)$
ac_B	$\mu_{BP}=\sum_{j=1}^{m}W_j\left(\eta\left(E\left(L_{ij}\left(p\right)\right)-E\right.\right.$	$\mu_{BN}=\sum_{j=1}^{m}W_j\left(\eta\left(E_j^{\max}-E\left(L_{ij}\left(p\right)\right)\right)\right)$
ac_N	$\mu_{NP}=\sum_{j=1}^{m}W_j\left(E\left(L_{ij}\left(p\right)\right)-E_j^{\min}\right.$	$\mu_{NN}=0$

In **Table 3**, $L_{ij}\left(p\right)$ are PLTSs and represent attribute values of alternative $A_i\left(i=1,2,\cdots,n\right)$ on attribute C_j $\left(j=1,2,\cdots,m\right)$. $E\left(L_{ij}\left(p\right)\right)$ is the expectation score of $L_{ij}\left(p\right)$ obtained by Equation (4). $E_j^{\max}$ and $E_j^{\min}$ are respectively the maximum and minimum expectation scores on attribute C_j, respectively. $\eta\in\left(0,0.5\right)$ is the risk aversion coefficient. By weighting the relative loss functions with attribute weights W_j $\left(j=1,2,\cdots,m\right)$, the aggregated relative loss functions are generated and shown in Table 4.

According to the above aggregated relative loss functions, the expected loss function $R(ac_\alpha|[A])$ is represented by the Bayesian decision theory as

$$R(ac_\alpha|[A]) = \mu_{\alpha P}\Pr(X|[A]) + \mu_{\alpha N}\Pr(\neg X|[A]). \tag{13}$$

According to the principle of Bayesian minimum risk decision-making, the optimal action is the one with the mimimum expected loss. Thus, the corresponding decision rule can be formulated as

$$\begin{aligned}
&(P0)\text{ if } R(ac_P|[A]) \le R(ac_B|[A]) \text{ and } R(ac_P|[A]) \le R(ac_N|[A]), \text{then } A \in POS(X);\\
&(B0)\text{ if } R(ac_B|[A]) \le R(ac_P|[A]) \text{ and } R(ac_B|[A]) \le R(ac_N|[A]), \text{then } A \in BND(X);\\
&(N0)\text{ if } R(ac_N|[A]) \le R(ac_P|[A]) \text{ and } R(ac_N|[A]) \le R(ac_B|[A]), \text{then } A \in NEG(X).
\end{aligned} \tag{14}$$

In Equation (14), $POS(X)$, $BND(X)$ and $NEG(X)$ respectively represent the positive region, boundary region and the negative region. $A \in POS(X)$ means that alternative A is accepted. Similarly, $A \in BND(X)$ indicates that alternative A is delayed. $A \in NEG(X)$ expresses that alternative A is rejected.

Yao [8] extended the three-way decision model to a multi-step decision-making framework and proposed the sequential three-way decision model. Sequential three-way decision is a dynamic and multi-level decision-making approach for modeling human problem-solving processes, and it mainly relies on the multi-granularity structure of decision objects. Its definition can be described as follows.

Definition 12 [35] Suppose that a decision information table is given as $Y = \{A, R = C \cup D, V, g\}$, where A denotes a finite nonempty set of objects, also called the universe; C denotes a finite nonempty set of conditional attributes; D denotes the set of decision attributes; V denotes the set of attribute values; and: $A \times R \to V$ is a mapping such that each object in A has a corresponding value for each attribute. If there exists a collection of conditional attribute sets satisfying $C^1 \subseteq C^2 \subseteq \cdots \subseteq C^m \subseteq C$ then a multi-level granularity structure can be constructed based on equivalence relations, which is expressed as follows

$$GS = (GS^1, GS^2, \cdots, GS^m), \tag{15}$$

$$GS^T = (U^T, C^T \cup D^T, V^T, g^T). \tag{16}$$

For each granularity level GS^T, there exists a pair of thresholds (x^T, y^T) for partitioning the universe, where the thresholds satisfy $x^T \ge y^T$. Under this granularity structure, the decision rules can be expressed as

$$\begin{aligned}
&(P1)\text{ if } \Pr(X^T|[A^T]) \ge x^T, \text{ then } A^T \in POS^{GS^T}(X);\\
&(B1)\text{ if } y^T < \Pr(X^T|[A^T]) < x^T, \text{ then } A^T \in BND^{GS^T}(X);\\
&(N1)\text{ if } \Pr(X^T|[A^T]) \le y^T, \text{ then } A^T \in NEG^{GS^T}(X).
\end{aligned} \tag{17}$$

Accordingly, at each granularity level, the positive region, boundary region, and negative region can be defined as

$$\begin{aligned} &POS^{GS^T}(X)=\left\{A\in A^T \middle| \Pr\left(X^T\middle|[A]\right)\geq x^T\right\};\\ &BND^{GS^T}(X)=\left\{A\in A^T \middle| y^T<\Pr\left(X^T\middle|[A]\right)<x^T\right\};\\ &NEG^{GS^T}(X)=\left\{A\in A^T \middle| \Pr\left(X^T\middle|[A]\right)\leq y^T\right\}. \end{aligned} \tag{18}$$

As the granularity becomes finer from coarse to fine, the positive region and negative region gradually expand, whereas the boundary region gradually shrinks. The three regions under different granularity structures satisfy the following relationships

$$\begin{aligned} &POS^{GS^T}(X)\subseteq POS^{GS^{T+1}}(X);\\ &BND^{GS^{T+1}}(X)\subseteq BND^{GS^T}(X);\\ &NEG^{GS^T}(X)\subseteq NEG^{GS^{T+1}}(X). \end{aligned} \tag{19}$$

2.6. TODIM

The TODIM method [36] is a common approach to MADM problems, which has been widely used to different fields. Liang *et al.* [37] used the distance between PLTSs as a tool to express the dominance relation among them, which extends the TODIM method to the probabilistic environment.

Suppose there are n alternatives $A=\{A_1,A_2,\cdots,A_n\}$, m attributes $C=\{C_1,C_2,\cdots,C_m\}$, the weight vector of every attribute is $W=\{W_1,W_2,\cdots,W_m\}$, satisfies $W_j\in[0,1]$, $\sum_{j=1}^{m}W_j=1$.

The first step is to calculate the ratio of all weights to the most important weight which is also the ratio to the maximum weight:

$$W_{jr}=\frac{W_j}{W_r}, \tag{20}$$

where, W_r is the largest attribute weight.

The second step is to calculate the formula for the degree of dominance between the two alternatives under each attribute $C_j\left(j=1,2,\cdots,m\right)$:

$$dom_j\left(A_i,A_k\right)=\begin{cases}\sqrt{\dfrac{W_{jr}d\left(L_{ij},L_{kj}\right)}{\sum_{j=1}^{m}W_{jr}}}, & \text{if } L_{ij}>L_{kj}\\ -\dfrac{1}{\theta}\sqrt{\dfrac{\left(\sum_{j=1}^{m}W_{jr}\right)d\left(L_{ij},L_{kj}\right)}{W_{jr}}}, & \text{if } L_{ij}<L_{kj}\\ 0. & \text{if } L_{ij}\sim L_{kj}\end{cases} \tag{21}$$

$$d\left(L_{ij},L_{kj}\right)=\frac{\left|H\left(L_{ij}\right)-H\left(L_{kj}\right)\right|+\left|v\left(L_{ij}\right)-v\left(L_{kj}\right)\right|}{\tau}, \tag{22}$$

where $\theta>0$, which reflects the impact of the loss. When $\theta>1$, the formula shows that it reduces the degree of loss aversion. Conversely, it increases the degree of loss aversion. And $d\left(L_{ij},L_{kj}\right)$ denotes distance between L_{ij} and L_{kj}, $H\left(L(p)\right)=\sum_{l=1}^{\#L(p)}I\left(L^l\right)p^l$ represent the scores of $L(p)$ and

$v(L(P)) = \sum_{l=1}^{\#L(p)} p^l \left| I(L^l) - H(L(p)) \right|$ denote the variance values of $L(p)$.

The overall dominance between two alternatives is obtained by summing up the dominance under all attributes according to Equation (21):

$$dom(A_i, A_k) = \sum_{j=1}^{m} dom_j(A_i, A_k). \tag{23}$$

In order to obtain the overall performance of the alternatives $A = \{A_1, A_2, \cdots, A_n\}$, the dominance between $A_k (k = 1, 2, \cdots, n)$ and the other alternatives A_k is then added up in full:

$$dom(A_i) = \sum_{i=1}^{n} dom(A_i, A_k). \tag{24}$$

For comparison purposes, the overall performance of each alternatives $A_i (i = 1, 2, \cdots, n)$ normalized as follows:

$$\overline{dom(A_i)} = \frac{dom(A_i) - \min_{k=1}^{n} dom(A_k)}{\max_{k=1}^{n} dom(A_k) - \min_{k=1}^{n} dom(A_k)}, \tag{25}$$

where, $\max_{k=1}^{n} dom(A_k)$ and $\min_{k=1}^{n} dom(A_k)$ are the maximum and minimum dominance value of $A_i (i = 1, 2, \cdots, n)$.

3. Deviation-Range Entropy of Probabilistic Linguistic Term Sets

Fuzzy entropy can be used to characterize the degree of uncertainty of a probabilistic linguistic term set. As discussed in Section 2.3, the fuzzy entropy proposed by Xu *et al.* [21] and that proposed by Wu *et al.* [26] both suffer from insufficient discriminative power and incomplete information representation. Therefore, it is necessary to develop a fuzzy entropy formula with stronger discriminative ability and more comprehensive information utilization.

The uncertainty of a probabilistic linguistic term set is largely determined by its envelope. In general, a longer envelope interval indicates a higher degree of uncertainty. In contrast, when a probabilistic linguistic term set contains only one linguistic term, its uncertainty is zero. The entropy proposed in this paper is inspired by the statistical measurement of the dispersion of random variables, where a probabilistic linguistic evaluation is regarded as a discrete random variable defined on a semantic scale. Specifically, the probability-weighted mean absolute deviation is introduced to characterize probabilistic dispersion, while the span of linguistic terms is used to capture the potential uncertainty space. The product form is then adopted to reflect the joint effect of these two types of uncertainty. To ensure a unified interpretation range and satisfy the normalization requirement that the maximum entropy should be equal to 1, the entropy is further normalized by its maximum value, thereby yielding a normalized semantic entropy index with values in the interval $[0,1]$.

Definition 13 Let $S = \{s_t \mid t = -\tau, \cdots, -1, 0, 1, \cdots, \tau\}$ be a linguistic term set. and let $L(p)=\left\{L^l\left(p^l\right)\middle|L^l\in S,p^l\geq 0,l=1,\cdots,\#L(p),\sum_{l=1}^{\#L(p)}p^l\leq 1\right\}$ be a PLTS defined on S. Then, the Deviation-range entropy $U(L(p))$ of a probabilistic linguistic term set is defined as

$$U(L(p))=\frac{\left(\sum_{l=1}^{\#L(p)}p^l\left|\gamma^l-\overline{\gamma}\right|\right)\times\left(1-e^{-R}\right)}{0.5\times\left(1-e^{-1}\right)}. \tag{26}$$

where, p^l denotes the probability associated with the l th linguistic term, and γ^l denotes the membership degree of the l th linguistic term, which can be obtained from Equation (1). The quantity $\overline{\gamma}$ denotes the average of $\gamma^l\left(l=1,2,\cdots,\#L(p)\right)$, namely, $\overline{\gamma}=\frac{\sum_{l=1}^{\#L(p)}\gamma^l}{\#L(p)}$. In addition, $R=\left|\gamma^L-\gamma^1\right|$ denotes the envelope interval of the probabilistic linguistic term set, where γ^L and γ^1 correspond to the membership degrees of the maximum and minimum linguistic terms contained in the probabilistic linguistic term set, respectively.

The entropy of $L(p)$ satisfies the following properties:

1) $U(L(p))=0$ if and only if $L(p)=\{s_t(1)\}$;
2) $U(L(p))=1$ if and only if $L(p)=\{s_{-\tau}(0.5),s_\tau(0.5)\}$;
3) $U(L(p))\to 0$ when $L(p)=\left\{s_t^l(p),s_t^{l+1}(1-p)\right\}$ and $s_t^l\to s_t^{l+1}$;
4) $U(L(p)^c)=U(L(p))$.

The proofs of these properties are given below.

1) Necessity: When $L(p)$ contains only one linguistic term s_t^l, have $\gamma^l=\frac{I\left(s_t^l\right)+\tau}{2\tau}$, $\overline{\gamma}=\frac{\gamma^l}{1}=\gamma^l$. Therefore, $U(L(p))=0$.

Sufficiency: Suppose that $U(L(p))=0$ and assume $\#L(p)>1$. From Equation (26), have $\left(\sum_{l=1}^{\#L(p)}p^l\left|\gamma^l-\overline{\gamma}\right|\right)\times\left(1-e^{-R}\right)=0$. Thus, either $\sum_{l=1}^{\#L(p)}p^l\left|\gamma^l-\overline{\gamma}\right|=0$, or $1-e^{-R}=0$. The former implies that $\left|\gamma^l-\overline{\gamma}\right|=0$ for all $l=1,2,\cdots,\#L(p)$; the latter implies that $R=0$, namely, $\left|\gamma^L-\gamma^1\right|=0$. Hence, $\gamma^L=\gamma^1$, which means that all included linguistic terms have the same membership degree. This contradicts the assumption $\#L(p)>1$. Therefore, $\#L(p)=1$, and thus $L(p)=\{s_t(1)\}$.

2) Necessity: When $L(p)=\{s_{-\tau}(0.5),s_\tau(0.5)\}$, have

$\overline{\gamma}=\frac{\frac{-\tau+\tau}{2\tau}+\frac{\tau+\tau}{2\tau}}{2}=0.5$. Therefore,

$U(L(p))=\frac{(0.5\times 0.5+0.5\times 0.5)\times\left(1-e^{-1}\right)}{0.5\times\left(1-e^{-1}\right)}=1$.

Sufficiency: Suppose that $U(L(p))=1$. Then, from Equation (26), obtain $\left(\sum_{l=1}^{\#L(p)}p^l\left|\gamma^l-\overline{\gamma}\right|\right)\times\left(1-e^{-R}\right)=0.5\times\left(1-e^{-1}\right)$. Hence, $1-e^{-R}=1-e^{-1}$, $\sum_{l=1}^{\#L(p)}p^l\left|\gamma^l-\overline{\gamma}\right|=0.5$. This yields $R=1$, that is, $\left|\gamma^L-\gamma^1\right|=1$. Since $\gamma^l\in[0,1]$, it follows that $\gamma^L = 1$ and $\gamma^1 = 0$. Furthermore, together with $\sum_{l=1}^{\#L(p)} p^l \left|\gamma^l - \bar{\gamma}\right| = 0.5$, the maximum is attained only when the two endpoint linguistic terms occur with probabilities 0.5, namely, $p^1 = p^L = 0.5$. Therefore, $L(p) = \{s_{-\tau}(0.5), s_\tau(0.5)\}$.

3) Let $L(p) = \{s_t^l(p), s_t^{l+1}(1-p)\}$. Then,

$$U(L(p)) = \frac{\left(p\left|\gamma^l - \gamma^{l+1}\right| + \frac{\gamma^{l+1} - \gamma^l}{2}\right) \times \left(1 - e^{-\left|\gamma^{l+1} - \gamma^l\right|}\right)}{0.5 \times \left(1 - e^{-1}\right)}.$$ When $s_t^l \to s_t^{l+1}$, we have $\gamma^{l+1} - \gamma^l \to 0$, and thus $U(L(p)) \to 0$.

4) Since $L(p)^c = f^{-1}\left(\bigcup_{l=1,2,\cdots,\#L(p)} \left\{\left(1 - f\left(s_t^l\right)\right)\left(p^l\right)\right\}\right)$, have

$$U\left(L(p)^c\right) = \frac{\left(\sum_{l=1}^{\#L(p)} p^l \left|\left(1-\gamma^l\right) - \left(1 - \bar{\gamma}\right)\right|\right) \times \left(1 - e^{-R}\right)}{0.5 \times \left(1 - e^{-1}\right)}.$$ Because $\left|\left(1-\gamma^l\right) - \left(1-\bar{\gamma}\right)\right| = \left|\gamma^l - \bar{\gamma}\right|$, it follows that

$$U\left(L(p)^c\right) = \frac{\left(\sum_{l=1}^{\#L(p)} p^l \left|\gamma^l - \bar{\gamma}\right|\right) \times \left(1 - e^{-R}\right)}{0.5 \times \left(1 - e^{-1}\right)} = U(L(p)).$$

This completes the proof.

Remark 1. Compared with the fuzzy entropy proposed by Xu *et al.* [21] and that proposed by Wu *et al.* [26], the deviation-range entropy proposed in this paper has two key advantages: 1) it captures more aspects of information contained in probabilistic linguistic term sets; 2) it has an enhanced discriminative capability. The detailed comparison is represented as Table 5.

Table 5. Comparison results by different fuzzy entropy functions.

PLTSs	$U(L(p))$	U^X	U^W
$L_1(p) = \{s_1(0.5), s_2(0.5)\}$	0.0405	0.7701	0.1297
$L_3(p) = \{s_0(0.5), s_3(0.5)\}$	0.3112	0.7701	0.2194
$L_4(p) = \left\{s_{-1}\left(\frac{1}{5}\right), s_0\left(\frac{3}{5}\right), s_1\left(\frac{1}{5}\right)\right\}$	0.0598	1	0.1863
$L_5(p) = \left\{s_{-1}\left(\frac{7}{15}\right), s_0\left(\frac{1}{15}\right), s_1\left(\frac{7}{15}\right)\right\}$	0.1395	1	0.1863

As shown in Table 5, for the probabilistic linguistic term sets $L_1(p)$ and $L_3(p)$, the fuzzy entropy proposed by Xu *et al.* [21] fails to distinguish between them. In contrast, both the proposed probabilistic linguistic deviation-range entropy and the fuzzy entropy proposed by Wu *et al.* [26] yield $U(L_1(p)) < U(L_3(p))$, indicating that $L_3(p)$ has a higher degree of uncertainty than $L_1(p)$. This result is consistent with intuition, since the envelope interval of the linguistic terms contained in $L_3(p)$ is longer. Furthermore, although the fuzzy entropy proposed by Wu *et al.* [26] can also distinguish $L_1(p)$

from $L_3(p)$, have
$U(L_3(p)) - U(L_1(p)) = 0.2707 > U^W(L_3(p)) - U^W(L_1(p)) = 0.0897$, which indicates that the proposed probabilistic linguistic deviation-range entropy has stronger discriminative power in distinguishing these two probabilistic linguistic term sets. Similarly, for the probabilistic linguistic term sets $L_4(p)$ and $L_5(p)$, the fuzzy entropies proposed by Xu *et al.* [21] and Wu *et al.* [26] both regard them as indistinguishable, whereas the proposed probabilistic linguistic deviation-range entropy yields $U(L_4(p)) < U(L_5(p))$. This indicates that, although $L_4(p)$ and $L_5(p)$ contain the same linguistic terms, the proposed probabilistic linguistic deviation-range entropy can still distinguish them because of their different probability distributions. This further confirms that the proposed entropy has stronger discriminative ability in characterizing differences in probabilistic linguistic information.

4. A FCM-TODIM-STWD Method under the PLTS Environment

In this section, a FCM-TODIM-STWD method is proposed in the probabilistic linguistic term set (PLTS) environment. First, based on the attribute evaluation information provided by decision makers and the PLTS aggregation operators introduced in Section 2.4, the evaluation values of alternatives given by different experts under different attributes are aggregated to form the corresponding attribute evaluation preferences. Next, the proposed deviation-range entropy is employed to determine the attribute weights. Then, the C-means algorithm is used to cluster the experts, and the evaluation values within each cluster are further aggregated to construct a cluster evaluation matrix. Meanwhile, the weights of the clusters are calculated according to the membership degrees of experts within each cluster, thereby forming the final comprehensive evaluation matrix. On this basis, TODIM is integrated with sequential three-way decision-making to construct a decision procedure in which attribute information is utilized sequentially according to the descending order of attribute weights. Finally, the conditional probabilities and loss functions of the alternatives are calculated to classify the alternatives.

4.1. Description of Problems

In large-group multi-attribute decision making problems, different decision makers may provide different evaluation information. For the convenience of subsequent modeling, the related elements are first defined as follows.

1) $A = \{A_1, A_2, \cdots, A_n\}$ denotes a nonempty finite set consisting of n objects, called the set of alternatives.

2) $C = \{C_1, C_2, \cdots, C_m\}$ denotes a nonempty finite set consisting of m attributes, called the attribute set.

3) $W = (W_1, W_2, \cdots, W_m)^{\mathrm{T}}$ denotes the attribute weight vector, where W_j represents the importance of the j th attribute and satisfies $0 \le W_j \le 1$ and $\sum_{j=1}^{m} W_j = 1$.

4) $e=\{e_1,e_2,\cdots,e_q\}$ denotes a nonempty finite set consisting of q experts, called the expert set.

Assume that all experts in the set $e=\{e_1,e_2,\cdots,e_q\}$ are assigned equal weights, that is, each expert has a weight of $\frac{1}{q}$. The decision matrix provided by the h th expert is denoted by $F^h=\left[L_{ij}^h(p)\right]_{n\times m}$, where $L_{ij}^h(p)$ is a probabilistic linguistic term set representing the evaluation of alternative A_i under attribute C_j given by the h th expert.

To characterize the overall evaluation preference of each expert under each attribute, Equation (11) is employed to aggregate all evaluation values given by expert e_h for all alternatives under attribute C_j, thereby obtaining an integrated probabilistic linguistic term set $\bar{L}_{hj}(p)$. Accordingly, the probabilistic linguistic decision matrix of experts under all attributes can be constructed as $\bar{F}=\left[\bar{L}_{hj}(p)\right]_{q\times m}$, which is given by

$$\bar{F}=\begin{bmatrix}\bar{L}_{11}(p) & \bar{L}_{12}(p) & \cdots & \bar{L}_{1m}(p)\\ \bar{L}_{21}(p) & \bar{L}_{22}(p) & \cdots & \bar{L}_{2m}(p)\\ \vdots & \vdots & \ddots & \vdots\\ \bar{L}_{q1}(p) & \bar{L}_{q2}(p) & \cdots & \bar{L}_{qm}(p)\end{bmatrix}_{q\times m}.$$

4.2. Weight Determination Based on Deviation-Range Entropy

In some existing probabilistic linguistic three-way decision methods, attribute weights are directly provided by decision makers (DMs) [22] [23]. However, such a subjective weighting scheme may cause the decision results to be heavily influenced by the subjective preferences of decision makers. It is well known that fuzzy entropy can reflect the variation degree of indicator data. A larger entropy value indicates smaller data variation and less information provided, and thus a smaller weight should be assigned. Conversely, a smaller entropy value indicates greater data variation and more information provided, and thus a larger weight should be assigned. Therefore, based on the deviation-range entropy formula presented in Section 3 (Equation (26)), this paper employs entropy values to determine the objective weights of attributes as

$$W_j=\frac{1-\frac{1}{q}\sum_{h=1}^{q}U_{hj}}{\sum_{j=1}^{m}\left(1-\frac{1}{q}\sum_{h=1}^{q}U_{hj}\right)}. \tag{27}$$

According to Equation (27), the attribute weight vector can be obtained as $W=(W_1,W_2,\cdots,W_m)^{\mathrm{T}}$.

Remark 2. It should be noted that the purpose of aggregating, for each expert, all alternative-level evaluations under the same attribute is not to replace the discrimination among alternatives, but to construct an expert-level attribute profile for weight determination. In a large-group setting, attribute weights are intended to reflect how informative and behaviorally stable each attribute is across the whole evaluation process, rather than to capture only local dispersion at the individual alternative level. This preprocessing step reduces the possibility that the weight estimation is dominated by isolated alternative-specific fluctuations. Meanwhile, the discrimination among alternatives is still preserved in the subsequent stages, because the final ranking and three-way classification are performed on the aggregated alternative-by-attribute decision matrix after expert clustering. Therefore, the expert-level aggregation used here serves the purpose of weight estimation, while the alternative-level information remains involved in the final decision analysis.

4.3. Expert Clustering Method Based on Fuzzy *C*-Means

The Fuzzy *C*-means (FCM) clustering algorithm [38] determines the fuzzy classification of samples by optimizing an objective function and obtaining the membership degrees of each sample point with respect to different cluster centers.

In large group multi-attribute decision-making problems, let $e=\{e_1,e_2,\cdots,e_q\}$ denote the expert set, and suppose that the experts are partitioned into c clusters, where the number of clusters is determined by $c=\sqrt{\frac{q}{2}}$ [39]. The corresponding cluster centers are denoted by $H=\{G_o \mid o=1,2,\cdots,c\}$. Then, Equation (4) is employed to process the decision matrix $\bar{F}=\left[\bar{L}_{hj}(p)\right]_{q\times m}$, thereby obtaining the expectation score matrix $\bar{E}=\left[\bar{E}_{hj}(p)\right]_{q\times m}$. On this basis, by jointly considering the maximization of sample memberships to cluster centers and the minimization of distances between samples and cluster centers, the following objective function is constructed:

$$\min J=\sum_{h=1}^{q}\sum_{o=1}^{c}(\kappa_{ho})^{f}\sqrt{\sum_{j=1}^{m}W_j\left(E_{hj}-u_{oj}\right)^2},\quad \text{s.t.}\begin{cases}\sum_{o=1}^{c}\kappa_{ho}=1, & \forall h=1,2,\cdots,q,\\ \kappa_{ho}\in[0,1], & \forall h,o.\end{cases} \tag{28}$$

where $u_{hj}=\frac{\sum_{h=1}^{q}(\kappa_{ho})^{f}E_{hj}}{\sum_{h=1}^{q}(\kappa_{ho})^{f}}$, κ_{ho} denotes the membership degree of expert e_h belonging to cluster G_o, and f is the fuzzifier parameter controlling the influence of membership degrees, which is usually set to $f=2$ [38]. Under the normalization constraint of memberships, this objective function minimizes the weighted distance and thus makes each sample tend to belong to the nearest cluster with a higher membership degree, thereby improving intra-cluster compactness while achieving fuzzy partitioning.

Furthermore, the Lagrangian function $\bar{J}=J+\sum_{h=1}^{q}\Lambda_t\left(\sum_{o=1}^{c}\kappa_{ho}-1\right)$ is constructed to solve the elements κ_{ho} in the membership matrix M, where Λ_h denotes the Lagrange multiplier. Accordingly, the membership degree can be obtained as

$$\kappa_{ho} = \left[\sum_{o'=1}^{c} \left(\frac{\sqrt{\sum_{j=1}^{m} W_j \left(E_{hj} - u_{oj} \right)^2}}{\sqrt{\sum_{j=1}^{m} W_j \left(E_{hj} - u_{o'j} \right)^2}} \right)^{\frac{2}{f-1}} \right]^{-1}, \tag{29}$$

where $h = 1,2,\cdots,q$ and $o,o' = 1,2,\cdots,c$.

The corresponding membership matrix M can be expressed as

$$M = \begin{bmatrix} & G_1 & G_2 & \cdots & G_c \\ e_1 & \kappa_{11} & \kappa_{12} & \cdots & \kappa_{1c} \\ e_2 & \kappa_{21} & \kappa_{22} & \cdots & \kappa_{2c} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e_q & \kappa_{q1} & \kappa_{q2} & \cdots & \kappa_{qc} \end{bmatrix}.$$

According to the maximum membership principle, each expert is assigned to the cluster for which the membership degree is the largest. The classification rule is given by

$$\kappa_{G_o}\left(e_h\right) = \max_{1 \le o \le c} \kappa_{ho}, h = 1,2,\cdots,q, \tag{30}$$

where G_o denotes the cluster to which expert e_h is assigned, and $\kappa_{G_o}\left(e_h\right)$ denotes the corresponding maximum membership degree.

After expert classification, let $W^G = \left\{W^{G_1}, W^{G_2}, \cdots, W^{G_c}\right\}^{\mathrm{T}}$ denote the weight vector of the clusters. Let $\kappa_{G_o}\left(e_h\right)$ denote the maximum membership degree corresponding to expert e_h being assigned to cluster G_o. Then, the weight of the o th cluster can be defined as

$$W^{G_o} = \frac{\sum_{e_h \in G_o} \kappa_{G_o}\left(e_h\right)}{\sum_{o=1}^{c} \sum_{e_h \in G_o} \kappa_{G_o}\left(e_h\right)}, o = 1,2,\cdots,c. \tag{31}$$

where $0 \le W^{G_o} \le 1$ and $\sum_{o=1}^{c} W^{G_o} = 1$.

According to the PLA operator defined in Equation (11), the decision matrices provided by experts, $F^h = \left[L_{ij}^h\left(p\right)\right]_{n\times m}$, can be aggregated into the cluster decision matrix $F^{G_o} = \left[L_{ij}^{G_o}\left(p\right)\right]_{n\times m}$. Furthermore, according to the $PLWA$ operator defined in Equation (12), the cluster decision matrices $F^{G_o} = \left[L_{ij}^{G_o}\left(p\right)\right]_{n\times m}$ are further aggregated in a weighted manner, thereby yielding the final decision matrix $F = \left[L_{ij}\left(p\right)\right]_{n\times m}$.

4.4. Sequential Three-Way Decision Based on TODIM

TODIM, proposed by Gomes and Lima [40], is a behavioral decision-making method whose main advantage lies in its ability to characterize the psychological behavior of decision makers. Based on this idea, TODIM is employed in this paper to calculate the dominance degree of alternatives with respect to attributes, and the obtained dominance information is further transformed into conditional probabilities. Then, alternatives are classified according to the threshold rules of three-way decision.

In the specific implementation process, the decision matrix aggregating multi-expert opinions is first obtained by using the fuzzy C-means clustering method described in Section 4.3, namely, $F = \left[L_{ij}(p)\right]_{n\times m}$. According to the attribute weights derived from Equation (27), the relative weight of attribute C_j with respect to the reference attribute C_r is further defined as

$$W_{jr} = \frac{W_j}{W_r}, \tag{32}$$

where W_j denotes the weight of attribute C_j, and $W_r = \max W_j$.

As shown in Section 3, a smaller deviation-range entropy of a probabilistic linguistic term set indicates that it contains more informative decision-making information, and thus the corresponding attribute should be assigned a larger weight. In practical decision-making problems, considering all attributes simultaneously often increases the complexity of the decision process. To address this issue, the idea of sequential three-way decision is introduced in this paper. Based on the attribute weights obtained from Equation (27), the attributes are sorted in descending order and incorporated into the decision process one by one until all alternatives are classified. Different from the strategy of using only the current attribute information at the T th level, the proposed method accumulates the information of all attributes from Level 1 to Level T. Specifically, the overall dominance degree at Level T is composed of the overall dominance degree accumulated up to Level $T-1$ and the newly added dominance degree at the current level. The relative loss function is accumulated in the same recursive manner, while the reference points of the newly added loss at each level are still determined according to the current attribute. If all alternatives can be completely classified by using only the first several high-weight attributes, the decision procedure can be simplified effectively. This method not only takes into account the behavioral preferences of decision makers, but also effectively handles uncertainty in the decision process, and is therefore suitable for complex decision-making scenarios.

Let the reordered attribute sequence be $C^1, C^2, \cdots, C^m$, where C^T denotes the attribute incorporated into the decision process at Level T. Let $E^T\left(L_{ij}(p)\right)$ and $E^T\left(L_{kj}(p)\right)$ denote the expectation of alternatives A_i and A_k under the attribute C^T, respectively. Then, the dominance degree of alternative A_i over alternative A_k at the current level is defined as

$$dom^T\left(A^{T_i}, A^{T_k}\right) = \begin{cases} \sqrt{\dfrac{W_{jr}^T\left(E^T\left(L_{ij}(p)\right) - E^T\left(L_{kj}(p)\right)\right)}{\sum_{j=1}^m W_{jr}}}, & E^T\left(L_{ij}(p)\right) > E^T\left(L_{kj}(p)\right) \\ 0, & E^T\left(L_{ij}(p)\right) = E^T\left(L_{kj}(p)\right) \\ \dfrac{1}{-\theta}\sqrt{\dfrac{\left(\sum_{j=1}^m W_{jr}^T\right)\left(E^T\left(L_{kj}(p)\right) - E^T\left(L_{ij}(p)\right)\right)}{W_{jr}^T}}, & E^T\left(L_{ij}(p)\right) < E^T\left(L_{kj}(p)\right) \end{cases} \tag{33}$$

where θ is the attenuation coefficient of loss, reflecting the degree of loss aversion of decision makers, and W_{jr}^{T} denotes the relative weight of the attribute at Level T with respect to the reference attribute.

Then, the incremental overall dominance degree of alternative A_i over all the other alternatives under the current attribute is defined as

$$\Delta\delta^{T}\left(A^{T_i}\right)=\sum_{k\neq i}^{n} dom^{T}\left(A^{T_i},A^{T_k}\right). \tag{34}$$

On this basis, the cumulative overall dominance degree at Level T is recursively defined as

$$\delta^{T}\left(A^{T_i}\right)=\begin{cases}\Delta\delta^{1}\left(A^{T_i}\right), & T=1,\\ \delta^{T-1}\left(A^{(T-1)_i}\right)+\Delta\delta^{T}\left(A^{T_i}\right), & T\geq 2.\end{cases} \tag{35}$$

Based on Equation (35), the conditional probability of alternative A^{T_i} at Level T is calculated as

$$Pr^{T}\left(X\mid A^{T_i}\right)=\frac{\delta^{T}\left(A^{T_i}\right)-\min\limits_{1\leq i\leq n}\delta^{T}\left(A^{T_i}\right)}{\max\limits_{1\leq i\leq n}\delta^{T}\left(A^{T_i}\right)-\min\limits_{1\leq i\leq n}\delta^{T}\left(A^{T_i}\right)}. \tag{36}$$

Furthermore, let $E^{T}\left(L_{ij}\left(p\right)\right)$ denote the expectation of alternative A^{T_i} under the current attribute C^{T}, and define the reference points of the current attribute as

$$E_{\max}^{T}=\max_{1\leq i\leq n}E^{T}\left(L_{ij}\left(p\right)\right),\ E_{\min}^{T}=\min_{1\leq i\leq n}E^{T}\left(L_{ij}\left(p\right)\right). \tag{37}$$

Based on the reference points of the current attribute, the incremental relative loss function at Level T is constructed as shown in **Table 6**.

Table 6. The dynamic aggregation relative loss function based on the expectation function of PLTSs.

	X	$\neg X$
$ac_{P}^{T_i}$	$\Delta\mu_{PP}^{T_i}=0$	$\Delta\mu_{PN}^{T_i}=W^{T}\left(E_{\max}^{T}-E^{T}\left(L_{ij}\left(p\right)\right)\right)$
$ac_{B}^{T_i}$	$\Delta\mu_{BP}^{T_i}=W^{T}\eta\left(E^{T}\left(L_{ij}\left(p\right)\right)-E_{\min}^{T}\right)$	$\Delta\mu_{BN}^{T_i}=W^{T}\eta\left(E_{\max}^{T}-E^{T}\left(L_{ij}\left(p\right)\right)\right)$
$ac_{N}^{T_i}$	$\Delta\mu_{NP}^{T_i}=W^{T}\left(E^{T}\left(L_{ij}\left(p\right)\right)-E_{\min}^{T}\right)$	$\Delta\mu_{NN}^{T_i}=0$

where W^{T} denotes the weight of the attribute at Level T, and $E_{\max}^{T}$ and $E_{\min}^{T}$ denote the maximum and minimum expectation of all alternatives under the current attribute, respectively.

On this basis, the cumulative relative loss at Level T is recursively defined as

$$\mu_{\alpha\beta}^{T_i}=\begin{cases}\Delta\mu_{\alpha\beta}^{1_i}, & T=1,\\ \mu_{\alpha\beta}^{(T-1)_i}+\Delta\mu_{\alpha\beta}^{T_i}, & T\geq 2,\end{cases} \tag{38}$$

where $\alpha\in\left\{P,B,N\right\}$ and $\beta\in\left\{P,N\right\}$.

Based on the cumulative relative loss function, the thresholds are calculated as

$$x^{T_i} = \frac{\mu_{PN}^{T_i} - \mu_{BN}^{T_i}}{\left(\mu_{PN}^{T_i} - \mu_{BN}^{T_i}\right) + \left(\mu_{BP}^{T_i} - \mu_{PP}^{T_i}\right)}, \quad y^{T_i} = \frac{\mu_{BN}^{T_i} - \mu_{NN}^{T_i}}{\left(\mu_{BN}^{T_i} - \mu_{NN}^{T_i}\right) + \left(\mu_{NP}^{T_i} - \mu_{BP}^{T_i}\right)}. \tag{39}$$

Then, alternatives are classified according to the decision rules $(P1)$ - $(N1)$:

$$\begin{aligned} &(P1) \text{ if } Pr^T\left(X \middle| \left[A^{T_i}\right]\right) \geq x^{T_i}, \text{ then } A^{T_i} \in POS^{GS^T}(X); \\ &(B1) \text{ if } y^{T_i} < Pr^T\left(X \middle| \left[A^{T_i}\right]\right) < x^{T_i}, \text{ then } A^{T_i} \in BND^{GS^T}(X); \\ &(N1) \text{ if } Pr^T\left(X \middle| \left[A^{T_i}\right]\right) \leq y^{T_i}, \text{ then } A^{T_i} \in NEG^{GS^T}(X). \end{aligned} \tag{40}$$

According to the decision rules $(P1)$ - $(N1)$, each alternative is assigned to one of the three regions at Level T, namely, the positive region, the boundary region, or the negative region. Alternatives that have entered the positive or negative region will not participate in subsequent classification, whereas those remaining in the boundary region will continue to be evaluated by incorporating the next attribute level until all alternatives are classified or all attributes have been included in the decision process.

5. FCM-TODIM-STWD Method for Probabilistic Linguistic Term Sets

The specific procedure of the FCM-TODIM-STWD method for probabilistic linguistic term sets is summarized as follows.

***Step* 1:** Decision makers evaluate the alternatives and construct the evaluation matrix of the h th expert, denoted by $F^h = \left[L_{ij}^h(p)\right]_{n\times m}$.

***Step* 2:** Aggregate the evaluation information of each expert over all alternatives under each attribute to form the expert-attribute preference matrix $\bar{F} = \left[\bar{L}_{hj}(p)\right]_{q\times m}$ by Equation (11).

***Step* 3:** Transform the expert-attribute preference matrix $\bar{F} = \left[\bar{L}_{hj}(p)\right]_{q\times m}$ into the entropy matrix $Q = \left[U\left(\bar{L}_{hj}(p)\right)\right]_{q\times m} = \left[U_{hj}\right]_{q\times m}$ by Equation (26).

***Step* 4:** According to Equation (27), calculate the attribute weights $W_j\,(j = 1, 2, \cdots, m)$.

***Step* 5:** Convert the expert-attribute preference matrix $\bar{F} = \left[\bar{L}_{hj}(p)\right]_{q\times m}$ into the expectation score matrix $\bar{E} = \left[E\left(\bar{L}_{hj}(p)\right)\right]_{q\times m} = \left[\bar{E}_{hj}\right]_{q\times m}$ by Equation (4).

***Step* 6:** Using the fuzzy C-means clustering method described in Section 4.3, partition the expert set $\{e_1, e_2, \cdots, e_q\}$ into $\sqrt{\frac{q}{2}}$ clusters.

***Step* 7:** Aggregate the evaluation matrices of experts within each cluster to obtain the cluster evaluation matrix $F^{G_o} = \left[L_{ij}^{G_o}(p)\right]_{n\times m}$ by Equation (11).

***Step* 8:** Calculate the cluster weights $W^{G_o}, o = 1, 2, \cdots, c$ by Equation (31).

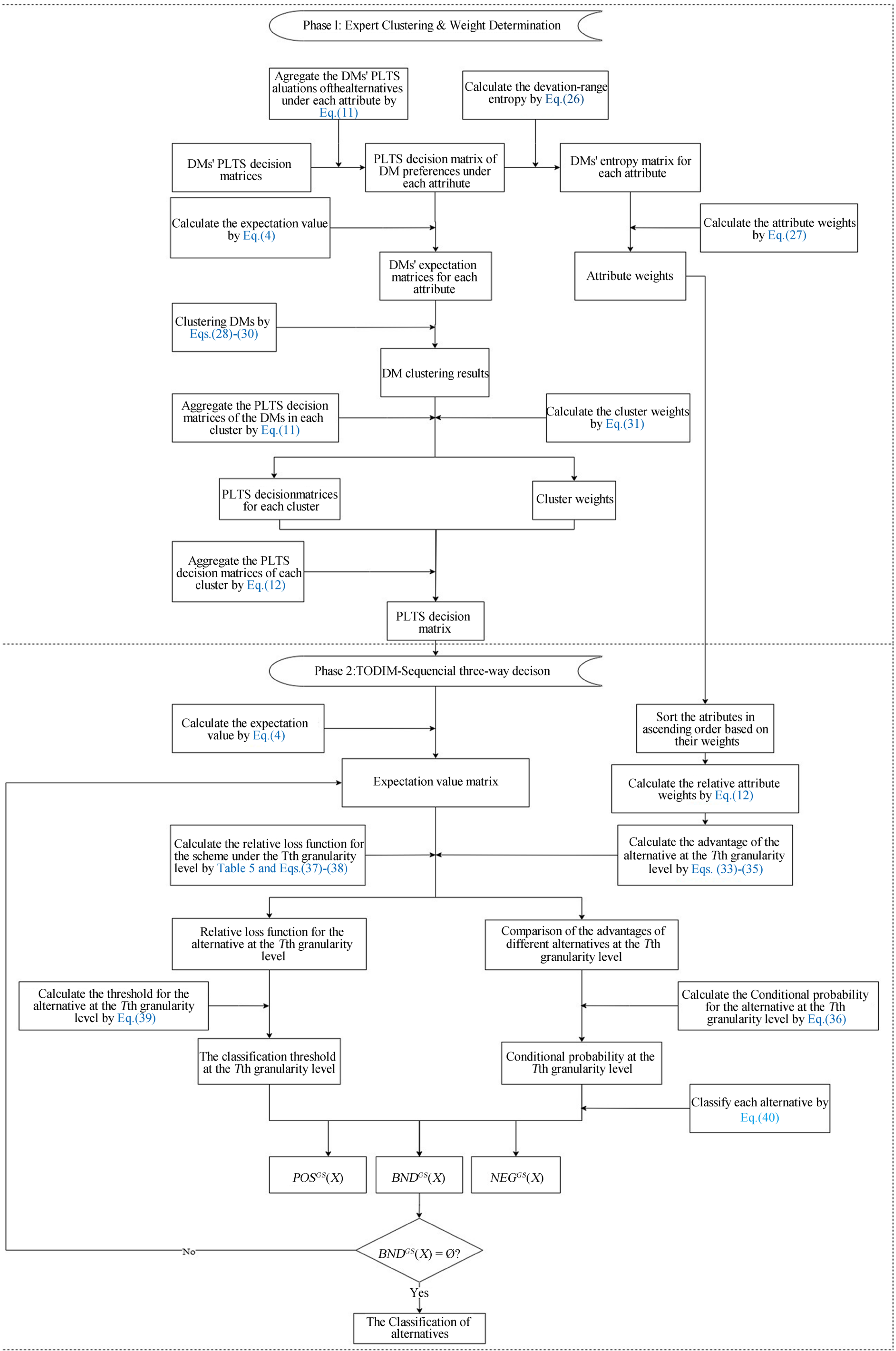


Figure 1. Decision-making process framework.

***Step* 9:** Aggregate the cluster evaluation matrices in a weighted manner to obtain the final decision matrix $F=\left[L_{ij}\left(p\right)\right]_{n\times m}$ by Equation (12).

***Step* 10:** Transform the final decision matrix $F=\left[L_{ij}\left(p\right)\right]_{n\times m}$ into the expectation score matrix $E=\left[E_{ij}\right]_{n\times m}$ by Equation (4).

***Step* 11:** Using the TODIM-STWD method described in Section 4.4, classify the alternatives.

The above decision-making process is depicted in Figure 1.

Algorithm 1 Deviation-range entropy-weight and FCM-based generation of the final decision matrix.

Input: Expert evaluation matrices $\{F^h=[L_{ij}^h(p)]_{n\times m}\}, h=1,2,\cdots,q$.
Output: The final decision matrix $F=[L_{ij}(p)]_{n\times m}$.
for all $A_i\in A, C_j\in C, e_h\in e$ **do**
  Aggregate the evaluation information of experts under each attribute according to Eq. (11), and obtain the expert–attribute preference matrix $\overline{F}=[\overline{L}_{hj}(p)]_{q\times m}$.
end for
for all $A_i\in A, C_j\in C, e_h\in e$ **do**
  Transform $\overline{F}$ into the entropy matrix $Q=[U_{hj}]_{q\times m}$ by Eq. (26).
end for
for all $C_j\in C, e_h\in e$ **do**
  if The weight of attribute W_j $(j=1,2,\cdots,m)$ is known, **then**
    Go to the next step.
  else
    Compute the attribute weight vector $\boldsymbol{W}=(W_1,W_2,\ldots,W_m)^T$ by Eq. (27).
  end if
end for
for all $C_j\in C, e_h\in e$ **do**
  Transform $\overline{F}$ into the expectation matrix $\overline{E}=[\overline{E}_{hj}]_{q\times m}$ by Eq. (4).
end for
for all $C_j\in C, e_h\in e$ **do**
  Compute the updated membership matrix M by Eq. (29).
  if $\|M^{T+1}-M^T\|<10^{-3}$, **then**
    Determine the final membership matrix.
  else
    Let $M^T=M^{T+1}$ and continue the iterative computation until the condition $\|M^{T+1}-M^T\|<10^{-3}$ is satisfied.
  end if
end for
for all $C_j\in C, e_h\in e$ **do**
  Partition the experts into $c=\sqrt{\frac{q}{2}}$ clusters by Eq. (30), and obtain the cluster set $G_1,G_2,\ldots,G_c$.
end for
for all $A_i\in A, C_j\in C, e_h\in e$ **do**
  Aggregate the expert evaluation matrices within each cluster by Eq. (11), and obtain the cluster evaluation matrix $F^{G_o}=[L_{ij}^{G_o}(p)]_{n\times m}, o=1,2,\ldots,c$.
end for
for all $G_o\in G, e_h\in e$ **do**
  Compute the weights of clusters $W^{G_1},W^{G_2},\ldots,W^{G_c}$ by Eq. (31).
end for
for all $A_i\in A, C_j\in C, G_o\in G$ **do**
  Aggregate the cluster evaluation matrices in a weighted manner by Eq. (12), and obtain the final decision matrix $F=[L_{ij}(p)]_{n\times m}$.
end for

Algorithm 2 TODIM-based sequential three-way decision process.

Input: Final decision matrix $F = [L_{ij}(p)]_{n\times m}$, attribute weight vector $\boldsymbol{W} = (W_1, W_2, \ldots, W_m)^T$, risk aversion coefficient η, and loss aversion coefficient θ.
Output: Classification and ranking results of alternatives.

for all $A_i \in A, C_j \in C$ **do**
    Transform the final decision matrix F into the expectation score matrix $E = [E_{ij}]_{n\times m}$ by Eq. (4).
end for
for all $C_j \in C$ **do**
    Sort the attributes in descending order according to their weights, and obtain the sequence $C^1, C^2, \ldots, C^m$.
end for
for all $A_i \in A$ **do**
    Initialize the set of unclassified alternatives as $\mathcal{A}^1 = \{A_1, A_2, \ldots, A_n\}$.
end for
for all $A_i \in \mathcal{A}$ **do**
    Let $\delta^0(A_i) = 0$.
    Let $\mu_{\alpha\beta}^{0_i} = 0$, where $\alpha \in \{P, B, N\}$ and $\beta \in \{P, N\}$.
end for
for $T = 1$ to m **do**
    Select the current-level attribute C^T and its weight W^T.
    for all $A_i \in \mathcal{A}^T$ **do**
        for all $A_k \in \mathcal{A}^T, k \neq i$ **do**
            Compute the current-level dominance degree $dom^T(A^{T_i}, A^{T_k})$ by Eq. (33).
        end for
        Compute the incremental overall dominance degree $\Delta\delta^T(A^{T_i})$ by Eq. (34).
        Compute the cumulative overall dominance degree $\delta^T(A^{T_i})$ by Eq. (35).
        Compute the conditional probability $Pr^T(X \mid [A^{T_i}])$ by Eq. (36).
        Determine the reference points $E^T_{\max}$ and $E^T_{\min}$ by Eq. (37).
        Compute the incremental relative losses $\Delta\mu_{\alpha\beta}^{T_i}$ according to Table 6.
        Compute the cumulative relative losses $\mu_{\alpha\beta}^{T_i}$ by Eq. (38).
        Compute the thresholds x^{T_i} and y^{T_i} by Eq. (39).
        if $Pr^T(X \mid A^{T_i}) \geq x^{T_i}$ **then**
            $A^{T_i} \in POS^{GS^T}(X)$.
        else if $y^{T_i} < Pr^T(X \mid A^{T_i}) < x^{T_i}$ **then**
            $A^{T_i} \in BND^{GS^T}(X)$.
        else
            $A^{T_i} \in NEG^{GS^T}(X)$.
        end if
    end for
    Let $\mathcal{A}^{(T+1)} = BND^{GS^T}(X)$.
    if $\mathcal{A}^{(T+1)} = \varnothing$ **then**
        break
    end if
end for

The algorithm of the FCM-TODIM-STWD method is depicted as Algorithm 1 and Algorithm 2.

6. Case Studies

In this section, a case study is provided to illustrate the application of the proposed method. Furthermore, comparative and experimental analyses are conducted to verify the validity and advantages of the proposed method.

Protecting the environment is everyone's responsibility. The air quality is an important index reflecting the environment quality and influences the people's health. Hence, the desire for a clear blue sky is a shared aspiration of humanity. This section applies the historical data for twenty cities in 2024 (https://www.aqistudy.cn/historydata/https://www.aqistudy.cn/historydata/), to evaluate the air quality of China. In this data set, twenty cities are regarded as DMs with equal weights. 366 days in a year are treated as alternatives (objects). Six attributes, including $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, and O_3, are chosen to reflect the air quality. Therefore, the evaluation of the air quality can be considered as a LGMADM problem. Generally, if the daily mean AQI is below 50, the air quality is regarded as satisfactory; otherwise, it is regarded as unsatisfactory. In three-way decision, X denotes the target state to be identified rather than a positively valued state. In this case, X is chosen as the unsatisfactory air-quality state because it is the state of DM concern. Under this setting, assigning an alternative to $POS(X)$ means accepting that the day belongs to the unsatisfactory air-quality state; assigning it to $NEG(X)$ means rejecting that it belongs to the unsatisfactory air-quality state, that is, judging it as satisfactory; and assigning it to $BND(X)$ means that the current evidence is insufficient for a definite judgment and further observation is needed. To clearly describe this MADM problem, some symbols are expressed as follows.

1) 366 days as an alternative set: $A = \{A_1, A_2, \cdots, A_{366}\}$;
2) the set of six attributes for rating air quality: $C = \{C_1, C_2, \cdots, C_6\}$;
3) twenty cities as a set of DMs: $E = \{E_1, E_2, \cdots, E_{20}\}$;
4) Attribute weights: $W = \{W_1, W_2, \cdots, W_6\}$ and $\sum_{j=1}^{m} W_j = 1$;
5) DMs' weights: $W_h = \left\{\frac{1}{20}, \frac{1}{20}, \cdots, \frac{1}{20}\right\}$;
6) target state $X = \{AQI \geq 50\}$;
7) complementary state $\neg X = \{AQI < 50\}$.

To conduct subsequent group decision-making analysis in the probabilistic linguistic environment, it is first necessary to map the real-valued evaluations of each attribute into linguistic terms. Let the linguistic term set be $S = \{s_{-3}, s_{-2}, s_{-1}, s_0, s_1, s_2, s_3\}$. Here, a larger subscript of the linguistic term indicates a better evaluation. Since all air quality indicators are cost-type attributes, that is, smaller attribute values imply better air quality, a reverse assignment strategy is adopted in the mapping process: smaller real values are mapped to larger linguistic terms, whereas larger real values are mapped to smaller linguistic terms. Specifically, for the j th attribute, all original evaluation values provided by all experts under this attribute are merged into an overall sample. Based on this overall sample, the six quantiles $\frac{1}{7}, \frac{2}{7}, \cdots, \frac{6}{7}$ are calculated and denoted by $u_j^1, u_j^2, \cdots, u_j^6$. Accordingly, the value range of the attribute can be divided into the following seven quantile intervals: $(-\infty, u_j^1]$, $(u_j^1, u_j^2]$, $(u_j^2, u_j^3]$, $(u_j^3, u_j^4]$, $(u_j^4, u_j^5]$, $(u_j^5, u_j^6]$, $(u_j^6, +\infty)$. For any real-valued observation x_{ij}^e, its linguistic mapping rule is defined as

$$L\left(x_{ij}^{e}\right)=\begin{cases} s_3(1), & x_{ij}^{e}\le u_j^1,\\ s_2(1), & u_j^1< x_{ij}^{e}\le u_j^2,\\ s_1(1), & u_j^2< x_{ij}^{e}\le u_j^3,\\ s_0(1), & u_j^3< x_{ij}^{e}\le u_j^4,\\ s_{-1}(1), & u_j^4< x_{ij}^{e}\le u_j^5,\\ s_{-2}(1), & u_j^5< x_{ij}^{e}\le u_j^6,\\ s_{-3}(1), & x_{ij}^{e}> u_j^6. \end{cases}$$

In this way, each real-valued evaluation of each alternative under each attribute provided by each expert can be transformed into a deterministic single probabilistic linguistic term set, that is, an evaluation value represented in the form of a probabilistic linguistic term set. This treatment has two main advantages. On the one hand, it maps continuous real-valued data into a discrete evaluation space, thereby improving the interpretability of the model. On the other hand, by partitioning the value range based on quantiles rather than fixed equal-width intervals, the influence caused by differences in attribute scales and uneven data distributions can be alleviated to a certain extent.

Table 7. Entropy matrix of expert e_h under each attribute C_j.

DMs	$PM_{2.5}$	PM_{10}	SO_2	CO	NO_2	O_3
1	0.6489	0.6412	0.0486	0.1367	0.2612	0.6428
2	0.4856	0.4936	0.1067	0.0862	0.1326	0.5669
3	0.4856	0.4936	0.1067	0.0862	0.1326	0.5669
4	0.5418	0.4875	0.6328	0.4382	0.3398	0.4622
5	0.5237	0.5129	0.3850	0.6667	0.6235	0.6735
6	0.5390	0.5422	0.8299	0.3400	0.5533	0.5803
7	0.5895	0.5512	0.5173	0.6008	0.4625	0.4931
8	0.5502	0.5517	0.6185	0.5578	0.5489	0.6052
9	0.5428	0.5240	0.2968	0.3415	0.3854	0.5035
10	0.5568	0.5501	0.4775	0.6126	0.6513	0.5679
11	0.4965	0.4882	0.7787	0.5394	0.6510	0.5782
12	0.6097	0.6147	0.8300	0.5320	0.4316	0.5080
13	0.2762	0.1970	0.2925	0.2865	0.1558	0.4423
14	0.4921	0.4756	0.5386	0.5672	0.5261	0.5137
15	0.5909	0.6042	0.8745	0.3861	0.6030	0.5861
16	0.4907	0.4785	0.5549	0.6080	0.4757	0.5938
17	0.5380	0.5138	0.6737	0.6126	0.5424	0.6635
18	0.4636	0.5159	0.5264	0.5514	0.4955	0.4924
19	0.6785	0.5643	0.5133	0.6470	0.5611	0.5990
20	0.4934	0.4159	0.6626	0.5433	0.5513	0.6635

Remark 3. It should be emphasized that the original observations are not assumed to be probabilistic linguistic assessments at the individual-data level. Instead, each raw attribute value is first transformed into a singleton linguistic term with probability 1 for normalization and comparability. The probabilistic linguistic information emerges in the subsequent aggregation stage: for the same alternative and attribute, different experts may provide different singleton linguistic terms, and the aggregation operators in Equations (11)-(12) combine these heterogeneous linguistic assessments into a probabilistic linguistic term set with multiple possible terms and associated probabilities. Therefore, the present case study is still a meaningful PLTS application, because the PLTS information is generated endogenously through group aggregation rather than being imposed directly on the raw data.

***Step* 1:** Construct the expert evaluation matrix $F^h = \left[L_{ij}^h\right]_{n\times m}$ according to the evaluation values provided by experts. Since there are 366 alternatives and 6 decision attributes, the resulting expert evaluation matrices $F^h = \left[L_{ij}^h\right]_{n\times m}$ are omitted here due to space limitations.

***Step* 2:** Obtain the comprehensive evaluation value of expert h under attribute j. Based on Equation (11), the comprehensive evaluation matrix of all experts under all attributes can be obtained as $\bar{F} = \left[\bar{L}_{hj}(p)\right]_{20\times 6}$.

***Step* 3:** According to Equation (26), the entropy matrix $Q = \left[U_{hj}\right]_{q\times m}$ can be obtained, as shown in Table 7.

***Step* 4:** According to Equation (27), the attribute weights $W_j\left(j = 1,2,\cdots,m\right)$ can be obtained, as shown in Table 8.

Table 8. Attribute weights.

W_1	W_2	W_3	W_4	W_5	W_6
0.1584	0.1647	0.1639	0.1828	0.1838	0.1464

Table 9. Expectation score function matrix of expert e_h under each attribute C_j.

DMs	$PM_{2.5}$	PM_{10}	SO_2	CO	NO_2	O_3
1	0.4791	0.4454	0.6995	0.9148	0.2441	0.5200
2	0.5815	0.6189	0.5314	0.8069	0.1753	0.4959
3	0.5815	0.6189	0.5314	0.8069	0.1753	0.4959
4	0.4613	0.4832	0.4850	0.4572	0.4180	0.5624
5	0.4208	0.4763	0.6280	0.3980	0.6935	0.4877
6	0.4872	0.3770	0.5196	0.2090	0.5168	0.4504
7	0.4872	0.6293	0.4763	0.4148	0.6393	0.5729
8	0.4531	0.5155	0.3707	0.5323	0.5546	0.4649
9	0.4918	0.5729	0.3957	0.7413	0.2295	0.4695

Continued

10	0.4722	0.4827	0.4394	0.4954	0.6516	0.4139
11	0.4085	0.3625	0.4308	0.5738	0.5137	0.4690
12	0.5369	0.4868	0.5460	0.3807	0.4945	0.4158
13	0.8497	0.8479	0.8329	0.5219	0.7978	0.6689
14	0.6526	0.7054	0.7222	0.6453	0.5738	0.6111
15	0.5114	0.4253	0.5209	0.2923	0.5961	0.4772
16	0.7099	0.7117	0.4731	0.5656	0.6453	0.6166
17	0.3443	0.3065	0.5806	0.3078	0.5747	0.4504
18	0.4162	0.4413	0.3224	0.4567	0.6753	0.4645
19	0.5132	0.4062	0.6239	0.3520	0.6152	0.5137
20	0.2978	0.2268	0.3643	0.5897	0.5383	0.4372

***Step* 5**: Based on Equation (4), the expectation score function matrix $E=\left[E_{hj}\right]_{q\times m}$ can be obtained, as shown in **Table 9**.

Table 10. Membership degrees of expert e_h to cluster G_o and the assigned cluster.

DMs	κ_{G_1}	κ_{G_2}	κ_{G_3}	G_o
1	0.1285	0.7688	0.1027	2
2	0.0187	0.9675	0.0138	2
3	0.0187	0.9675	0.0138	2
4	0.4982	0.1114	0.3904	1
5	0.2807	0.0364	0.6829	3
6	0.2358	0.0560	0.7082	3
7	0.5759	0.0540	0.3701	1
8	0.8348	0.0394	0.1258	1
9	0.1071	0.8241	0.0687	2
10	0.7914	0.0246	0.1840	1
11	0.6169	0.1046	0.2785	1
12	0.2836	0.0444	0.6720	3
13	0.3739	0.2487	0.3774	3
14	0.3837	0.3049	0.3114	1
15	0.1245	0.0166	0.8589	3
16	0.4988	0.1898	0.3114	1
17	0.2661	0.0600	0.6738	3
18	0.6534	0.0538	0.2928	1
19	0.1056	0.0163	0.8780	3
20	0.4846	0.1687	0.3467	1

***Step* 6:** Determine the number of clusters according to the heuristic rule $c=\sqrt{\frac{q}{2}}$ [25]. For $q=20$, have $c\approx 2.236$; thus, $c=3$ is used in the present case study. According to Equation (29), the membership degrees of experts to clusters $G_o\left(o=1,2,3\right)$, denoted by $\kappa_{G_o}\left(e_h\right)$, as well as their assigned clusters, can be obtained, as shown in Table 10.

***Step* 7:** Aggregate the evaluation information of experts within each cluster to obtain the cluster evaluation matrix $F^{G_o}=\left[L_{ij}^{G_o}\left(p\right)\right]_{n\times m}$ by Equation (11).

***Step* 8:** Calculate the cluster weights $W^{G_o}, o=1,2,\cdots,c$ by Equation (31), as shown in Table 11.

Table 11. Cluster weights of G_o.

	G_1	G_2	G_3
W^{G_o}	0.3891	0.2572	0.3537

***Step* 9:** Aggregate the cluster evaluation information in a weighted manner to obtain the final decision matrix $F=\left[L_{ij}\left(p\right)\right]_{n\times m}$ by Equation (12).

***Step* 10:** Transform the final decision matrix $F=\left[L_{ij}\left(p\right)\right]_{n\times m}$ into the expectation score function matrix $E=\left[E_{ij}\right]_{n\times m}$ by Equation (4).

***Step* 11:** Use the TODIM-STWD method described in Section 4.4 to classify the alternatives.

***Step* 11-1:** According to Table 8, the weight of the reference attribute is $W_r=\max W_j=0.1838$. Then, according to Equation (32), the relative weights W_{jr} of attributes C_j with respect to the reference attribute C_r can be calculated, as shown in Table 12.

Table 12. Relative weights of attributes C_j.

W_1	W_2	W_3	W_4	W_5	W_6
0.8618	0.8961	0.8917	0.9946	1	0.7965

***Step* 11-2:** According to Table 8, $\max W_j=0.1838$. Therefore, the first granularity structure GS^1 performs decision-making based on attribute C_5. The subsequent granularity levels are carried out according to the attribute order C_4, C_2, C_3, C_1, and C_6, respectively.

***Step* 11-3:** Calculate the dominance degree of alternative A^{T_i} over alternative A^{T_k} under attribute C_5 by Equation (33).

***Step* 11-4:** Compute the incremental overall dominance degree under attribute C_5 by Equation (34).

***Step* 11-5:** Calculate the cumulative overall advantage of scheme A^{T_i} relative to all other schemes under attribute C_5 using Equation (35).

***Step* 11-6:** Calculate the conditional probability $Pr^T\left(X\mid A^{T_i}\right)$ of alternative

A^{T_i} under attribute C_5 by Equation (36).

Step **11-7:** Determine the reference points $E_{\max}^{T}$ and $E_{\min}^{T}$ under attribute C_5 by Equation (37).

Step **11-8:** Calculate the incremental dynamic aggregation relative loss function for scheme A^{T_i} under attribute C_5 based on Table 6.

Step **11-9:** Compute the cumulative dynamic aggregated relative loss function of scheme A^{T_i} under attribute C_5 by Equation (38).

Step **11-10:** Calculate the thresholds x^{T_i} and y^{T_i} of alternative A^{T_i} under attribute C_5 by Equation (39).

Step **11-11:** Classify alternative A^{T_i} under attribute C_5 by Equation (40).

Step **11-12:** If $BND^{GS^T}(X) \neq \varnothing$, then the alternatives in $BND^{GS^T}(X)$ are further classified and ranked at the next granularity level, and ***Steps*** **11-3-11-12** are repeated until $BND^{GS^T}(X) = \varnothing$.

Due to space limitations, some intermediate computational results are omitted from the main text. The proposed algorithm completes the classification and ranking of 366 alternatives after four iterations, indicating that the final results can be obtained by considering only four attributes, namely NO_2, CO, PM_{10}, and SO_2. Specifically, under attribute NO_2, all 366 alternatives are processed, and 34 alternatives enter the boundary region; under attribute CO, these 34 alternatives are further processed, and 8 alternatives remain in the boundary region; under attribute PM_{10}, these 8 alternatives are processed, and only 1 alternative remains in the boundary region; finally, under attribute SO_2, the remaining alternative is further processed and the boundary region becomes empty. At this point, the algorithm terminates.

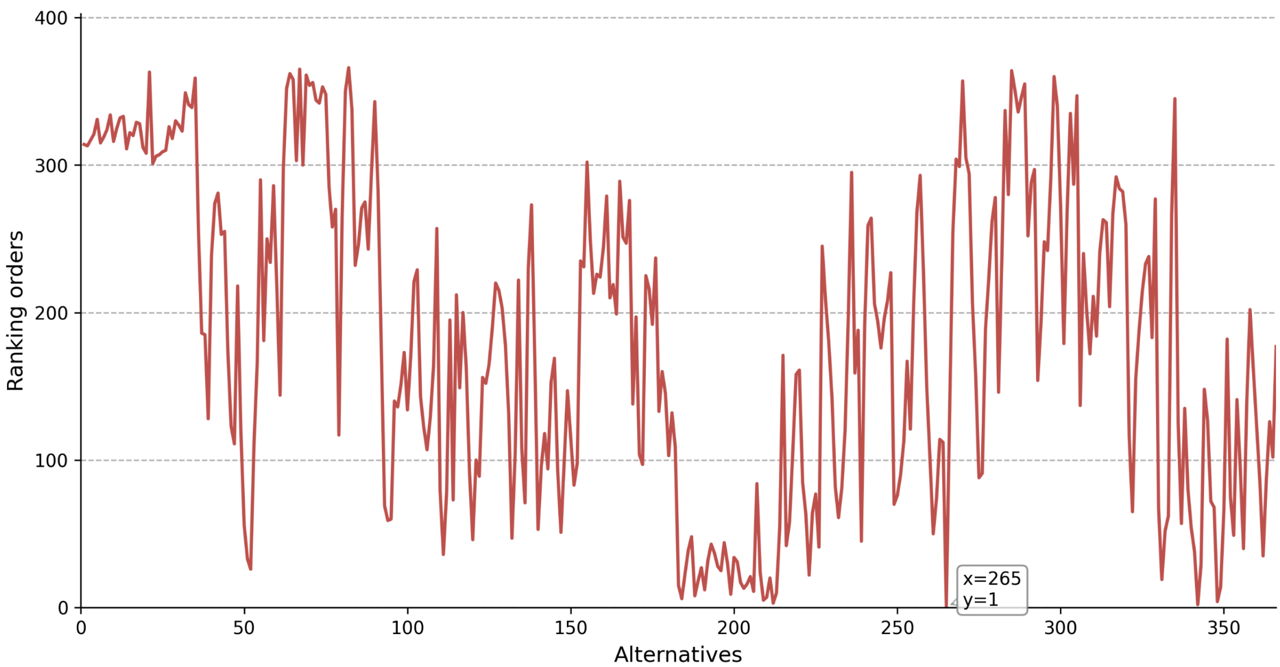


Figure 2. Rank results for air quality.

As shown in **Figure 2**, the horizontal axis represents the date, and the vertical axis represents the daily ranking of air quality, in which alternative A_{265} is ranked first. As shown in **Figure 3**, among the 366 days, 335 days are classified into the acceptance region, while 31 days are classified into the rejection region. These results indicate that the air quality in China in 2024 exhibits a relatively consistent classification tendency on most days. Combined with the classification results, it can be further seen that the issue of air quality still deserves continuous attention.

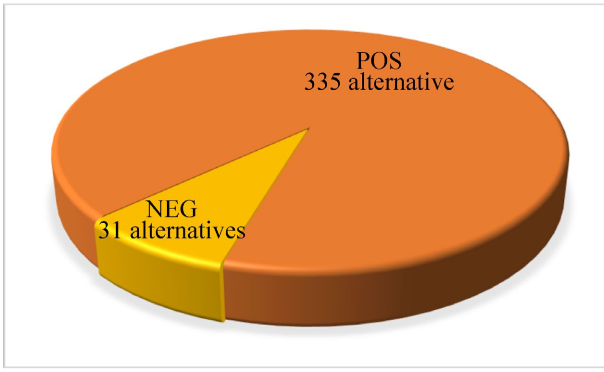


Figure 3. Classification results for air quality.

6.1. Comparative Analysis and Discussion

To verify the effectiveness and superiority of the proposed method in solving large group multi-attribute decision-making problems, comparative analyses are conducted in a multi-attribute decision-making environment by comparing the proposed method with several classical LGMADM methods as well as some three-way decision models.

6.1.1. Comparative Analysis with Classical LGMADM Methods

In the present case study, the FCM-TODIM-STWD method finally classifies all alternatives into two decision regions, thereby providing a new perspective for solving large-group multi-attribute decision-making problems. However, classical LGMADM methods usually can only generate ranking results and cannot further provide classifications of alternatives. Therefore, based on the AQI dataset in Section 6, three classical methods, namely TOPSIS [5], PROMETHEE [41], and MULTIMOORA [42], are selected to conduct a Spearman rank correlation coefficient (SRCC) [43] analysis between their ranking results and those of the proposed method. The results are shown in **Figure 4**. Combined with **Figure 4** and **Figure 5**, the feasibility of the proposed method can be verified.

1) The optimal alternative obtained by the proposed method is consistent with those identified by existing classical methods. As shown in **Figure 4**, the best alternative obtained by the proposed method is A_{265}, which is the same as the results produced by TOPSIS [5], PROMETHEE [41], and MULTIMOORA [42]. This indicates that the proposed method is feasible in terms of identifying the op-

timal alternative. In addition, compared with the above three classical LGMADM methods, the FCM-TODIM-STWD method can not only provide ranking results for alternatives, but also classify them into different decision regions. In this way, the ranking results can be further interpreted, and richer decision-support information can be provided to decision makers. This demonstrates that the proposed FCM-TODIM-STWD method has stronger interpretability.

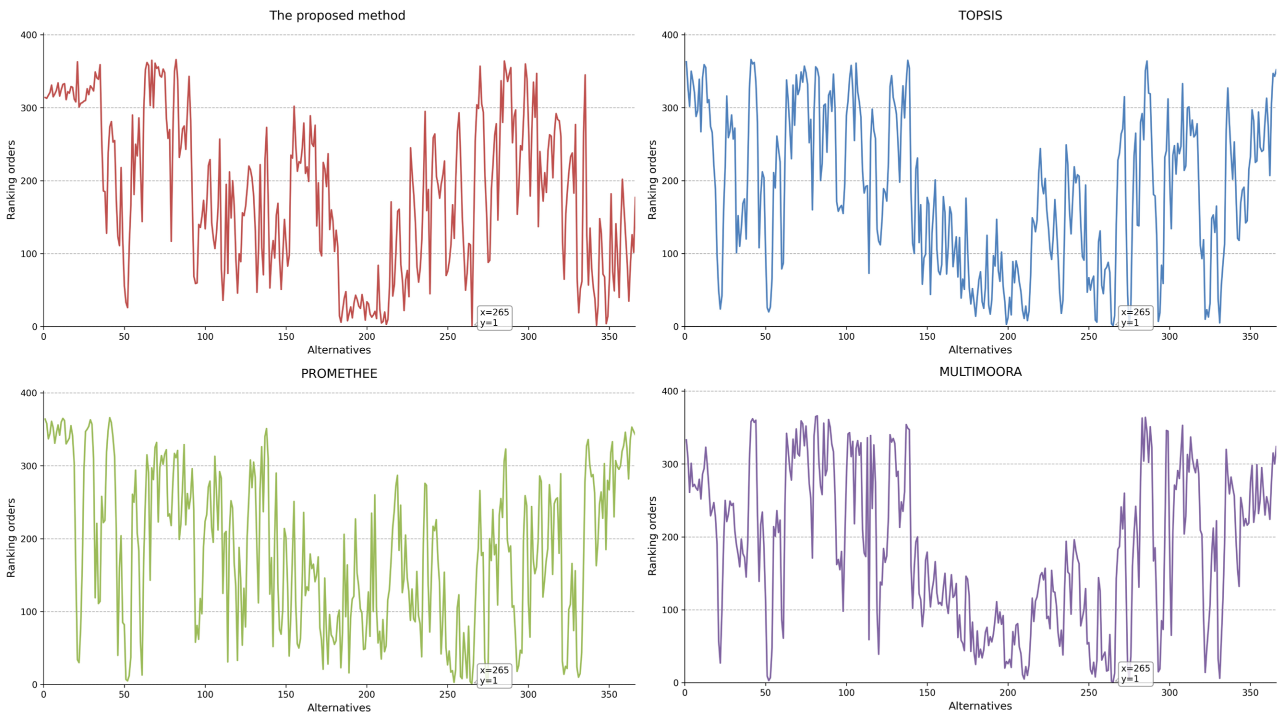


Figure 4. Ranking results of alternatives by different methods.

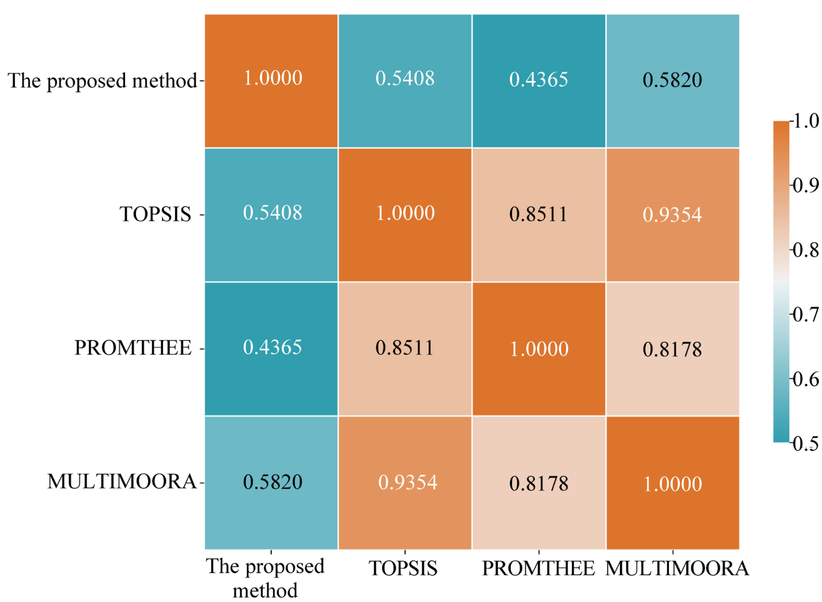


Figure 5. SRCCs between ranking results of different methods.

2) According to the threshold table of Spearman rank correlation, when the sample size is greater than 100 and the significance level of the two-tailed test is 0.01, if the correlation coefficient between two ranking results is greater than 0.257, the two samples can be regarded as highly correlated [44]. As shown in **Figure 5**, the Spearman rank correlation coefficients between the proposed method and TOPSIS [5], PROMETHEE [41], and MULTIMOORA [42] are all greater than 0.4. This indicates that the ranking results of the proposed method are statistically highly consistent with those of the classical methods, which further verifies the feasibility of the proposed method.

Based on the above comparative results, it can be concluded that the proposed method is reasonable and effective. Furthermore, from the perspective of methodological mechanisms, the proposed FCM-TODIM-STWD method differs from classical LGMADM methods in the following aspects.

1) The proposed FCM-TODIM-STWD method can not only generate ranking results, but also classify alternatives into different decision regions, thereby providing a more interpretable explanation of the ranking results and offering richer decision information to decision makers. In contrast, TOPSIS [5], PROMETHEE [41], and MULTIMOORA [42] usually only provide ranking results and cannot further reveal the decision region to which each alternative belongs.

2) The proposed FCM-TODIM-STWD method can incorporate the loss aversion coefficient in prospect theory to characterize the psychological behavioral features of decision makers, thereby reflecting decision preferences more realistically and providing decision information that is more consistent with real decision-making situations. By contrast, TOPSIS [5], PROMETHEE [41], and MULTIMOORA [42] mainly rank alternatives based on criterion values and do not explicitly consider the psychological behavioral factors of decision makers.

3) The proposed FCM-TODIM-STWD method introduces a sequential three-way decision mechanism. By performing discrimination step by step according to the attribute order, the method can preliminarily screen alternatives using early-stage attribute information, thereby effectively reducing information acquisition costs. Meanwhile, it provides a deferred-decision mechanism for alternatives in the boundary region, which helps avoid premature judgments under insufficient information. In contrast, TOPSIS [5], PROMETHEE [41], and MULTIMOORA [42] generally require all attribute information to be used at one time for computation, which not only leads to higher information acquisition costs but also lacks flexibility in handling uncertain situations.

4) The proposed FCM-TODIM-STWD method determines attribute weights based on the constructed entropy function, which can reduce the influence of subjective weighting to a certain extent. In contrast, MULTIMOORA [42] adopts a subjective weighting scheme; TOPSIS [5] determines attribute weights based on the idea of distance maximization; and PROMETHEE [41] also relies on distance information, determining attribute weights by comparing the proportion of the distances under a given attribute to the distances under all attributes. Compared

with such distance-based weighting methods, which usually require the calculation of distances among multiple alternatives under multiple attributes and are therefore relatively cumbersome, the weighting method proposed in this paper has the advantages of reduced subjectivity and relative computational simplicity.

6.1.2. Comparative Analysis of Several Three-Way Decision Models in a Multi-Attribute Environment

In this paper, a FCM-TODIM-STWD model is established for solving large group multi-attribute decision-making problems by integrating three-way decision theory with a sequential strategy. In fact, three-way decision theory, as an emerging technique for dealing with decision-making problems, has attracted increasing attention in recent years. Many scholars have proposed various three-way decision methods in different linguistic environments and successfully applied them to multi-attribute decision-making problems or large group multi-attribute decision-making problems. For example, Han *et al.* [19] first defined the relative loss function in the probabilistic linguistic term environment and proposed a three-way multi-attribute decision-making method under probabilistic linguistic information. Subsequently, Yang *et al.* [27] proposed a three-way multi-attribute decision-making model in a weak probabilistic linguistic term environment. Meanwhile, Jiang and Hu [12] combined different linguistic environments with sequential three-way decision theory and proposed a sequential three-way decision-making method for large group multi-attribute decision-making with heterogeneous information. In addition, Xiao *et al.* [45] proposed a sequential three-way decision method for large group multi-attribute decision-making based on interval-valued information, while Song *et al.* [46] constructed a sequential three-way decision model for large group multi-attribute decision-making in the interval-valued intuitionistic fuzzy environment.

Table 13. Comparison results between different methods.

Methods	Lang. env	Attribute weights	Conditional probability	Psych. behav	Multiple experts	Sequential
The proposed method	PLTS	Deviation–range entropy	TODIM	Yes	Yes	Yes
Han *et al.*'s method [19]	PLTS	Information entropy	Equivalence class	Yes	No	No
Yang *et al.*'s method [27]	WPLTS	Information entropy	Equivalence class	No	No	No
Jiang and Hu's method [12]	Heter.	Deviation from reference point	Cost function	No	Yes	Yes
Xiao *et al.*'s method [45]	Interval	Information entropy	No	Yes	Yes	Yes
Song *et al.*'s method [46]	IVIF	Information entropy;	Grey relational analysis;	Yes	Yes	Yes
		coefficient of variation	VIKOR	Yes	Yes	Yes
Han and Zhan's method [15]	PLTS	Attribute discriminant value	-	Yes	Yes	Yes
Chen and Yin's method [24]	PLTS	PL-BWM	Improved TODIM	Yes	No	No

Notes: PLTS = probabilistic linguistic term set; WPLTS = weak probabilistic linguistic term set; Heter. = heterogeneous information; Interval = interval-valued information; IVIF = interval-valued intuitionistic fuzzy; "-" means that this item is not explicitly emphasized in the main framework description.

However, the above methods cannot be directly applied to solve the large group multi-attribute sequential three-way decision-making problem in the probabilistic linguistic term environment described in Section 6. To further demonstrate the advantages of the established FCM-TODIM-STWD model, the main characteristics of the proposed model and seven related methods are compared, as shown in **Table 13**.

As can be seen from **Table 13**, the proposed FCM-TODIM-STWD model provides a relatively comprehensive framework for large-group multi-attribute decision-making in the PLTS environment. Compared with the methods of Han *et al.* [19] and Yang *et al.* [27], the proposed model simultaneously incorporates multiple experts and a sequential strategy into the decision process, thereby enabling more sufficient use of decision information. Compared with the methods of Jiang and Hu [12], Xiao *et al.* [45], and Song *et al.* [46], which are developed in heterogeneous, interval-valued, or interval-valued intuitionistic fuzzy environments, the proposed method can directly deal with large-group multi-attribute decision-making problems in the PLTS environment. In addition, compared with Han and Zhan's method [15] and Chen and Yin's method [24], the proposed model further emphasizes expert clustering and attribute-driven sequential screening, and thus is more suitable for the present large-group multi-attribute classification task. Therefore, the proposed method can be regarded as a unified decision-making framework that integrates probabilistic linguistic representation, behavioral evaluation, multi-expert aggregation, and sequential three-way classification.

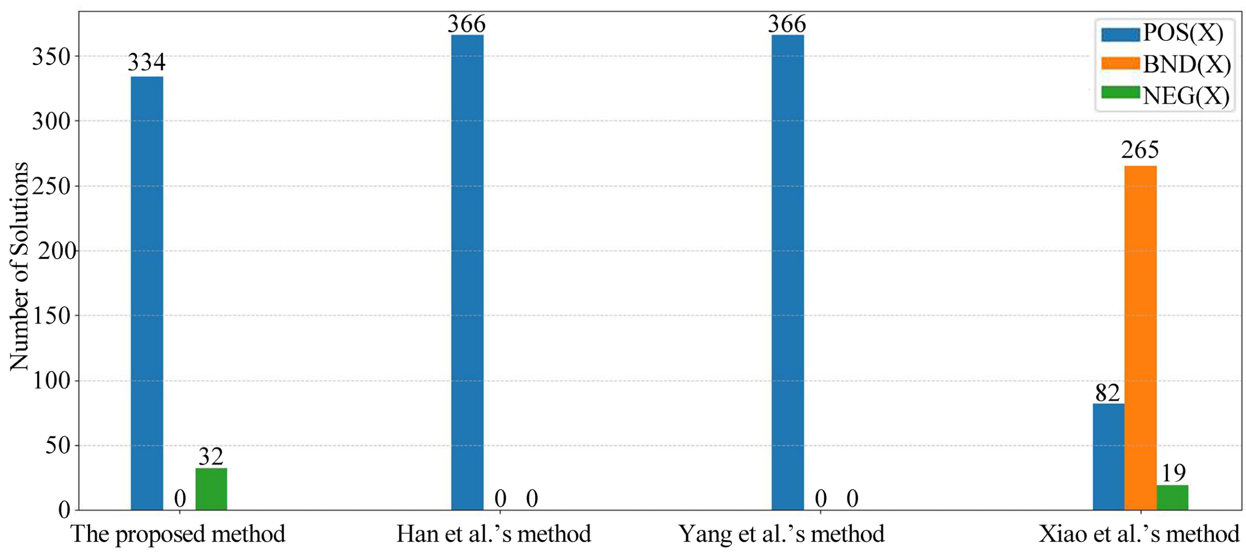


Figure 6. Classification results by different methods.

For the quantitative comparison, only the methods of Han *et al.* [19], Yang *et al.* [27], and Song *et al.* [46] are selected as numerical baselines. The main reason is that these three methods are sufficiently representative and relatively comparable to the proposed model from three complementary perspectives. Specifically, Han *et al.*'s method is chosen as a direct PLTS-based three-way decision baseline; Yang *et al.*'s method is included as a closely related probabilistic-linguistic three-way decision model; and Song *et al.*'s method is further selected as a representative sequential group three-way decision approach. By contrast, the methods of Jiang and Hu [12] and Xiao *et al.* [45] are developed for substantially different information environments and decision settings, Han and Zhan's method [15] mainly focuses on group consensus, and Chen and Yin's method [24] does not provide a directly comparable large-group sequential classification framework. Therefore, these methods are retained in the qualitative comparison of **Table 13**, but are not included in the full quantitative experiment. The classification results of the proposed method and the above three baseline methods are shown in **Figure 6**.

According to the AQI dataset, the classification results by different methods are depicted in **Figure 6**. It is observed from **Figure 6** that the classifications clearly vary with different methods. Hence, how to judge the quality of classifications is a key issue. Precision rate (*PR*), recall rate (*RR*) and error rate (*ER*) [47] are popular indices for assessing the classification performances, *i.e.*,

$$PR = \frac{X \to P}{X \to P + \neg X \to P} \times 100\%, \tag{41}$$

$$RR = \frac{X \to P}{X \to P + X \to N} \times 100\%, \tag{42}$$

$$ER = \frac{X \to N + \neg X \to P}{K} \times 100\%, \tag{43}$$

where, $X \to P$ expresses the number of alternatives in the set X that are assigned to the *POS* domain, $\neg X \to P$ represents the number of alternatives in the set $\neg X$ that are assigned to the *POS* domain, $X \to N$ denotes the number of alternatives in the set X that are assigned to the *NEG* domain, and K refers to the total number of alternatives. The results are listed in **Table 14** and visualized in **Figure 7**.

Remark 4. For evaluation purposes, a reference state partition is first constructed according to the daily mean AQI over the 20 cities. Specifically, if the daily mean AQI of a given day is no less than 50, the day is assigned to the target state X; otherwise, it is assigned to the complementary state $\neg X$. On this basis, the indices *PR*, *RR*, and *ER* are defined from the cross-counts between the reference state partition $(X, \neg X)$ and the three-way classification results (*POS*, *BND*, *NEG*), rather than from a standard binary confusion matrix. Therefore, these indices are used to evaluate the consistency between the predefined state partition and the resulting three-way classification.

Remark 5. To avoid confusion, the quantitative comparison is carried out from two aspects. Ranking consistency is assessed by the best alternative and SRCC, whereas *PR*, *RR*, and *ER* are used only to evaluate the consistency between the reference state partition and the resulting three-way classification.

Based on the results shown in **Figure 6** and **Table 14**, the following observations can be obtained.

Table 14. *PR*, *RR* and *ER* for different classification methods.

Methods	*PR*	*RR*	*ER*
The proposed method	92.22%	96.25%	10.38%
Han *et al.*'s method [19]	87.43%	100%	12.57%
Yang *et al.*'s method [27]	87.43%	100%	12.57%
Song *et al.*'s method [46]	82.93%	78.16%	9.02%

1) As shown in **Figure 6**, the proposed FCM-TODIM-STWD method finally classifies all air quality results into two decision regions, namely the acceptance region and the rejection region, by using a sequential strategy, whereas the method of Song *et al.* [46] divides the results into three regions, namely the acceptance region, the boundary region, and the rejection region. In general, the more alternatives are assigned to the boundary region, the higher the uncertainty in the decision results. Compared with the method of Song *et al.* [46], the proposed method produces an empty boundary region in the final stage, which indicates that the FCM-TODIM-STWD method can effectively reduce uncertainty in the decision-making process through the sequential strategy.

2) According to Equations. (41)-(43), a method with better classification performance should generally have higher precision and recall rates and a lower *ER*. As shown in **Table 14** and **Figure 7**, the proposed method achieves the highest *PR* among the compared methods. Its *RR* is slightly lower than those of Han *et al.* [19] and Yang *et al.* [27], but its *ER* is lower than theirs. By contrast, although Song *et al.*'s [46] method yields the lowest *ER*, it also shows the lowest *PR* and *RR* and retains a nonempty boundary region. Therefore, its lower *ER* should be interpreted together with its lower *PR*, lower *RR*, and the retention of a nonempty boundary region. Overall, when *PR*, *RR*, and *ER* are considered jointly, the proposed FCM-TODIM-STWD method provides the most balanced classification performance among the compared methods.

In summary, based on the air quality dataset, this subsection has compared the proposed FCM-TODIM-STWD method with several related methods from both qualitative and quantitative perspectives, thereby demonstrating its effectiveness and advantages. In fact, with the continuous emergence of various complex decision-making scenarios, emergency events, and increasing decision costs, how to make reasonable decisions within a shorter time and at a lower decision cost has become an important issue for decision-making departments. The FCM-TODIM-STWD method proposed in this paper incorporates the psychological behavioral factors of decision makers and further introduces a sequential three-way decision mechanism, with the aim of reducing the amount of attribute information required in the decision process and lowering the evaluation cost of experts. However, such a multi-stage sequential decision strategy has not been sufficiently considered in previous large group multi-attribute decision-making methods and practices. Therefore, the present study not only yields a conclusion of theoretical significance, but also provides a decision-making method with practical application value.

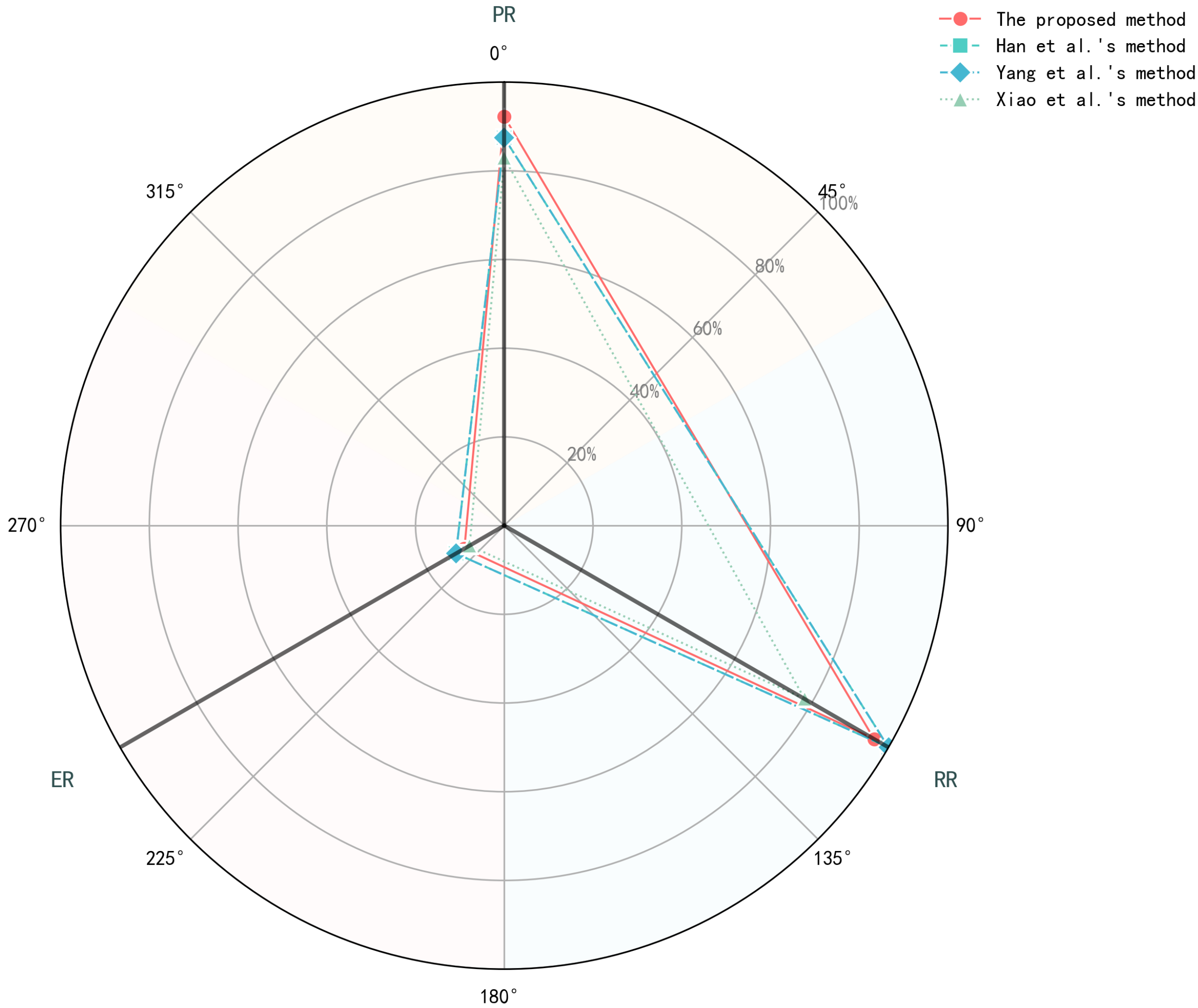


Figure 7. PR, RR and ER for different classification methods.

7. Experimental Analysis

Section 6 verified the effectiveness of the proposed method based on the 2024 AQI dataset from 20 cities. In this section, several experiments are further designed to evaluate the performance of the proposed method. Section 7.1 conducts a sensitivity analysis of parameter θ in the TODIM method, while Section 7.2 investigates how parameter η affects the decision results. In addition, ablation experiments are designed in this section to explore the influence of psychological behavior and sequential three-way decision on the decision results.

7.1. Sensitivity Analysis of Parameter θ in TODIM

In the TODIM method, the loss aversion coefficient θ is an important parameter satisfying $\theta > 0$. It can be seen from Equation (33) that when $\theta > 1$, the degree of loss aversion decreases, whereas when $0 < \theta < 1$, the degree of loss aversion increases. Therefore, this paper further investigates the effect of the loss aversion coefficient θ on the ranking and classification results. The dataset used is still the air quality dataset of 20 cities in 2024. Under the condition that all other parameters remain fixed, the values of θ are set to 2, 3, 5, 10, 50, and 100, respectively. The corresponding best alternatives and classification results are shown in Table 15 and Figure 8. In addition, Figure 8 also reports the precision rate, recall rate, error rate, and the number of granularity levels required for sequential decision-making under different values of θ.

Table 15. *PR*, *RR* and *ER* for different θ.

The values of θ	The best alternatives	*PR*	*RR*	*ER*	*T*
$\theta = 2$	A_{265}	92.22%	96.25%	10.38%	4
$\theta = 3$	A_{265}	92.12%	95%	11.48%	4
$\theta = 5$	A_{265}	91.95%	92.81%	13.39%	5
$\theta = 10$	A_{265}	91.42%	86.56%	18.85%	5
$\theta = 50$	A_{265}	90.61%	69.38%	33.06%	4
$\theta = 100$	A_{265}	90.87%	65.31%	36.07%	5

It can be observed from Table 15 that different values of θ do not affect the selection of the best alternative, and the optimal alternative remains A_{265} throughout. However, the classification results are significantly influenced by the value of θ. As shown in Figure 9, as θ increases, the number of alternatives in the positive region gradually decreases, whereas the number of alternatives in the negative region correspondingly increases. Meanwhile, with the increase of θ, the precision rate shows an overall decreasing trend, the recall rate decreases significantly, the error rate rises markedly, and the number of granularity levels required for decision-making varies between four and five.

In essence, parameter θ reflects the attenuation degree of decision makers when facing loss-related information. As θ increases, the sensitivity of decision makers to negative information gradually weakens, meaning that unfavorable differences in local attributes are no longer excessively amplified. However, this does not imply that alternatives therefore become more likely to be accepted. In practical decision-making scenarios, a smaller θ usually corresponds to a more cautious comparison strategy that is more sensitive to losses, whereas a larger θ makes decision makers more inclined to screen alternatives directly from the perspective of overall competitiveness. For those alternatives whose overall performance is insufficient but that might originally retain some room for consideration because of stronger loss sensitivity, this room will shrink as θ increases, and such alternatives will therefore be more likely to be assigned to the rejection region.

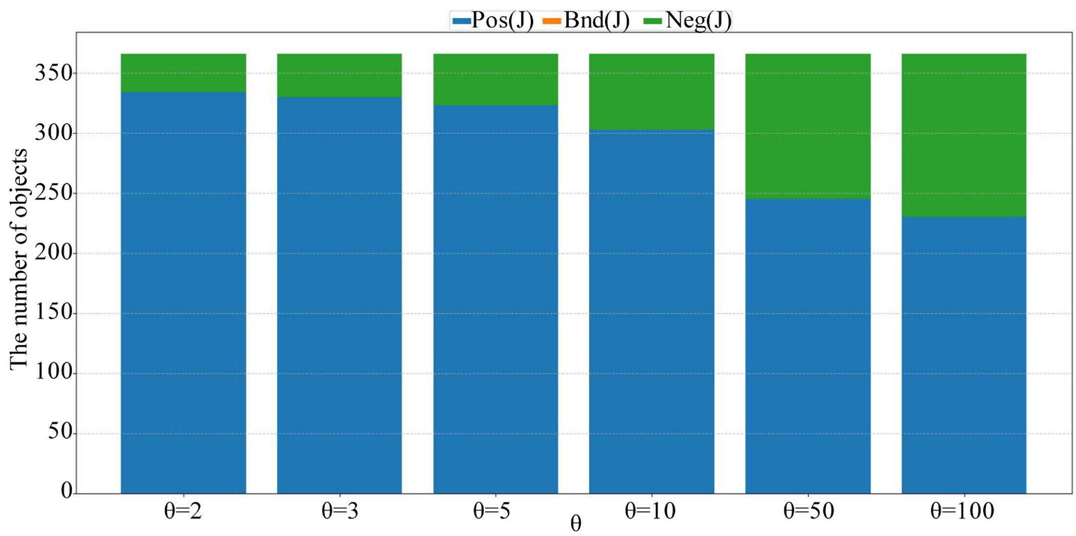


Figure 8. Comparison of classification results for different values of θ.

7.2. Sensitivity Analysis of Parameter η in the Relative Loss Function

As shown in **Table 6**, the risk aversion coefficient η needs to be introduced when transforming score values into the relative loss function, and this parameter is used to characterize the degree of risk aversion of decision makers. Given that $\eta \in (0, 0.5)$, this paper considers $\eta = 0.05, 0.10, \cdots, 0.45$ while keeping all other parameters fixed, and then examines the influence of η on the ranking and classification results.

Figure 9 and **Table 16** show that when η takes different values, the ranking results remain highly stable overall, and the optimal alternative consistently remains A_{265}. This indicates that the FCM-TODIM-STWD method is not highly sensitive to parameter η in terms of ranking results. In practical decision-making, even if decision makers exhibit different degrees of risk aversion, the identification of the best alternative remains relatively stable.

As can be observed from **Table 16** and **Figure 10**, as the value of η increases, the number of objects assigned to the acceptance region gradually increases, whereas the numbers of objects assigned to the rejection region and boundary region decrease correspondingly. Meanwhile, the precision rate and recall rate show an overall increasing trend, the error rate decreases significantly, and the number of granularity levels required for sequential decision-making is reduced from 6 to 2. This indicates that a larger η is conducive to improving classification efficiency and reducing the hierarchical cost of the decision-making process.

In essence, parameter η controls the adjustment degree of the relative loss function in the boundary region, and its variation directly affects the allocation of objects among the acceptance region, rejection region, and boundary region. As η increases, the relative losses associated with assigning objects to the rejection region and the boundary region also increase. According to the principle of loss minimization, decision makers tend to assign more objects to the acceptance region, thereby reducing the numbers of objects in the rejection and boundary regions. Therefore, the above experimental results are consistent with the underlying decision-making mechanism.

Figure 9. SRCCs between ranking results of different values of η.

Table 16. PR, RR and ER for different η.

The values of η	The best alternatives	*PR*	*RR*	*ER*	*T*
$\eta = 0.05$	A_{265}	89.53%	62.18%	33.88%	6
$\eta = 0.1$	A_{265}	89.83%	67.95%	33.88%	6
$\eta = 0.15$	A_{265}	90.23%	75%	28.96%	6
$\eta = 0.2$	A_{265}	90.75%	82.81%	22.4%	6
$\eta = 0.25$	A_{265}	91.09%	86.25%	19.4%	4
$\eta = 0.3$	A_{265}	92.22%	96.25%	10.38%	4
$\eta = 0.35$	A_{265}	91.98%	93.12%	13.11%	4
$\eta = 0.4$	A_{265}	92.17%	95.62%	10.93%	3
$\eta = 0.45$	A_{265}	92.19%	95.94%	10.66%	2

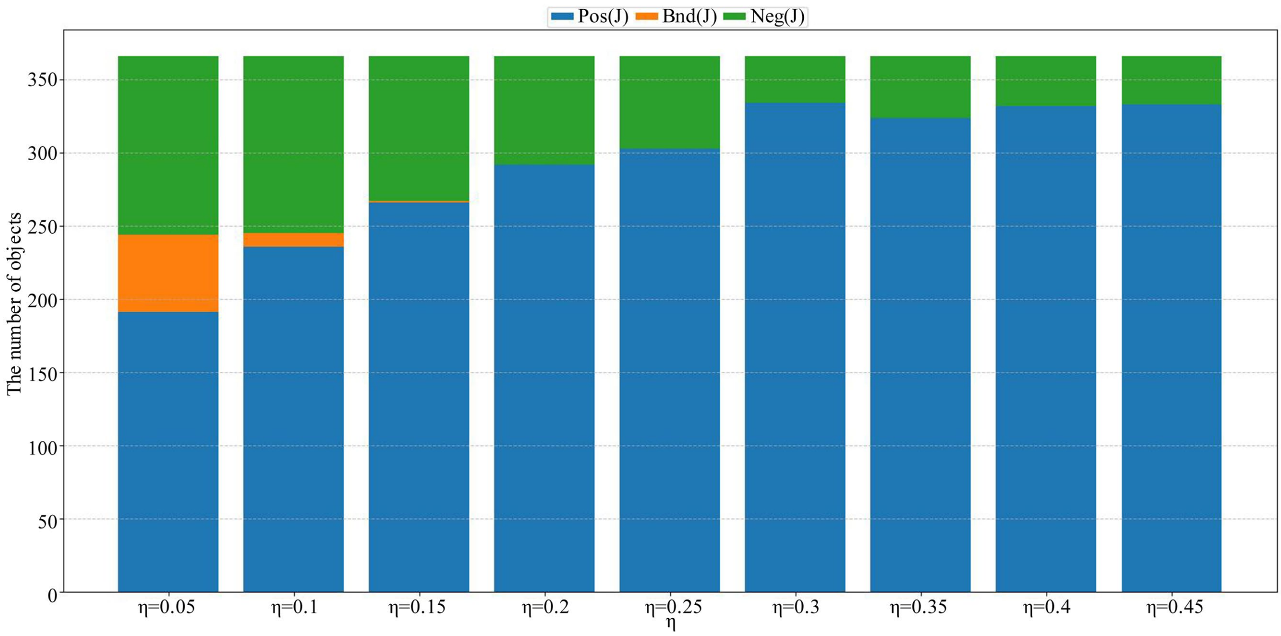


Figure 10. Comparison of classification results for different values of η.

7.3. Sensitivity Analysis of the Cluster Number

In the proposed FCM-TODIM-STWD method, the cluster number plays an important role in the expert partition stage, since it affects the construction of cluster-level evaluation information and may further influence the final ranking and three-way classification results. In the baseline setting of this study, the cluster number is determined according to the heuristic rule described in Section 4.3. To further examine whether the clustering stage is overly dependent on a single heuristic setting, a sensitivity analysis with respect to the cluster number is conducted.

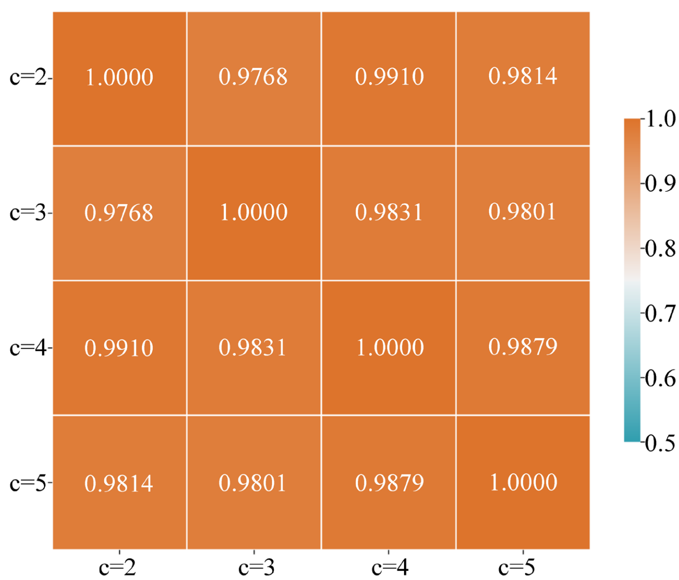


Figure 11. SRCCs between ranking results of different values of c.

Specifically, while keeping all the other parameters unchanged, the values of the cluster number are set to $c = 2,3,4,5$. The corresponding classification results are reported in **Table 17**, and the SRCCs between the ranking results obtained under different values of c are shown in **Figure 11**.

Table 17. *PR*, *RR* and *ER* for different c.

The values of c	The best alternatives	*PR*	*RR*	*ER*	*T*
$c = 2$	A_{265}	92.17%	95.62%	10.93%	4
$c = 3$	A_{265}	92.22%	96.25%	10.38%	4
$c = 4$	A_{265}	92.22%	96.25%	10.38%	4
$c = 5$	A_{265}	92.22%	96.25%	10.38%	4

As shown in **Figure 11**, the ranking results obtained under nearby values of c are highly consistent. In particular, all SRCC values are greater than 0.9, indicating a strong positive association among the ranking orders generated under different cluster-number settings. This suggests that the ranking structure remains largely unchanged when the value of c varies within the tested range.

Table 17 further shows that the best alternative is consistently identified as A_{265} for all the tested values of c. Therefore, the recognition of the optimal alternative is not affected by small changes in the cluster number. From the perspective of classification performance, the values of *PR*, *RR*, and *ER* fluctuate only slightly across different settings of c, which indicates that the three-way classification results are also insensitive to small perturbations in the cluster-number parameter.

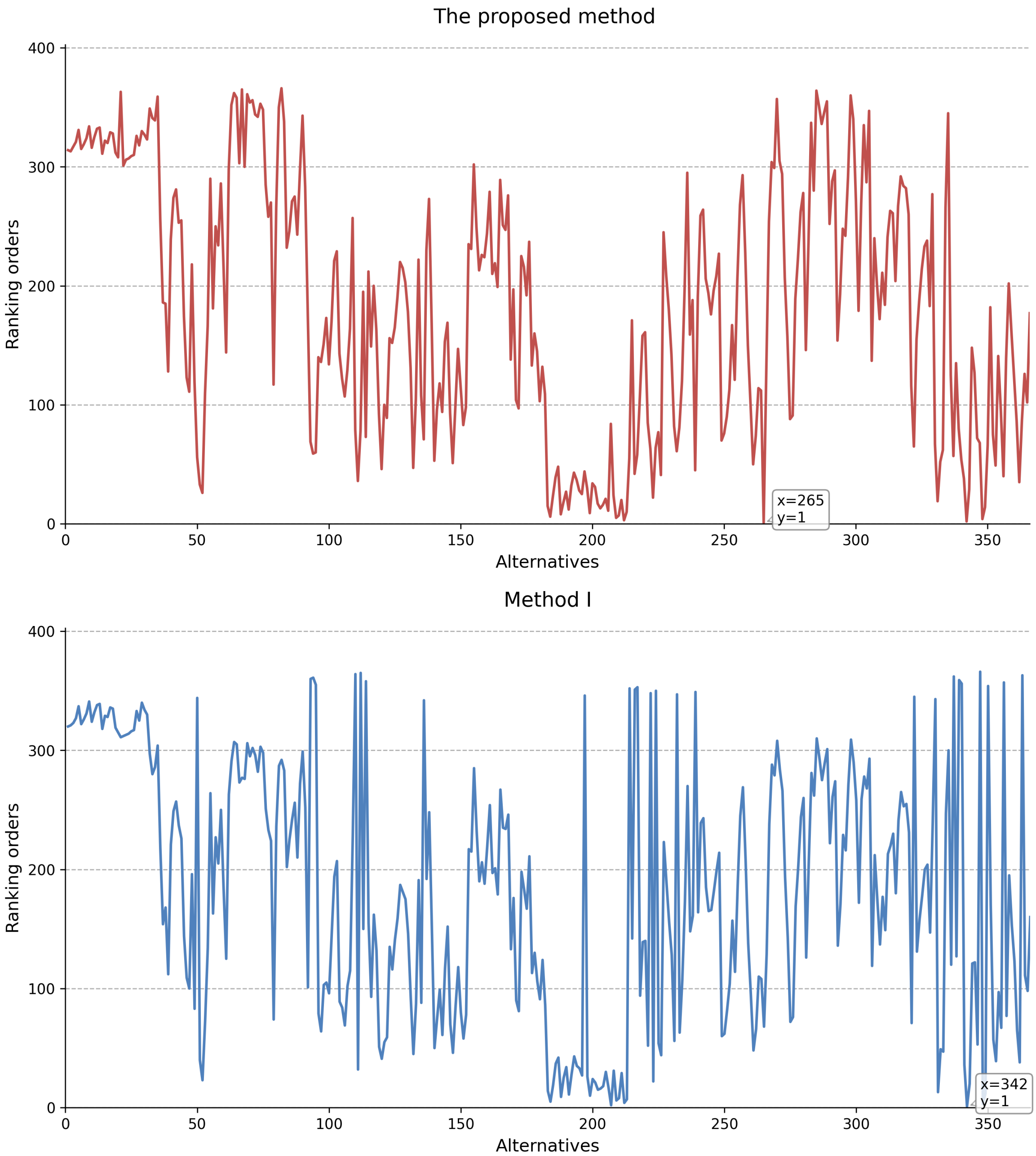
Overall, these results demonstrate that the proposed FCM-TODIM-STWD method is locally robust with respect to the choice of the cluster number. Although the parameter c is introduced through a heuristic rule, the final decision results are not driven by one specific cluster-number setting. Instead, the clustering stage mainly serves as a structural aggregation mechanism for large-scale expert information, without causing substantial changes in the final ranking or classification outcomes under nearby parameter settings.

7.4. Ablation Experiments Based on the Proposed FCM-TODIM-STWD Method

This paper proposes a FCM-TODIM-STWD method by integrating TODIM with sequential three-way decision-making. To further verify the effectiveness and advantages of the proposed method, ablation experiments are designed in this subsection by removing the TODIM module and the sequential three-way decision module, respectively. Specifically, the model without TODIM, which no longer takes the psychological behavioral factors of decision makers into account, is denoted as Method I, whereas the model in which sequential three-way decision is replaced by classical three-way decision is denoted as Method II. Based on the dataset in Section 6, Figure 12 and Figure 13 present the ranking and classification results of the three methods, respectively, and Table 18 reports the corresponding *PR*, *RR*, and *ER*.

Table 18. *PR*, *RR* and *ER* with different methods.

Methods	*PR*	*RR*	*ER*
The proposed FCM-TODIM-STWD method	92.22%	96.25%	10.38%
Method I (without TODIM)	92%	14.37%	75.96%
Method II (without STWD)	87.93%	60.62%	42.08%



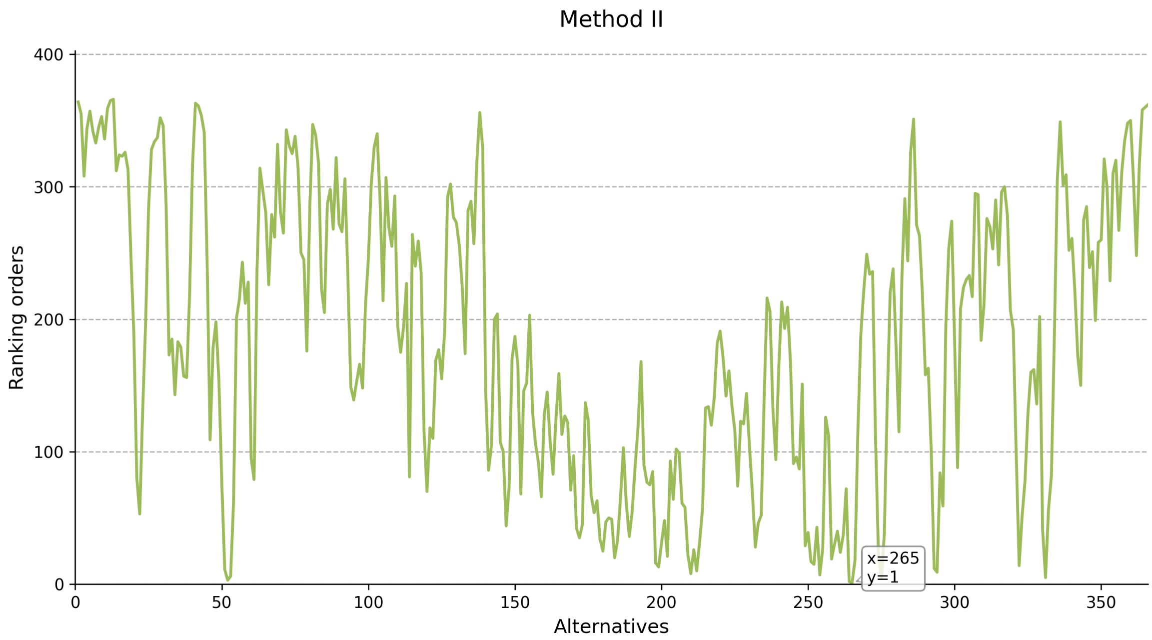


Figure 12. Ranking results of alternatives by different methods.

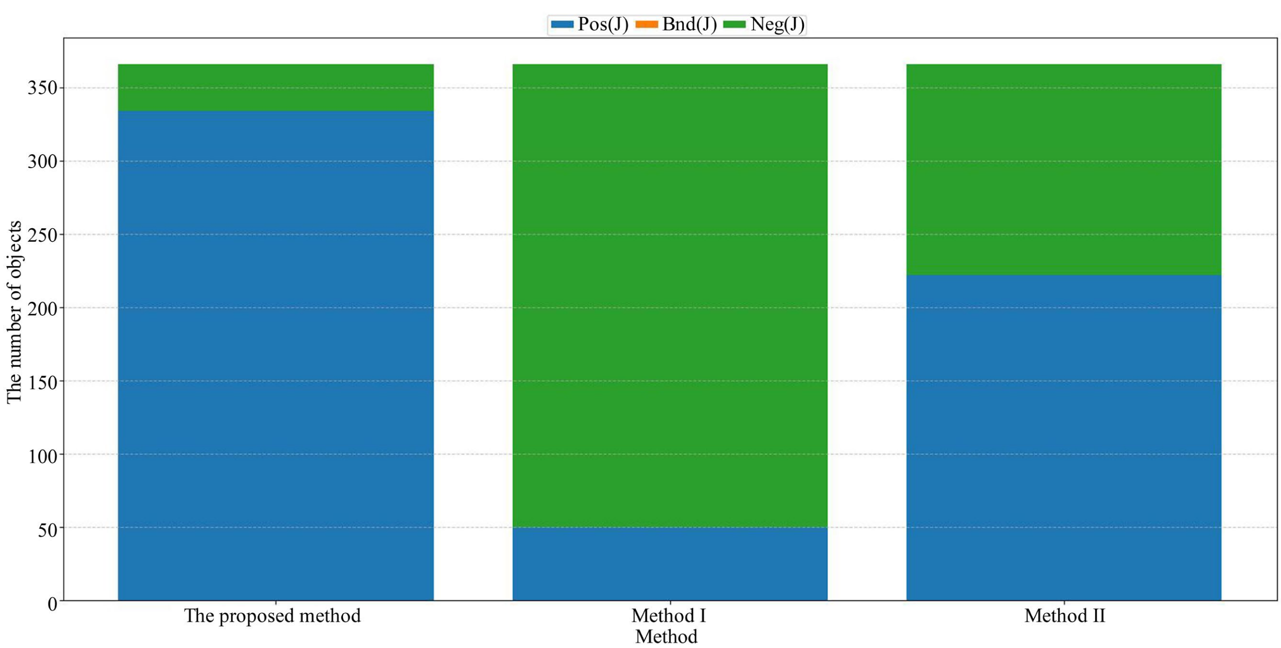


Figure 13. Classification results of alternatives by different methods.

By combining Figure 12 and Table 18, it can be observed that, compared with Method II and the proposed method, Method I yields a different optimal alternative, indicating that the introduction of the TODIM module affects the final ranking results. Meanwhile, Table 18 shows that the proposed FCM-TODIM-STWD method achieves the highest precision rate and recall rate, as well as the lowest error rate, among the three methods. In contrast, removing the TODIM module leads to a dramatic decline in recall and a substantial increase in error rate, while replacing sequential three-way decision with classical three-way decision also weakens classification performance. These results indicate that the TODIM mechanism and the sequential screening strategy jointly contribute to the behavioral rationality, discrimination ability, and overall effectiveness of the proposed method.

8. Conclusions

Probabilistic linguistic term sets (PLTSs) provide an effective tool for flexibly representing uncertain evaluation information supplied by decision makers (DMs). Sequential three-way decision (STWD) offers an effective mechanism for progressive classification in complex decision environments. This paper focuses on large-group multi-attribute decision making (LGMADM) problems in the context of PLTSs and proposes a FCM-TODIM-STWD method. The main contributions of this paper are summarized as follows.

1) A novel deviation-range entropy of PLTSs is proposed to determine attribute weights objectively. Compared with existing entropy measures, the proposed entropy captures richer uncertainty information and has stronger discriminating power.

2) A TODIM-based sequential three-way decision mechanism is developed, in which the psychological behavior of DMs is incorporated into the decision process through the loss-aversion mechanism of TODIM. Thus, the proposed method can reflect behavioral preferences more realistically.

3) According to the obtained attribute weights, attributes are incorporated into the decision process sequentially in descending order of importance. Alternatives in the boundary region are progressively refined until all alternatives are classified. This strategy reduces unnecessary information acquisition during the decision process.

4) The feasibility and superiority of the proposed method are verified through a case study, comparative analysis, sensitivity analysis, and ablation experiments. The results show that the proposed method can provide interpretable decision results and exhibits good classification performance in the PLTS-based LGMADM environment.

With the increasing complexity of decision scenarios, LGMADM has become an important research topic. Future studies will further extend the proposed model to dynamic decision environments, heterogeneous information fusion problems, and consensus adjustment processes, so as to improve the applicability and scalability of the proposed framework.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Garcia-Zamora, D., Labella, A., Ding, W., Rodriguez, R.M. and Martinez, L. (2022)

Large-Scale Group Decision Making: A Systematic Review and a Critical Analysis. *IEEE/CAA Journal of Automatica Sinica*, **9**, 949-966. https://doi.org/10.1109/jas.2022.105617

[2] Zhong, X. and Xu, X. (2020) Clustering-Based Method for Large Group Decision Making with Hesitant Fuzzy Linguistic Information: Integrating Correlation and Consensus. *Applied Soft Computing*, **87**, Article ID: 105973. https://doi.org/10.1016/j.asoc.2019.105973

[3] Ma, Z., Zhu, J., Ponnambalam, K. and Zhang, S. (2019) A Clustering Method for Large-Scale Group Decision-Making with Multi-Stage Hesitant Fuzzy Linguistic Terms. *Information Fusion*, **50**, 231-250. https://doi.org/10.1016/j.inffus.2019.02.001

[4] Song, Y. and Li, G. (2019) A Large-Scale Group Decision-Making with Incomplete Multi-Granular Probabilistic Linguistic Term Sets and Its Application in Sustainable Supplier Selection. *Journal of the Operational Research Society*, **70**, 827-841. https://doi.org/10.1080/01605682.2018.1458017

[5] Pang, Q., Wang, H. and Xu, Z. (2016) Probabilistic Linguistic Term Sets in Multi-Attribute Group Decision Making. *Information Sciences*, **369**, 128-143. https://doi.org/10.1016/j.ins.2016.06.021

[6] Teng, F., Du, C., Shen, M. and Liu, P. (2022) A Dynamic Large-Scale Multiple Attribute Group Decision-Making Method with Probabilistic Linguistic Term Sets Based on Trust Relationship and Opinion Correlation. *Information Sciences*, **612**, 257-295. https://doi.org/10.1016/j.ins.2022.07.092

[7] Yao, Y. (2010) Three-Way Decisions with Probabilistic Rough Sets. *Information Sciences*, **180**, 341-353. https://doi.org/10.1016/j.ins.2009.09.021

[8] Yao, Y. (2013) Granular Computing and Sequential Three-Way Decisions. In: Lingras, P., *et al.*, Eds., *International Conference on Rough Sets and Knowledge Technology*, Springer, 16-27. https://doi.org/10.1007/978-3-642-41299-8_3

[9] Yao, Y. (2018) Three-Way Decision and Granular Computing. *International Journal of Approximate Reasoning*, **103**, 107-123. https://doi.org/10.1016/j.ijar.2018.09.005

[10] Qian, J., Liu, C., Miao, D. and Yue, X. (2020) Sequential Three-Way Decisions via Multi-Granularity. *Information Sciences*, **507**, 606-629. https://doi.org/10.1016/j.ins.2019.03.052

[11] Xu, W., Ai, S. and Yang, H. (2026) Sequential Three-Way Decision with Adaptive Thresholds and Its Applications in Two Binary Classifiers. *Expert Systems with Applications*, **295**, Article ID: 128194. https://doi.org/10.1016/j.eswa.2025.128194

[12] Jiang, H. and Hu, B.Q. (2024) A Sequential Multiple Attribute Three-Way Group Decision-Making Method to Heterogeneous MAGDM Problems with Unknown Weight Information. *Expert Systems with Applications*, **256**, Article ID: 124869. https://doi.org/10.1016/j.eswa.2024.124869

[13] Wang, M., Liang, D. and Xu, Z. (2020) Sequential Three-Way Multiple Attribute Group Decisions with Individual Attributes and Its Consensus Achievement Based on Social Influence. *Information Sciences*, **518**, 286-308. https://doi.org/10.1016/j.ins.2020.01.024

[14] Wang, Y., Sun, B., Zhang, X. and Wang, Q. (2020) BWM and Multimoora-Based Multigranulation Sequential Three-Way Decision Model for Multi-Attribute Group Decision-Making Problem. *International Journal of Approximate Reasoning*, **125**, 169-186. https://doi.org/10.1016/j.ijar.2020.07.003

[15] Han, X. and Zhan, J. (2023) A Sequential Three-Way Decision-Based Group Consensus Method under Probabilistic Linguistic Term Sets. *Information Sciences*, **624**, 567-

589. https://doi.org/10.1016/j.ins.2022.12.111

[16] Luo, N., Zhang, Q., Xie, Q., Wang, Y., Yin, L. and Wang, G. (2025) Sequential Three-Way Group Decision-Making for Double Hierarchy Hesitant Fuzzy Linguistic Term Set. *Information Sciences*, **687**, Article ID: 121403. https://doi.org/10.1016/j.ins.2024.121403

[17] Mao, Y., Xu, Y. and Fan, Y. (2025) A Sequential Three-Way Decision Model Based on Linguistic Z-Numbers. *PLOS One*, **20**, e0320350. https://doi.org/10.1371/journal.pone.0320350

[18] Shi, J., Liu, Q., Shi, C., Lv, M. and Pang, W. (2024) Intuitionistic Fuzzy Sequential Three-Way Decision Model in Incomplete Information Systems. *Symmetry*, **16**, Article No. 1244. https://doi.org/10.3390/sym16091244

[19] Han, X., Zhang, C. and Zhan, J. (2022) A Three-Way Decision Method under Probabilistic Linguistic Term Sets and Its Application to Air Quality Index. *Information Sciences*, **617**, 254-276. https://doi.org/10.1016/j.ins.2022.10.108

[20] Wang, Y., Zhan, J. and Zhang, C. (2023) A Three-Way Decision Method Based on Prospect Theory under Probabilistic Linguistic Term Sets. *Information Sciences*, **645**, Article ID: 119342. https://doi.org/10.1016/j.ins.2023.119342

[21] Xu, G.-L., Wan, S.-P. and Dong, J.-Y. (2020) An Entropy-Based Method for Probabilistic Linguistic Group Decision Making and Its Application of Selecting Car Sharing Platforms. *Informatica*, **31**, 621-658. https://doi.org/10.15388/20-infor423

[22] Wang, T., Huang, B., Li, H., Liu, D. and Yu, H. (2023) Three-Way Decision for Probabilistic Linguistic Conflict Analysis via Compounded Risk Preference. *Information Sciences*, **631**, 65-90. https://doi.org/10.1016/j.ins.2023.02.067

[23] Li, X., Xu, Z. and Wang, H. (2023) Three-Way Group Decision Making Based on Evidential Reasoning with Probabilistic Linguistic Term Assessments. *International Journal of Fuzzy Systems*, **25**, 1742-1755. https://doi.org/10.1007/s40815-023-01486-1

[24] Chen, J. and Yin, C. (2024) A New Probabilistic Linguistic Decision-Making Process Based on PL-BWM and Improved Three-Way TODIM Methods. *Egyptian Informatics Journal*, **28**, Article ID: 100567. https://doi.org/10.1016/j.eij.2024.100567

[25] Wang, F., Pan, Y., Xu, G. and Wei, Q. (2025) Probabilistic Linguistic Multi-Attribute Group Decision Making Method Considering the Important Degrees of Experts and Attributes. *Expert Systems with Applications*, **265**, Article ID: 125899. https://doi.org/10.1016/j.eswa.2024.125899

[26] Wu, X., Liao, H., Lev, B. and Zavadskas, E.K. (2021) A Multiple Criteria Decision-Making Method with Heterogeneous Linguistic Expressions. *IEEE Transactions on Engineering Management*, **70**, 1857-1870. https://doi.org/10.1109/tem.2021.3072590

[27] Yang, H., Liu, X. and Guo, Z. (2025) A Three-Way Decision Method Based on COPRAS in the Weak Probabilistic Linguistic Term Set Information Systems. *International Journal of Machine Learning and Cybernetics*, **16**, 1265-1290. https://doi.org/10.1007/s13042-024-02333-x

[28] Du, J., Liu, S., Tao, L. and Dong, W. (2023) Three-Way Failure Mode and Effect Analysis Approach for Reliability Management in Multigranular Unbalanced Linguistic Contexts. *Computers & Industrial Engineering*, **175**, Article ID: 108909. https://doi.org/10.1016/j.cie.2022.108909

[29] Xu, Z.S. (2004) EOWA and EOWG Operators for Aggregating Linguistic Labels Based on Linguistic Preference Relations. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **12**, 791-810.

https://doi.org/10.1142/s0218488504003211

[30] Gou, X., Xu, Z. and Liao, H. (2017) Multiple Criteria Decision Making Based on Bonferroni Means with Hesitant Fuzzy Linguistic Information. *Soft Computing*, **21**, 6515-6529. https://doi.org/10.1007/s00500-016-2211-1

[31] Wu, X., Liao, H., Xu, Z., Hafezalkotob, A. and Herrera, F. (2018) Probabilistic Linguistic MULTIMOORA: A Multicriteria Decision Making Method Based on the Probabilistic Linguistic Expectation Function and the Improved Borda Rule. *IEEE Transactions on Fuzzy Systems*, **26**, 3688-3702. https://doi.org/10.1109/tfuzz.2018.2843330

[32] Zhang, Y., Xu, Wang, H. and Liao, H. (2016) Consistency-Based Risk Assessment with Probabilistic Linguistic Preference Relation. *Applied Soft Computing*, **49**, 817-833. https://doi.org/10.1016/j.asoc.2016.08.045

[33] Zhang, X. and Liu, F. (2025) Probabilistic Linguistic Three-Way Decisions: Integrating Prospect Theory with Fuzzy Possibilistic C-Means Clustering. *Fuzzy Sets and Systems*, **517**, Article ID: 109442. https://doi.org/10.1016/j.fss.2025.109442

[34] Zhu, J., Ma, X., Zhan, J. and Yao, Y. (2022) A Three-Way Multi-Attribute Decision Making Method Based on Regret Theory and Its Application to Medical Data in Fuzzy Environments. *Applied Soft Computing*, **123**, Article ID: 108975. https://doi.org/10.1016/j.asoc.2022.108975

[35] Pawlak, Z. (1982) Rough Sets. *International Journal of Computer & Information Sciences*, **11**, 341-356. https://doi.org/10.1007/bf01001956

[36] Autran Monteiro Gomes, L.F. and Duncan Rangel, L.A. (2009) An Application of the TODIM Method to the Multicriteria Rental Evaluation of Residential Properties. *European Journal of Operational Research*, **193**, 204-211. https://doi.org/10.1016/j.ejor.2007.10.046

[37] Liang, W., Luo, S. and Zhao, G. (2018) Evaluation of Cleaner Production for Gold Mines Employing a Hybrid Multi-Criteria Decision Making Approach. *Sustainability*, **11**, Article No. 146. https://doi.org/10.3390/su11010146

[38] Shi, Z., Wu, D., Guo, C., Zhao, C., Cui, Y. and Wang, F. (2021) Fcm-rdpa: TSK Fuzzy Regression Model Construction Using Fuzzy C-Means Clustering, Regularization, Droprule, and Powerball Adabelief. *Information Sciences*, **574**, 490-504. https://doi.org/10.1016/j.ins.2021.05.084

[39] Kodinariya, T.M., Makwana, P.R., *et al.* (2013) Review on Determining Number of Cluster in k-Means Clustering. *International Journal*, **1**, 90-95.

[40] Gomes, L. and Lima, M. (1991) Todimi: Basics and Application to Multicriteria Ranking. *Foundations of Computing and Decision Sciences*, **16**, 1-16.

[41] Liu, P. and Li, Y. (2018) The PROMTHEE II Method Based on Probabilistic Linguistic Information and Their Application to Decision Making. *Informatica*, **29**, 303-320. https://doi.org/10.15388/informatica.2018.169

[42] Liu, P. and Li, Y. (2019) An Extended MULTIMOORA Method for Probabilistic Linguistic Multi-Criteria Group Decision-Making Based on Prospect Theory. *Computers & Industrial Engineering*, **136**, 528-545. https://doi.org/10.1016/j.cie.2019.07.052

[43] Croux, C. and Dehon, C. (2010) Influence Functions of the Spearman and Kendall Correlation Measures. *Statistical Methods & Applications*, **19**, 497-515. https://doi.org/10.1007/s10260-010-0142-z

[44] Deng, T., Wang, W., Wang, C. and Zhan, J. (2025) Three-Way Decision Oriented Rational Behavior Multi-Attribute Decision-Making. *Information Fusion*, **126**, Article ID: 103538. https://doi.org/10.1016/j.inffus.2025.103538

[45] Xiao, Y., Zhan, J., Zhang, C. and Liu, P. (2024) A Sequential Three-Way Decision-Based Group Consensus Method with Regret Theory under Interval Multi-Scale Decision Information Systems. *IEEE Transactions on Emerging Topics in Computational Intelligence*, **8**, 1670-1686. https://doi.org/10.1109/tetci.2024.3354116

[46] Song, J., Leng, X. and Liu, Z. (2023) Green Supplier Selection Based on Sequential Group Three-Way Decision Making. *Mathematics*, **11**, Article No. 4605. https://doi.org/10.3390/math11224605

[47] Liang, D., Wang, M. and Xu, Z. (2021) A Novel Approach of Three-Way Decisions with Information Interaction Strategy for Intelligent Decision Making under Uncertainty. *Information Sciences*, **581**, 106-135. https://doi.org/10.1016/j.ins.2021.09.037