An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol.16, No.6, pp.2424-2444 (2026)
https://doi.org/10.36922/IJOCTA026240118

RESEARCH ARTICLE

A probabilistic hesitant fuzzy BWM-TODIM framework with coordinated consistency adjustment and a probability-sensitive dissimilarity measure for logistics provider selection

Fei Wang*, Baoyi Zhang, and Siyuan Liu

Department of Engineering Management, School of Big Data and Intelligent Engineering, Hebei University of Economics and Business, Shijiazhuang, Hebei, China

wangfei0221@hueb.edu.cn, *1021647174@qq.com*, *2629544940@qq.com*

ARTICLE INFO

Article history:
Received: June 12, 2026
Revised: September 15, 2026
Accepted: September 21, 2026
Published Online: September 30, 2026

Keywords:
Probabilistic hesitant fuzzy set
Best–worst method
TODIM
Dissimilarity measure
Multicriteria decision-making

ABSTRACT

Selecting a third-party logistics provider is a multicriteria decision problem involving delivery reliability, service quality, environmental performance, cost control, and transport risk. In such problems, expert assessments are often uncertain and hesitant, and different possible membership values may carry different confidence levels. To better represent this type of evaluation information, this study develops a probabilistic hesitant fuzzy BWM-TODIM framework for logistics provider selection. In the weighting stage, probabilistic hesitant fuzzy information is incorporated into the best–worst method (BWM) to derive criterion weights, and a coordinated inconsistency adjustment mechanism is introduced to reduce the loss of probability-weighted preference information and the resulting weight disturbance during consistency improvement. In the ranking stage, a probability-sensitive dissimilarity measure is constructed by adding a probability–membership interaction term, so that probabilistic hesitant fuzzy elements with similar membership structures but different probability distributions can be more clearly distinguished. The obtained dissimilarity information is then embedded into the TODIM procedure to support pairwise gain–loss dominance analysis under the decision maker's risk preferences. A published logistics-provider selection case is used to illustrate the proposed framework. Since the source case does not provide direct BWM pairwise comparisons, the required comparison vectors are transparently reconstructed from the reported criterion-importance probabilistic hesitant fuzzy elements, and the decision matrix is normalized before calculation. The empirical analysis and supplementary tests show that the proposed methodological modifications can improve the interpretability of weighting and ranking under probabilistic hesitant fuzzy information. Sensitivity analyses, including dissimilarity-measure ablation, inconsistency stress testing, criterion-weight perturbation, probability perturbation, and comparison with alternative methods, further indicate that the framework is useful for identifying close competing alternatives and for examining the stability of the final ranking.

CC BY

1. Introduction

Decision-making involves selecting a preferred alternative from several candidates according to the decision maker's judgments and preferences. As research on decision-making has advanced, the limitations of precise numerical values for representing complexity and uncertainty have become increasingly apparent.[1-3] To address this limitation, Zadeh[4] introduced fuzzy sets

*Corresponding Author

to represent imprecise information. Torra[5] subsequently proposed hesitant fuzzy sets, which allow several possible membership values to be assigned to an element. However, conventional hesitant fuzzy sets do not explicitly distinguish unequal occurrence probabilities among these possible values. In the actual decision-making process, different decision makers have different tendencies; consequently, conventional hesitant fuzzy sets may not adequately represent differences in decision makers' preferences. On this basis, Xu *et al.*[6] proposed probabilistic hesitant fuzzy (PHF) sets. The concept endows each member of the set with a different probability and describes the preference information with probability, which can comprehensively describe the preference and hesitation of decision-makers.

Multicriteria decision-making (MCDM) is an important part of modern decision science. It refers to the decision problem of selecting the optimal alternative or ranking alternatives under multiple criteria, and it is widely used in management and engineering.[7-9] At present, PHF sets have been applied to solve MCDM problems to a certain extent. For example, Ding *et al.*[10] proposed an emergency logistics supplier selection model considering the psychological behavior of decision makers in a complex decision-making environment based on a PHF interactive and multicriteria decision-making (Tomada de Decisão Interativa Multicritério [TODIM]) method. Liu *et al.*[11] proposed a TODIM multi-attribute decision-making method based on a PHF comprehensive distance measure. Liao *et al.*[12] combined the TODIM method with cumulative prospect theory for group decision making under a PHF environment. Zhou *et al.*[13] proposed a probabilistic q-rung hesitant fuzzy TODIM method based on a new distance measure. These studies show that TODIM has been widely extended in fuzzy and PHF environments because it can incorporate the psychological preferences of decision makers.[14-16]

The determination of criteria weights is an essential step in solving MCDM problems to reflect the importance of each criterion to the final decision.[17-19] At present, the method of determining weights in a PHF environment has been studied. For example, Yang *et al.*[20] combined social trust and opinion similarity to carry out linear weighting to complete the initial grouping and weight distribution of decision makers, and based on prospect theory, compared the gap between individual and group consensus to adjust the weight of non-cooperative decision makers, and used the TODIM–MAIRCA method to implicitly determine the attribute weight in the scheme ranking. But in this method, the balance between trust and similarity needs to be set as parameters subjectively. The weight results are highly sensitive to the values of the parameters, and the calculation complexity of the whole method is high. The classification of non-cooperative behavior is relatively rough, which cannot effectively distinguish the influence of general and serious behavior on the weight. Yang *et al.*[21] proposed a method to determine the criterion weight value by exponential PHF entropy integration. They also proposed a new method for calculating PHF entropy, which addresses the difficulty of comparing probabilistic hesitant fuzzy elements (PHFEs) with different cardinalities. However, the criterion weight value is obtained by averaging the PHF entropy of each alternative under this criterion. The evaluation information of each alternative is uneven, and the entropy difference between alternatives under individual criteria is too large. If the average value is adopted, the evaluation value of very few alternatives will be too low, and the average value will be pulled down, so that the final weight value is lower than the evaluation value of each alternative, and the accurate criterion weight cannot be obtained. Saha *et al.*[22] calculated the weight of supplier selection criteria by using the optimization model of cross-entropy and discrete measurement. This method assigns higher weights to criteria that more effectively distinguish among alternatives and exhibit greater variation, thereby maximizing dispersion in the determination of criterion weights and making fuller use of the discriminatory information contained in the evaluation data. Although this method can derive criterion weights from the discriminatory information in the evaluation data, its computational procedure is relatively complex, and large between-alternative differences may disproportionately affect the resulting weights even when the corresponding criterion is not considered highly important. Wang *et al.*[23] used the entropy weight method to calculate the index weight. This method equates the data dispersion degree of indicators with the actual importance, resulting in the weight of indicators with small differences but high criticality being seriously underestimated, and indicators dominated by extreme values may obtain artificially high weights.

Numerous methods are available for determining criterion weights, and the best worst multi-attribute decision-making method (best–

worst method [BWM]) is one of them. Among the subjective weighting methods, the entropy weight method, and many other weight determination methods mainly used in the existing research, the BWM method has the advantages of more accurate and more convenient calculation. Rezaei[24] improved the analytic hierarchy process (AHP) in 2015 and proposed the BWM method, which is based on pairwise comparison to obtain the weights of different criteria. Compared with AHP, BWM requires fewer pairwise comparisons and provides a consistency measure. Taking this as the theoretical background, Guo *et al.*[25] extended the traditional BWM to a triangular fuzzy environment. Ali[26] combined BWM with hesitant fuzzy numbers. Xu *et al.*[27] proposed interval hesitant fuzzy BWM from the perspective of interval hesitant fuzzy numbers. In addition, various types of fuzzy sets have been successfully incorporated into BWM, including the trapezoidal interval type-2 fuzzy BWM proposed by Wan *et al.*,[28] the generalized interval-valued trapezoidal fuzzy BWM proposed by Ali *et al.*,[29] the intuitionistic fuzzy BWM proposed by Mou *et al.*,[30] and the interval rough number-based BWM proposed by Deveci.[31]

However, BWM in a fuzzy environment still has some limitations. Although the above scholars introduced a variety of fuzzy sets into the BWM method, they did not consider the different membership probabilities of fuzzy sets and could not accurately show the preference information of decision makers. In the hesitant fuzzy BWM proposed by Ali,[26] the hesitant fuzzy value used in the comparison standard between the criteria is greater than 1, which is inconsistent with the definition of a hesitant fuzzy number. Li[32] proposed PHF-BWM, but the calculation is relatively complex, and the consistency of attribute weights needs to be further improved, so there is room for improvement. On this basis, this paper proposes a new BWM based on PHF numbers. The PHFE and BWM are combined to calculate the criterion weights, and a coordinated adjustment is proposed for cases with unacceptably high inconsistency.

The choice of BWM and TODIM in this paper is based on their complementary roles rather than on the assumption that they are universally superior to all alternatives. For weighting, AHP-related pairwise comparison methods require more comparison information when the number of criteria increases, while BWM uses the best and worst criteria as reference points and provides a consistency-check mechanism.[24] Entropy-based and dispersion-oriented methods can reflect objective data variation, but they may weaken the subjective importance information expressed by experts.[21-23,33] For ranking, technique for order preference by similarity to ideal solution (TOPSIS), measurement of alternatives and ranking according to compromise solution (MARCOS), VIKOR (multicriteria optimization and compromise solution), and evaluation based on distance from average solution (EDAS) rank alternatives using ideal-solution, utility/reference-solution, or average-solution principles, respectively.[23,31,34,35] TODIM, by contrast, evaluates alternatives through pairwise gains and losses and can account for loss aversion.[10,12-16,36,37] Therefore, PHF-BWM is used to derive criterion weights, and TODIM is used to transform pairwise PHFE dissimilarities into dominance values.

The comparison in **Table 1** further clarifies that the framework does not treat BWM and TODIM as universally optimal choices. BWM is selected for the weighting stage because the study requires expert-oriented criterion weights and an explicit consistency check under PHF information. TODIM is selected for the ranking stage because the proposed PHFE dissimilarity measure naturally produces pairwise differences, which can be transformed into gain–loss dominance values. Thus, EDAS, TOPSIS-type, and objective weighting methods are retained as comparison references, while BWM and TODIM form the main computational route of the present framework.

In the TODIM method, after determining the weight of each criterion, it is necessary to measure the distance of the fuzzy evaluation information of each alternative, that is, to calculate the difference between alternatives to determine the preferred one. Dissimilarity measurement is an important tool for PHFEs because it can show the degree of difference between PHF information.[33,36,38] Liu *et al.*[38] studied a PHF recognition method based on a comprehensive characteristic distance measure. Su *et al.*[39] discussed related uncertainty and information measures for PHF information. Fang *et al.*[40] proposed a PHF multi-attribute group decision-making method based on a new distance measure. Ding *et al.*[36] further proposed a PHF hybrid weighted distance measure that can consider position weight and possible value. Considering that the membership values of two PHFEs may be similar while their probability allocations differ, this paper adds the product of probability differences and membership levels,

Table 1. Comparison of candidate MCDM methods and rationale for selecting BWM and TODIM

Stage	Alternative methods and references	Main feature	Role in this study
Subjective weighting	AHP, fuzzy BWM, hesitant fuzzy BWM, PHF-BWM[24-26,28,30,32]	Derive criterion weights from expert preference information.	BWM is adopted because it requires fewer pairwise comparisons than AHP and provides a consistency-check mechanism.
Objective weighting	Entropy, cross-entropy, dispersion, and MEREC methods[21-23,33]	Determine weights from data dispersion, contrast, or information variation.	These methods are retained as references or comparisons because they may weaken subjective importance judgments.
Ideal-solution ranking	TOPSIS, MARCOS, VIKOR[23,31,34]	Rank alternatives according to ideal, compromise, or utility-based reference solutions.	They are not used as the main ranking method because the present study focuses on pairwise gain–loss dominance.
Average-solution ranking	EDAS[35]	Evaluate alternatives by positive and negative distances from the average solution.	EDAS is used as an external comparison because the source case adopts a PHF-EDAS framework.
Dominance-based ranking	TODIM, fuzzy TODIM, PHF-TODIM[10,12-16,36,37]	Model pairwise dominance under gain–loss and loss-aversion behavior.	TODIM is selected because it can directly incorporate PHFE dissimilarities and behavioral risk preference.
Proposed combination	PHF-BWM + probability-sensitive TODIM[10,32,35-37]	Links subjective PHF weighting with pairwise dominance ranking.	BWM determines criterion weights, while TODIM transforms weighted PHFE dissimilarities into final dominance values.

Abbreviations: AHP: Analytic hierarchy process; BWM: Best–worst method; EDAS: Evaluation based on distance from average solution; MARCOS: Measurement of alternatives and ranking according to compromise solution; MCDM: Multicriteria decision-making; MEREC: Method based on the removal effects of criteria; PHF: Probabilistic hesitant fuzzy; PHF-BWM: Probabilistic hesitant fuzzy best–worst method; PHF-EDAS: Probabilistic hesitant fuzzy evaluation based on distance from average solution; PHF-TODIM: Probabilistic hesitant fuzzy TODIM; PHFE: Probabilistic hesitant fuzzy element; TODIM: Tomada de Decisão Interativa Multicritério (interactive and multicriteria decision-making); TOPSIS: Technique for order preference by similarity to ideal solution; VIKOR: VIseKriterijumska Optimizacija I Kompromisno Resenje (multicriteria optimization and compromise solution).

so that the obtained dissimilarity becomes more responsive to differences in probability allocation.

Recent fuzzy-TODIM applications further show that the method is still being extended in different uncertain environments. Bhowmik *et al.*,[36] for instance, proposed an extended fuzzy-TODIM approach based on Schweizer-Sklar and power averaging operators to evaluate green energy sources in an interval-valued intuitionistic environment. Their study also used sensitivity analysis to examine the robustness of the ranking results. Compared with these operator-oriented fuzzy-TODIM studies, the present paper focuses on two narrower issues in PHF decision making: coordinated consistency adjustment in PHF-BWM and probability-sensitive dissimilarity in the TODIM ranking process.

The empirical case is based on a published assessment of five logistics service providers for a food and beverage company in Chennai, India.[35] The source study reports seven economic, environmental, and service criteria, three expert-importance matrices, expert weights, and an aggregate PHFE alternative matrix. It does not report BWM comparison vectors. For that reason, this paper uses the source criterion-importance PHFEs only as input for a transparent reconstruction of the Best-to-Others

Table 2. Summary of related studies and links to the present work

Research stream	Representative studies	Main contribution	Remaining limitation	Link to this study
PHF-TODIM and behavioral MCDM	Ding *et al.*;[10] Liao *et al.*;[12] Zhou *et al.*[13]	Introduce prospect-theory or TODIM-based ranking in PHF settings.	The weighting stage and consistency adjustment are usually not discussed as the main problem.	This paper combines TODIM ranking with a PHF-BWM weighting and adjustment process.
Fuzzy-TODIM applications	Bhowmik *et al.*[37]	Extend TODIM with fuzzy information, Schweizer-Sklar and power averaging operators.	The focus is on fuzzy operator aggregation rather than PHFE probability allocation.	This paper studies probability-sensitive PHFE dissimilarity in the TODIM procedure.
PHF distance/ dissimilarity measures	Ding *et al.*;[36] Liu *et al*;[38] Su *et al.*;[39] Fang *et al.*[40]	Measure differences between PHFEs.	The interaction between probability allocation and membership level still needs further attention.	This paper adds a probability–membership interaction term and tests it by ablation.
Fuzzy and hesitant BWM	Guo and Zhao;[25] Ali and Rashid;[26] Wan *et al.*[28]	Extend BWM to different fuzzy environments.	Some models do not explicitly model probability information in hesitant judgments.	This paper uses PHFE comparison information in BWM weight calculation.
PHF-BWM weighting	Li *et al.*[32]	Introduces BWM under PHF information.	Consistency adjustment may still change the implied preference and weight structure.	This paper uses a coordinated adjustment idea and verifies it through stress testing.
Logistics/supplier selection	Ortiz-Barrios *et al.*;[9] Saha *et al.*;[22] Krishankumar *et al.*[35]	Apply fuzzy MCDM methods to supplier and logistics decisions.	Many applications focus on a single final ranking and provide limited module-level validation.	This paper separates weighting, dissimilarity, behavioral parameters, and perturbation effects.

Abbreviations: BWM: Best–worst method; MCDM: Multicriteria decision-making; PHF: Probabilistic hesitant fuzzy; PHF-BWM: Probabilistic hesitant fuzzy best–worst method; PHF-TODIM: Probabilistic hesitant fuzzy Tomada de Decisão Interativa Multicritério; PHFE: Probabilistic hesitant fuzzy element; TODIM: Tomada de Decisão Interativa Multicritério (interactive and multicriteria decision-making).

and Others-to-Worst vectors. This secondary-data adaptation is stated explicitly and checked against the source entropy weights.

Based on the above discussion, this study makes two methodological contributions and one application-oriented contribution. First, to reduce the possible loss of probabilistic preference information during PHF-BWM inconsistency adjustment, a coordinated adjustment mechanism is introduced by considering membership information, probability information, and weight drift together. Second, to distinguish PHFEs with similar membership supports but different probability allocations, a probability-sensitive PHFE dissimilarity measure is proposed by adding a probability–membership interaction term. Third, to improve traceability and reproducibility, the framework is applied to a published logistics-provider case and is examined through dissimilarity-measure ablation, inconsistency stress testing, weighting comparison, behavioral-parameter sensitivity, probability perturbation, and rank-reversal diagnostics.

The remainder of this paper is organized as follows. Section 2 introduces the basic PHFS concepts and the data-preparation rules. Section 3 presents the PHF-BWM weighting procedure and

coordinated inconsistency adjustment. Section 4 gives the probability-sensitive dissimilarity measure and the TODIM calculation. Section 5 describes the logistics-provider case and the reconstruction of BWM comparison vectors. Section 6 reports the main weighting and ranking results. Section 7 provides validation, sensitivity analysis, and discussion. Section 8 concludes the paper. **Figure 1** illustrates the overall methodological framework and analysis workflow of this study.

2. Preliminaries

Xu *et al.*[6] used the term hesitant probabilistic fuzzy set in the original definition. In recent MCDM studies, the term PHF set is also commonly used for the same type of information representation. To keep consistency with the PHF-MCDM terminology used in the present paper, this study uses PHF set and PHFE throughout. Unless otherwise specified, the two terms are treated as referring to the same representation.

This section introduces the basic concepts and operations used in the PHF-BWM-TODIM framework. The PHFS definition describes the coexistence of hesitation and probability in expert assessments. The score and variance functions provide a simple comparison rule for PHFEs, while probability normalization, cost-criterion conversion, and element ordering ensure that the empirical data have a consistent form for later calculations.

2.1. Probabilistic hesitant fuzzy set

A PHFE contains two types of information. The membership values express the possible degrees to which an object satisfies a criterion, and the corresponding probabilities describe the relative support or likelihood assigned to these possible values. Compared with a single membership value, this representation retains the hesitation of experts; compared with an ordinary hesitant fuzzy element, it further records that different possible membership values may not have the same importance. In this paper, PHFEs appear in two places: the PHF-BWM comparison information used to obtain criterion weights, and the alternative evaluation matrix used in the TODIM ranking process.

Let X be a non-empty set. A PHF set H_p on X is written as:

$$H_P = \left\{ \left\langle x, h_P(x) \right\rangle \middle| \; x \in X \right\} \tag{1}$$

where:

$$h_P(x) = \left\{ \gamma^l \middle| \; p^l; l = 1, 2, \ldots, L \right\} \tag{2}$$

is a PHFE, $0 \le \gamma^l \le 1$, $0 \le p^l \le 1$, and:

$$\sum_{l=1}^{L} p^l = 1 \tag{3}$$

The published case permits $\sum_{l=1}^{L} p^l \le 1$ to represent partial ignorance.[39] Before the data are used in the present model, the probabilities are normalized as:

$$\overline{p}^l = \frac{p^l}{\sum_{r=1}^{L} p^r} \tag{4}$$

This preprocessing retains the relative proportions of the reported probabilities while removing the unallocated probability mass. It is used only to make the secondary data compatible with the complete-probability form adopted in this study and does not introduce an additional

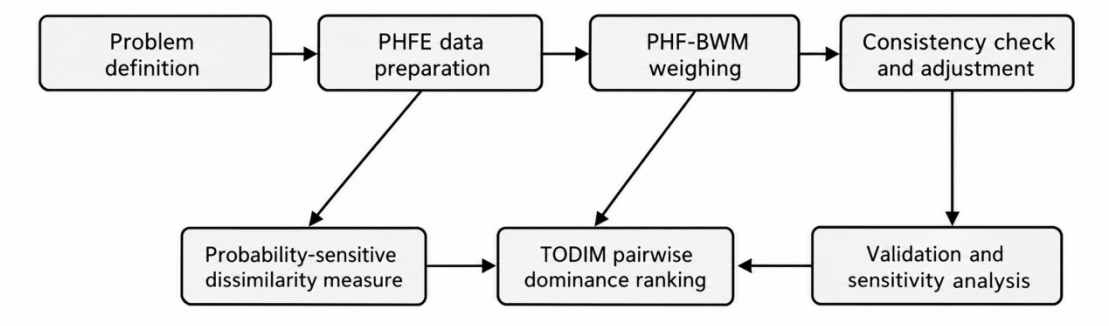


Figure 1. Research framework of the proposed PHF-BWM-TODIM method
Abbreviations: PHF-BWM: Probabilistic hesitant fuzzy best–worst method; PHFE: Probabilistic hesitant fuzzy element; TODIM: Tomada de Decisão Interativa Multicritério (interactive and multicriteria decision-making).

decision rule.

For a PHFE $h=\{\gamma^l|\ p^l\}$, the score and variance are:

$$S(h)=\sum_{l=1}^{L}\gamma^l p^l \tag{5}$$

$$V(h)=\sum_{l=1}^{L}p^l\left(\gamma^l-S(h)\right)^2 \tag{6}$$

For two PHFEs h_1 and h_2, $h_1>h_2$ when $S(h_1)>S(h_2)$. If the scores are equal, the PHFE with the smaller variance is preferred.

The score function is used as the primary comparison basis because it summarizes the probability-weighted membership level of a PHFE. The variance is used only when two PHFEs have the same score, in which case a smaller variance indicates a more concentrated and less dispersed assessment. This comparison rule is also used later to determine whether one alternative performs better, worse, or equally under a given criterion before the TODIM dominance value is calculated.

2.2. Cost-criterion conversion and element ordering

Before applying TODIM, all criteria must have the same preference direction. Benefit criteria are kept unchanged because larger membership values indicate better performance. Cost criteria are converted by complementing the membership values, so that a lower original cost or risk is transformed into a higher benefit-form evaluation.

The TODIM calculation is carried out after all criteria have been converted to benefit form. A cost PHFE is transformed by:

$$h^c=\left\{1-\gamma^l|\ p^l\right\} \tag{7}$$

Within each PHFE, membership values are ordered from small to large, and each probability remains attached to its original membership. The logistics-provider matrix has two components in every entry, so no cardinality extension is required.

Ordering is required because the dissimilarity measure compares the components of two PHFEs position by position. When the order of membership values changes, the attached probabilities must move together with their original membership values; otherwise, the probability information would be separated from the assessment to which it belongs. Since every entry in the logistics-provider decision matrix contains two components, no artificial element extension is performed in the empirical calculation.

3. Probabilistic hesitant fuzzy best–worst method

This section describes how the criterion weights are obtained under PHF information. The PHF-BWM procedure includes comparison-vector construction, membership and probability optimization, consistency checking, and coordinated adjustment when the consistency ratio exceeds the accepted threshold.

3.1. Comparison information and weight calculation

Let $C=\{C_1,C_2,\ldots,C_n\}$ be the criterion set. The best criterion is denoted by C_B and the worst criterion by C_W. The probabilistic Best-to-Others and Others-to-Worst vectors are:

$$A_{OB}=\left(n_{1B},n_{2B},\ldots,n_{nB}\right) \tag{8}$$

$$A_{WO}=\left(n_{W1},n_{W2},\ldots,n_{Wn}\right) \tag{9}$$

Each comparison is a PHFE:

$$n_{jB}=\left\{n_{jB}^a|\ p_{jB}^a;a=1,\ldots,m\right\} \tag{10}$$

$$n_{Wj}=\left\{n_{Wj}^a|\ p_{Wj}^a;a=1,\ldots,m\right\} \tag{11}$$

When judgments are collected directly, the linguistic scale in **Table 3** may be used. If only published criterion-importance PHFEs are available, the BWM comparison vectors are reconstructed from relative importance scores as explained in Section 5.2. This reconstruction is a data-processing step and should not be interpreted as direct BWM elicitation.

Algorithm 1. PHF-BWM weighting and coordinated consistency adjustment

Step 1: identify the criteria, the best criterion, and the worst criterion.
Step 2: obtain or reconstruct the Best-to-Others and Others-to-Worst PHFE comparison vectors.
Step 3: solve the membership optimization model and the probability optimization model.
Step 4: calculate the crisp criterion weights using the PHFE score function.
Step 5: calculate the consistency ratio.
Step 6: if the consistency ratio is acceptable, retain the weights; otherwise, apply the coordinated adjustment and recalculate the weights and consistency ratio.

Table 3. Linguistic scale for direct PHF-BWM judgments

Linguistic judgment	PHFE
Equally important	$\{1\|1/3,1\|1/3,1\|1/3\}$
Relatively unimportant	$\{0.7\|0.2,0.8\|0.4,0.9\|0.4\}$
Quite unimportant	$\{0.5\|0.2,0.6\|0.3,0.7\|0.5\}$
Very unimportant	$\{0.3\|0.1,0.4\|0.6,0.5\|0.3\}$
Absolutely unimportant	$\{0.1\|0.2,0.2\|0.5,0.3\|0.3\}$

Abbreviations: PHF-BWM: Probabilistic hesitant fuzzy best–worst method; PHFE: Probabilistic hesitant fuzzy element.

Let the PHF weight of C_j be:

$$w_j = \left\{h_j^a |\ p_j^a; a = 1,\ldots,m\right\}. \tag{12}$$

For the membership components, the following minimax model is solved:

$$\min \xi$$

subject to

$$\left|\frac{h_j^a}{h_B^a} - n_{jB}^a\right| \le \xi, \qquad \left|\frac{h_W^a}{h_j^a} - n_{Wj}^a\right| \le \xi \tag{13}$$

$$\sum_{j=1}^{n}\sum_{a=1}^{m}\frac{h_j^a}{n/2} = 1, \qquad 0 \le h_j^a \le 1 \tag{14}$$

The probability components satisfy:

$$\left|\frac{p_j^a}{p_B^a} - p_{jB}^a\right| \le \xi_p, \qquad \left|\frac{p_W^a}{p_j^a} - p_{Wj}^a\right| \le \xi_p \tag{15}$$

$$\sum_{a=1}^{m} p_j^a = 1, \qquad 0 \le p_j^a \le 1 \tag{16}$$

The score of the fuzzy weight and its normalized crisp value are:

$$S\left(w_j\right) = \sum_{a=1}^{m} h_j^a p_j^a \tag{17}$$

$$w_j = \frac{S\left(w_j\right)}{\sum_{r=1}^{n} S\left(w_r\right)} \tag{18}$$

3.2. Consistency ratio

For component a, complete consistency requires:

$$n_{Wj}^a n_{jB}^a = n_{WB}^a. \tag{19}$$

Let:

$$\alpha_{WB} = max_a\left\{n_{WB}^a\right\} \tag{20}$$

The continuous consistency index is:

$$CI\left(\alpha_{WB}\right) = \frac{1 + 2\alpha_{WB} + \sqrt{1 + 8\alpha_{WB}}}{2} \tag{21}$$

The consistency ratio is:

$$CR = \frac{\xi^*}{CI\left(\alpha_{WB}\right)} \tag{22}$$

A comparison system is accepted when $CR \le 0.20$.

3.3. Coordinated inconsistency adjustment

When $CR > 0.20$, changing only the membership comparison may reduce inconsistency but can also alter the probabilistic meaning of the judgment and the final weights. The repair is therefore written as a constrained adjustment problem:

$$\min J = \xi + \lambda_1 \sum_{r,a}\left|n_r^{a'} p_r^{a'} - n_r^a p_r^a\right| + \lambda_2 \sum_{j=1}^{n}\left|w_{j'} - w_j^{(0)}\right| \tag{23}$$

subject to:

$$CR\left(A'\right) \le 0.20 \tag{24}$$

$$0 \le n_r^{a'} \le 1, \qquad 0 \le p_r^{a'} \le 1, \qquad \sum\nolimits_a p_r^{a'} = 1 \tag{25}$$

Here, A' is the repaired comparison system and $w_j^{(0)}$ is the preliminary weight before repair. The first penalty limits the change in probability-weighted preference information. The second limits the change in the criterion-weight vector. In the empirical stress test, $\lambda_1 = \lambda_2 = 1$ is used as the baseline, followed by a parameter check. This formulation does not force the weights to remain unchanged. Its purpose is to reduce unnecessary disturbance while restoring acceptable consistency.

4. Probability-sensitive dissimilarity measure and TODIM

This section presents the ranking part of the framework. A probability-sensitive dissimilarity measure is first defined for PHFEs, and then the measure is embedded into TODIM to support pairwise gain-and-loss comparison among logistics providers.

4.1. Proposed probabilistic hesitant fuzzy element dissimilarity measure

Let $h_1 = \{\gamma_1^l |\ p_1^l\}$ and $h_2 = \{\gamma_2^l |\ p_2^l\}$ be two ordered PHFEs with the same cardinality L. The proposed dissimilarity is:

$$d_P(h_1,h_2)=\frac{1}{3}\sum_{l=1}^{L}[\left|\gamma_1^l p_1^l-\gamma_2^l p_2^l\right| + \left|\gamma_1^l-\gamma_2^l\right| p_1^l p_2^l+\left|p_1^l-p_2^l\right|\gamma_1^l\gamma_2^l] \quad (26)$$

The first term measures the difference in probability-weighted membership. The second records the membership difference with the two occurrence probabilities. The third is the added interaction term. It increases the response when two PHFEs have similar membership support but allocate probability differently, particularly at larger membership levels. The measure is non-negative, symmetric, and satisfies $d_P(h,h)=0$. Because the triangle inequality is not established, the expression is used as a dissimilarity measure in the TODIM calculation rather than as a distance metric.

Proposition 1. The proposed expression is non-negative, symmetric, and equals zero when the two PHFEs are identical. In addition, after normalization by the three terms in the expression, its value is bounded within [0,1]. Since the triangle inequality is not established here, this paper uses the term dissimilarity measure rather than distance metric. This wording is consistent with the role of the expression in TODIM, where it describes the magnitude of pairwise difference rather than constructing a metric space.

4.2. TODIM calculation

Let $w_r=\max_j w_j$ be the reference weight and:

$$w_{jr}=\frac{w_j}{w_r} \quad (27)$$

For alternatives A_k and A_t, the dominance contribution under criterion C_j is:

$$\phi_j(A_k,A_t)=\begin{cases}\dfrac{\sqrt{w_{jr}d_p(h_{kj},h_{tj})}}{\sum_{s=1}^{n}w_{sr}}, & h_{kj}>h_{tj},\\ 0, & h_{kj}=h_{tj},\\ -\dfrac{1}{\theta}\sqrt{\dfrac{\left(\sum_{s=1}^{n}w_{sr}\right)d_p(h_{kj},h_{tj})}{w_{jr}}}, & h_{kj}<h_{tj}\end{cases} \quad (28)$$

The pairwise dominance and total dominance are:

$$\vartheta(A_k,A_t)=\sum_{j=1}^{n}\phi_j(A_k,A_t) \quad (29)$$

$$T_k=\sum_{t=1}^{m}\vartheta(A_k,A_t) \quad (30)$$

The normalized value is:

$$\Phi(A_k)=\frac{T_k-\min_i T_i}{\max_i T_i-\min_i T_i}. \quad (31)$$

Alternatives are ranked in descending order of $\Phi(A_k)$.

Algorithm 2. Probability-sensitive TODIM ranking

Step 1: input the normalized PHFE decision matrix and the PHF-BWM criterion weights.
Step 2: convert all cost criteria into benefit criteria and order PHFE elements.
Step 3: calculate the score of each PHFE to judge the gain or loss relation between two alternatives.
Step 4: calculate the proposed PHFE dissimilarity under each criterion.
Step 5: calculate the local dominance contribution and aggregate it into the pairwise dominance value.
Step 6: obtain the total and normalized dominance values and rank the alternatives.

5. Case study and data preparation

This section explains how the published logistics-provider data are used in the empirical analysis. The criterion-importance PHFEs provide the basis for reconstructing the PHF-BWM comparison vectors, whereas the published aggregate alternative matrix is used for TODIM ranking after probability normalization and cost-criterion conversion.

5.1. Case background

The data come from the logistics-provider study reported by Krishankumar *et al.*[35] The focal company is an anonymized food and beverage enterprise in Chennai. It selected five third-party logistics providers from an initial list of 10. Three experts—a finance and audit employee, a logistics manager, and a senior human resources employee—evaluated the providers under seven criteria. The criteria are on-time delivery (C_1), service quality (C_2), adoption of green practices (C_3), customer relationship and harmony (C_4), pollution-control strategy (C_5), total cost (C_6), and damage to raw materials during shipment (C_7). The first five criteria are benefits, while the last two are costs.

The case is suitable for illustrating the framework for three reasons. First, it is based on traceable published data and provides PHFE-based criterion-importance information and an aggregate PHFE decision matrix. Second,

it includes both benefit and cost criteria, which allows the cost-conversion rule to be demonstrated. Third, the criteria cover delivery, service, environmental performance, customer relationship, pollution control, total cost, and shipment damage, so the evaluation contains economic, operational, environmental, and risk-related dimensions. The three experts come from finance and audit, logistics management, and human resources. Although this expert group is not large, it provides different functional perspectives for an illustrative case.

The source study uses an attitude-based entropy procedure and PHF-EDAS. The present paper reuses the published criterion-importance PHFEs and the published aggregate alternative matrix but recalculates all weights, dissimilarities, dominance values, and rankings. The data source and the transformation rules are kept separate from the proposed method (**Figure 2**).

5.2. Reconstruction of PHF-BWM comparisons

The source article does not contain direct BWM comparisons. For expert e and criterion j, the normalized importance score is:

$$s_{ej} = \sum_l \gamma_{ej}^l \overline{p}_{ej}^l \tag{32}$$

The aggregated importance is:

$$G_j = \sum_{e=1}^{3} \lambda_e s_{ej} \tag{33}$$

where the published expert-weight vector is:

$$\boldsymbol{\lambda} = (0.3274, 0.3013, 0.3713) \tag{34}$$

The best and worst criteria are obtained from the largest and smallest G_j. To keep every reconstructed comparison within [0,1], a bounded ratio is used:

$$r_{jB}^{(e)} = \frac{\min\left(s_{ej}, s_{eB}\right)}{\max\left(s_{ej}, s_{eB}\right)} \tag{35}$$

$$r_{Wj}^{(e)} = \frac{\min\left(s_{eW}, s_{ej}\right)}{\max\left(s_{eW}, s_{ej}\right)} \tag{36}$$

The three expert-specific ratios form the membership support, and the expert weights form the associated probabilities:

$$n_{jB} = \left\{ r_{jB}^{(1)} |\ \lambda_1, r_{jB}^{(2)} |\ \lambda_2, r_{jB}^{(3)} |\ \lambda_3 \right\} \tag{37}$$

$$n_{Wj} = \left\{ r_{Wj}^{(1)} |\ \lambda_1, r_{Wj}^{(2)} |\ \lambda_2, r_{Wj}^{(3)} |\ \lambda_3 \right\} \tag{38}$$

This reconstruction converts absolute importance assessments into relative comparisons. It is not equivalent to asking the original experts to complete a BWM questionnaire; therefore, the resulting weights should be interpreted as reconstructed PHF-BWM weights based on the published criterion-importance information.

The reconstructed comparisons and resulting weights should therefore be interpreted carefully. The reconstructed comparison vectors are derived from the published criterion-importance PHFEs and thus preserve the relative importance information reported in the source study. However, the procedure compresses each expert-specific PHFE into a score before constructing relative ratios, and it cannot fully reproduce the cognitive process of direct BWM pairwise comparison. For this reason, the obtained weights are described as PHF-BWM weights reconstructed from secondary criterion-importance data. The reconstruction rule is reported explicitly, and the resulting weights are compared with the source entropy-based weighting logic in the validation

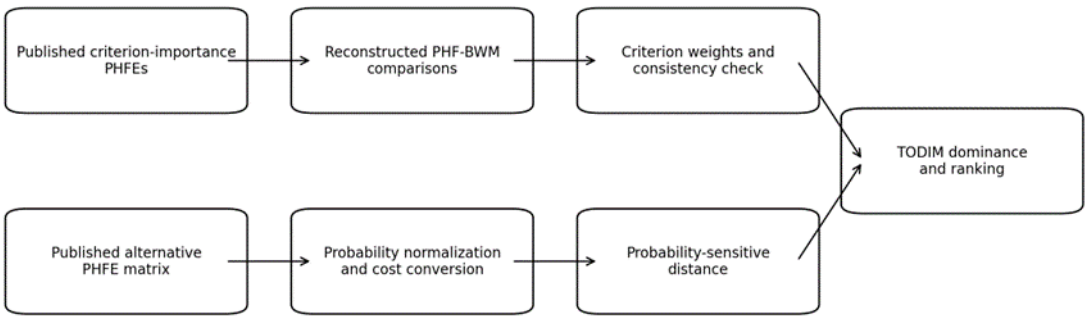


Figure 2. Data-processing procedure of the empirical case
Abbreviations: PHF-BWM: Probabilistic hesitant fuzzy best–worst method; PHFE: Probabilistic hesitant fuzzy element; TODIM: Tomada de Decisão Interativa Multicritério (interactive and multicriteria decision-making).

section.

Algorithm 3. Reconstruction of PHF-BWM comparisons from secondary PHFE data

Step 1: normalize the probabilities of each published criterion-importance PHFE.
Step 2: compute the expert-specific criterion score.
Step 3: aggregate the criterion scores using the published expert weights.
Step 4: determine the best and worst criteria from the largest and smallest aggregate scores.
Step 5: calculate bounded ratios between each criterion and the best and worst criteria.
Step 6: form the reconstructed PHFE comparison vectors by treating the three expert-specific ratios as membership supports and the expert weights as their associated probabilities.

The expert-specific criterion scores, aggregated scores, and resulting best and worst criterion designations used in the PHF-BWM reconstruction are presented in **Table 4**.

Based on the aggregated scores, C_7 was identified as the best criterion and C_5 as the worst criterion.

5.3. Alternative matrix preparation

The published aggregate matrix allows the sum of probabilities to be less than one. Each cell is first normalized by $\overline{p}^l = p^l / \sum_r p^r$. Cost criteria C_6 and C_7 are then transformed by complementing their membership values.

6. Main results

6.1. Criterion weights

The reconstructed comparisons were entered into the PHF-BWM model. The optimal deviation was $\xi^* = 0.104410$, the continuous consistency index was $CI = 2.841301$, and $CR = 0.036747 < 0.20$.

No inconsistency repair was required in the baseline case (**Table 5**).

C_7 receives the largest weight, followed by total cost. The differences among C_1–C_4 are small, suggesting that modest changes in these weights could affect the relative positions of the leading providers (**Figure 3**).

6.2. TODIM ranking

The baseline TODIM calculation uses $\theta = 0.5$. This value represents a relatively loss-averse behavioral setting, which is reasonable in logistics-provider selection because cost overruns, delivery failures, and shipment damage are usually viewed more seriously than equivalent gains. Nevertheless, theta is a behavioral parameter rather than a universal constant. For this reason, Section 7.3 reports a sensitivity analysis over several theta values.

With $\theta = 0.5$, the total and normalized dominance values are shown below (**Table 6**).

The resulting order is $TLP5 > TLP2 > TLP1 > TLP4 > TLP3$. TLP5 and TLP2 are separated by only 0.0079 in the normalized score. This small difference indicates that TLP5 and TLP2 are closely ranked, rather than clearly separated. TLP3 is clearly separated from the other providers.

7. Validation, sensitivity analysis, and discussion

7.1. Independent inconsistency stress test

The main case and the stress test have different

Table 4. Criterion-importance scores used in the best–worst method reconstruction

Criterion	Expert 1	Expert 2	Expert 3	Aggregated score	Role
C_1	0.4625	0.5471	0.6719	0.5657	
C_2	0.5800	0.5000	0.5618	0.5491	
C_3	0.4794	0.6529	0.5676	0.5645	
C_4	0.5526	0.6176	0.4735	0.5429	
C_5	0.5250	0.4971	0.5700	0.5333	Worst
C_6	0.5265	0.5750	0.6265	0.5782	
C_7	0.6526	0.5944	0.6294	0.6265	Best

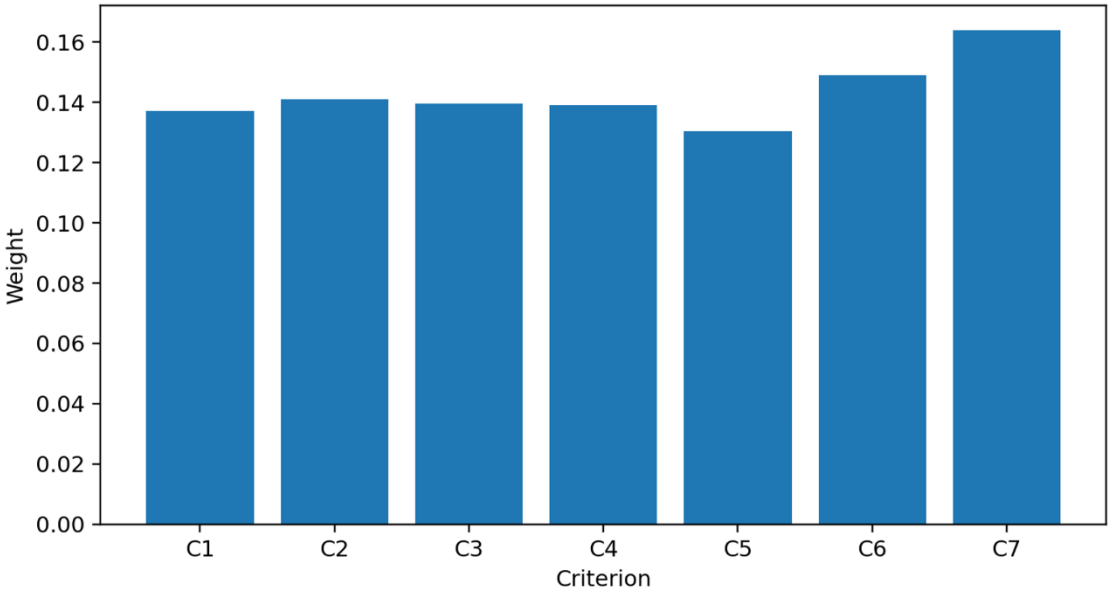


Figure 3. Criterion weights obtained by the reconstructed PHF-BWM

functions. The main case is used to obtain the baseline weights and final ranking, whereas the stress test is designed only to examine whether the adjustment mechanism behaves as expected when inconsistent PHF-BWM information occurs (**Table 7**).

Table 5. Criterion weights obtained by the reconstructed PHF-BWM

Criterion	Description	Weight	Rank
C_1	On-time delivery	0.1371	6
C_2	Service quality	0.1410	3
C_3	Adoption of green practices	0.1396	4
C_4	Customer relationship/ harmony	0.1392	5
C_5	Pollution control strategy	0.1304	7
C_6	Total cost	0.1489	2
C_7	Damage during shipment	0.1639	1

Abbreviation: PHF-BWM: Probabilistic hesitant fuzzy best–worst method.

The baseline comparison system is already consistent, so the coordinated adjustment is not activated in the main weighting result. To examine its behavior under unacceptable inconsistency, an independent stress test was conducted. The three components of the worst-to-best comparison were replaced by $c = 0.20$, 0.10, and 0.05. The corresponding pre-repair consistency ratios are 0.2342, 0.3170, and 0.3747. Under the coordinated repair model with $\lambda_1 = \lambda_2 = 1$, the common repaired membership was $x^* = 0.25494$ and the consistency ratio was reduced to the acceptance boundary of 0.20.

Table 6. TODIM dominance values and rankings

Alternative	Total dominance	Normalized value	Rank
TLP1	−20.0188	0.9009	3
TLP2	−19.4602	0.9921	2
TLP3	−25.5387	0.0000	5
TLP4	−22.7417	0.4565	4
TLP5	−19.4117	1.0000	1

Compared with a fixed adjustment to the nearest linguistic anchor 0.30, the coordinated model reduces the L1 weight drift by 44.25%, 21.40%, and 16.83% in the three stress scenarios. It also retains the complete-probability condition. These results indicate that the adjustment can reduce unnecessary disturbance in the comparison information and the derived criterion weights, although it does not eliminate weight changes completely (**Figure 4**).

A parameter check shows that balanced or stronger preservation penalties select the

Table 7. Role of the inconsistency-adjustment experiment

Item	Main case	Stress test
Purpose	Baseline weighting and ranking	Validation of repair behavior
Consistency ratio before repair	0.036747	0.2342, 0.3170, and 0.3747
Repair activated?	No	Yes
Used for final ranking?	Yes	No
Used for method validation?	Partly	Yes

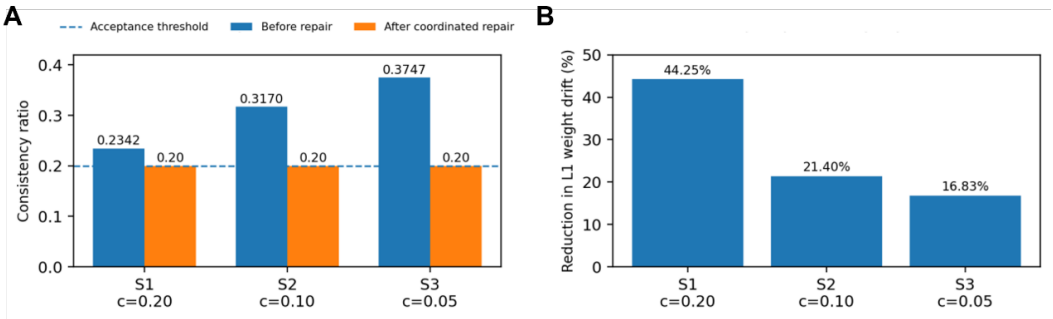


Figure 4. Visual summary of the independent inconsistency stress test. (A) Consistency recovery. (B) Disturbance reduction.

minimum feasible repair. Very weak penalties place too much emphasis on reducing the optimization deviation and produce an unnecessarily large change.

7.2. Dissimilarity-measure ablation

The third term of d_p was removed to form an ablated dissimilarity. Weighted Hamming and weighted Euclidean distances were also tested while the weights and TODIM parameters were held fixed (**Table 8; Figure 5**).

Table 8. Rankings under different dissimilarity settings

Distance setting	Ranking	Gap
Proposed interaction dissimilarity	TLP5 > TLP2 > TLP1 > TLP4 > TLP3	0.0079
Ablated dissimilarity	TLP5 > TLP1 > TLP2 > TLP4 > TLP3	0.0337
Weighted Hamming	TLP1 > TLP2 > TLP5 > TLP4 > TLP3	0.0371
Weighted Euclidean	TLP1 > TLP5 > TLP2 > TLP4 > TLP3	0.0669

The proposed interaction term affects the ordering of the top three alternatives. With the full dissimilarity measure, TLP5 ranks narrowly ahead of TLP2. Removing the interaction term keeps TLP5 first but moves TLP1 to second. Hamming and Euclidean settings place TLP1 first. The experiment therefore shows that probability allocation contributes materially to the result.

7.3. Sensitivity to the TODIM loss parameter

This section examines how changes in the TODIM loss parameter affect the normalized dominance values and ranking of the alternatives (**Table 9; Figure 6**).

For ϑ from 0.10 to 0.50, TLP5 ranks first. TLP2 becomes first under larger theta values, whereas TLP3 remains last throughout. This result indicates that the behavioral parameter matters when the leading alternatives are close. The baseline value is therefore interpreted as one reasonable behavioral setting, not as the only possible risk-attitude assumption.

7.4. Criterion-weight perturbation

Each criterion weight was varied by ±5%,

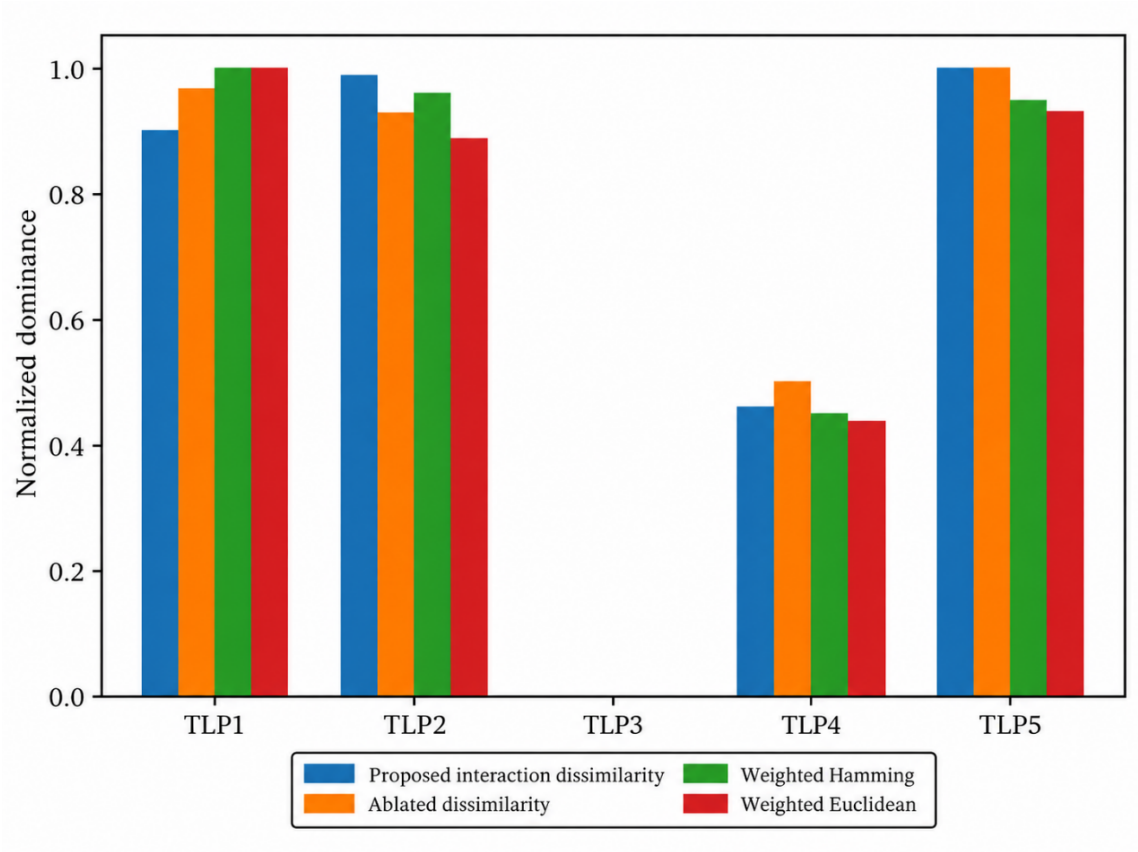


Figure 5. Alternative scores under four dissimilarity settings
Notes: TLP3 had a normalized dominance value of 0.0000 under all four settings; consequently, its bars have zero height and are not visible.

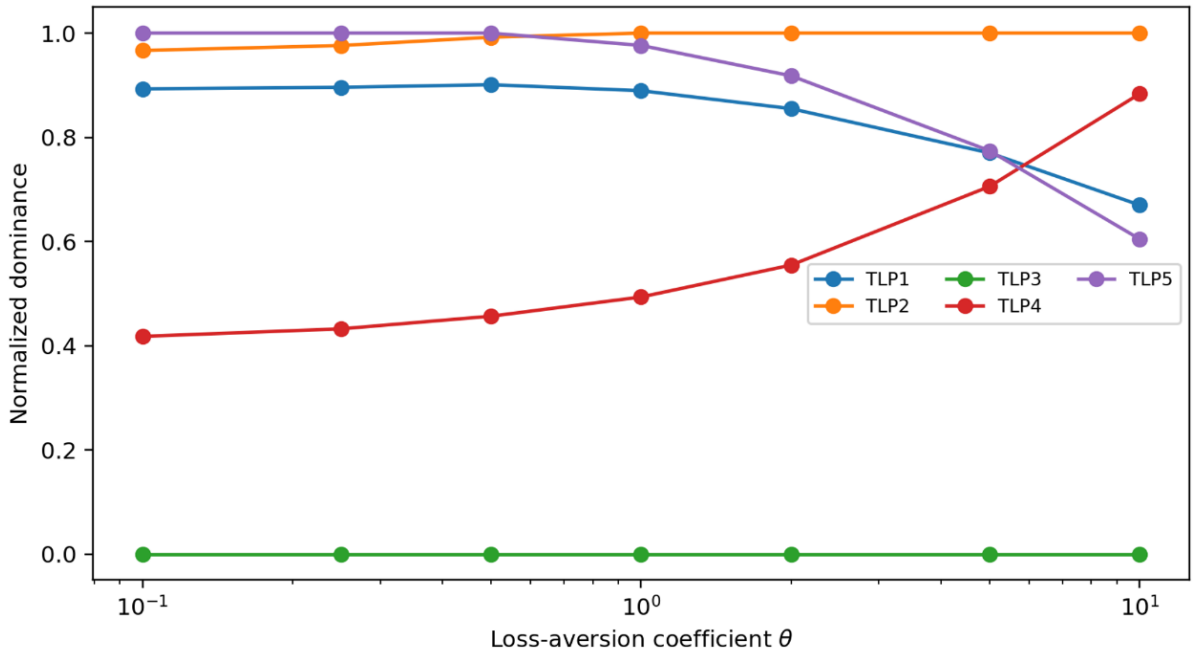


Figure 6. Sensitivity of normalized dominance values to the TODIM loss parameter

Table 9. Sensitivity of normalized dominance to TODIM loss parameter ϑ

ϑ	TLP1	TLP2	TLP3	TLP4	TLP5
0.1	0.8929	0.9665	0.0000	0.4177	1.0000
0.25	0.8959	0.9761	0.0000	0.4322	1.0000
0.5	0.9009	0.9921	0.0000	0.4565	1.0000
1	0.8894	1.0000	0.0000	0.4934	0.9763
2	0.8549	1.0000	0.0000	0.5546	0.9179
5	0.7700	1.0000	0.0000	0.7055	0.7742
10	0.6699	1.0000	0.0000	0.8833	0.6048

±10%, and ±20%, one at a time, followed by renormalization. This produced 42 scenarios (**Table 10**).

Table 10. Summary of one-at-a-time weight perturbation results

Indicator	Result
Number of one-at-a-time scenarios	42
TLP5 ranked first	26 scenarios
TLP2 ranked first	16 scenarios
Unchanged complete ranking	26 scenarios
Mean Spearman rank correlation	0.9571
Minimum Spearman rank correlation	0.7000

The mean Spearman correlation is high, but the first position switches between TLP5 and TLP2. Most changes are confined to the leading alternatives; TLP3 remains last. This pattern is consistent with the small baseline score gap.

7.5. Probability perturbation

The probability of the first element in each two-element PHFE was perturbed independently, and the second probability was set to one minus the first. Five hundred simulations were run for each perturbation range with a fixed random seed (**Table 11**).

TLP5 remains the most frequent winner, but its selection frequency falls from 59.8% under ±5% perturbation to 55.2% under ±10%. The mean rank correlation also declines. The proposed dissimilarity measure is responsive to probability changes, as intended, but that responsiveness also means that uncertain probability estimates should be subjected to sensitivity analysis.

7.6. Weighting and ranking-method comparison

Using the source article's entropy weights with the same dissimilarity measure and TODIM procedure gives TLP2 > TLP5 > TLP1 > TLP3 > TLP4. The published PHF-EDAS result is TLP1 > TLP4 > TLP2 > TLP5 > TLP3.

The different orders are expected because the weighting logic and the ranking logic differ. The proposed model uses reconstructed subjective importance relations and pairwise gain–loss dominance, while the source model uses entropy-style weights and distance from an average solution. Without a verified true ranking, the comparison is evidence of method sensitivity, not proof that one ranking is objectively correct.

7.7. Rank-reversal check

A deletion test was used as an additional diagnostic. When TLP3, the baseline last-ranked provider, was removed, the remaining order became TLP1 > TLP5 > TLP2 > TLP4.

The first position therefore changed from TLP5 to TLP1. This result indicates that the ranking may be sensitive to the alternative set in the present TODIM implementation.

7.8. Scenario-based numerical experiments

Because a single empirical case may be influenced

Table 11. Probability perturbation results

Probability perturbation	TLP1 first	TLP2 first	TLP3 first	TLP4 first	TLP5 first	Mean Spearman rank correlation	Baseline order frequency
±5%	0.188	0.212	0.000	0.002	0.598	0.8188	0.236
±10%	0.232	0.180	0.000	0.036	0.552	0.6932	0.142

by its own data structure, a supplementary scenario-based numerical experiment was conducted. This experiment was used as a computational stability check and was not treated as a substitute for independent real applications. It examined whether the procedure remains stable when the number of alternatives, the number of criteria, and the probability structure of PHFEs were changed.

In each setting, two-component PHFEs with complete probabilities were generated under a fixed random seed. The membership values were generated around the latent performance order with random noise, while the probabilities were assigned according to balanced, concentrated, dispersed, or mixed patterns. A known latent order, A1 $>$ A2 $>$ ... $>$ Am, was used only as a benchmark for evaluating the ranking output. The PHF-TODIM calculation and the probability-sensitive dissimilarity measure are applied consistently across all settings. Each setting is repeated 100 times (**Table 12; Figure 7**).

The results show that the procedure remains computationally usable under the tested scales. When the number of alternatives increases, the full-order retention rate decreases because adjacent alternatives become closer, but the mean Spearman correlation remains above 0.995 in all settings. The probability-structure tests also show that concentrated and dispersed probabilities do not invalidate the procedure, although dispersed and mixed settings make the complete ranking more sensitive. These findings provide supplementary evidence for the computational stability of the framework, while also showing that additional real-world cases are still needed for broader generalization.

7.9. Theoretical and managerial implications

From a theoretical perspective, the study shows that probability information in PHFEs should not be treated only as an auxiliary description. It can affect both the reconstructed weight information and the final TODIM dominance values. The coordinated adjustment also changes the focus of consistency repair from merely reducing the consistency ratio to reducing unnecessary changes in probabilistic preference information and criterion weights. The ablation and perturbation experiments further help separate the effects of weighting, dissimilarity, behavioral parameters, and probability uncertainty.

From a managerial perspective, the results suggest that logistics-provider selection should not rely mechanically on a single ranking value. In the baseline calculation, TLP5 is slightly ahead of TLP2, but the difference is narrow. Managers should therefore treat these two providers as close competitors and may conduct additional negotiation, site investigation, or risk checks before making the final contract decision. The ϑ-sensitivity analysis also shows that different assumptions about risk attitudes may produce different leading alternatives, especially when alternatives have similar overall performance.

7.10. Limitations

Several limitations should be acknowledged. First, the empirical analysis is based on a single logistics-provider case involving only three experts; therefore, the generalizability of the findings remains limited. Second, the source case does not provide direct BWM pairwise judgments; the BWM vectors are reconstructed from secondary criterion-importance PHFEs, which may not fully reproduce direct expert comparisons. Third, the baseline comparison system is already consistent, so the inconsistency adjustment is examined through a stress test rather than activated in the main case. Fourth, the TODIM parameter ϑ and the probability assignments may influence the leading alternatives. Fifth, the numerical

Table 12. Results of the scenario-based numerical experiments

Scenario	Scale	Prob. structure	A1 first	Top-three retained	Full order retained	Mean Spearman	Mean gap
S1	5 × 7	Balanced	100.00%	100.00%	100.00%	1.0000	0.2186
S2	8 × 7	Balanced	98.00%	95.00%	90.00%	0.9976	0.1056
S3	10 × 7	Balanced	95.00%	88.00%	77.00%	0.9971	0.0740
S4	5 × 10	Balanced	100.00%	100.00%	100.00%	1.0000	0.2054
S5	5 × 12	Balanced	100.00%	100.00%	100.00%	1.0000	0.2079
S6	8 × 10	Balanced	100.00%	98.00%	98.00%	0.9995	0.1061
S7	10 × 12	Balanced	98.00%	98.00%	91.00%	0.9989	0.0741
S8	5 × 7	Concentrated	100.00%	100.00%	100.00%	1.0000	0.1967
S9	5 × 7	Dispersed	100.00%	100.00%	100.00%	1.0000	0.2598
S10	8 × 10	Mixed	99.00%	94.00%	82.00%	0.9957	0.1365
S11	10 × 12	Concentrated	99.00%	99.00%	97.00%	0.9996	0.0686
S12	10 × 12	Dispersed	99.00%	94.00%	71.00%	0.9955	0.1071

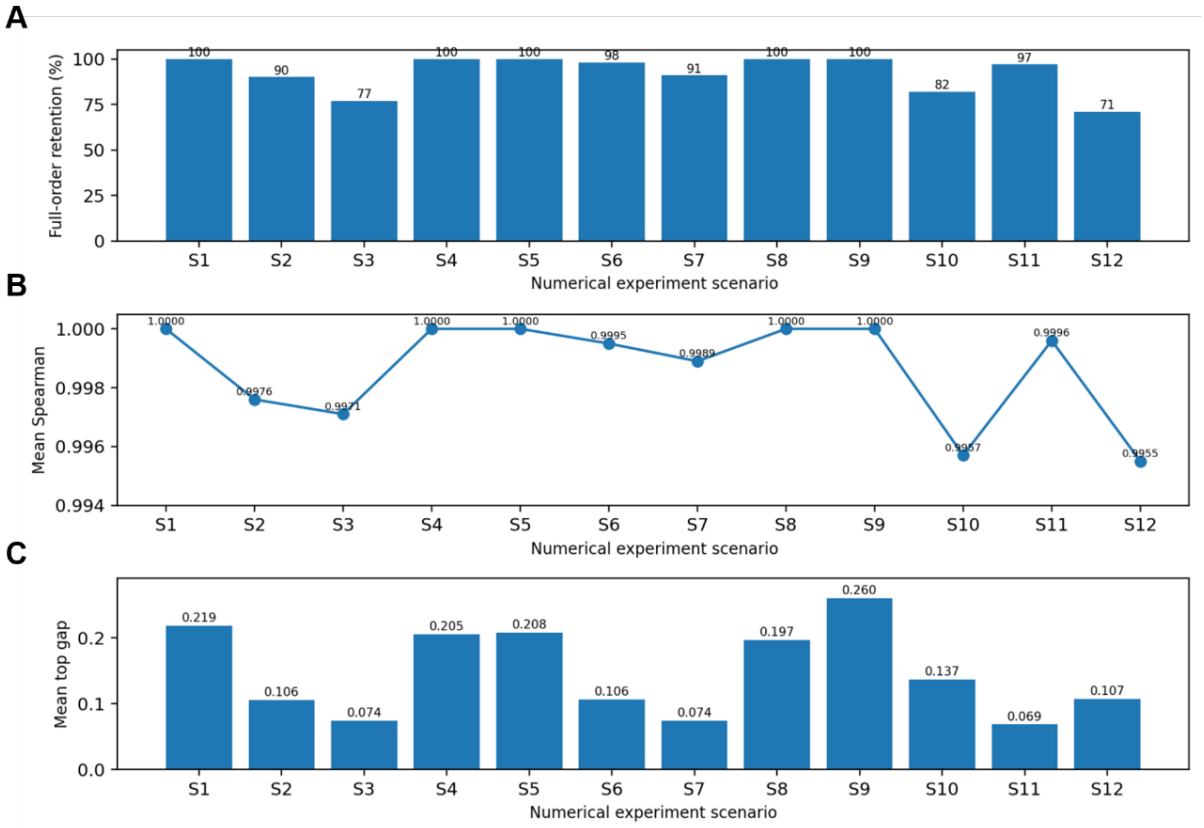


Figure 7. Visual summary of the scenario-based numerical experiments. (A) Ranking stability across data scales. (B) Spearman rank correlation across scenarios. (C) Average separation of the leading alternatives across scenarios.

experiments are synthetic checks rather than independent real cases, so they should be understood as supplementary evidence of computational behavior, not as direct empirical validation in another industry.

8. Conclusion

This study develops a PHF-BWM-TODIM framework for logistics-provider selection. In the weighting stage, a coordinated inconsistency adjustment is used to reduce the loss of probability-weighted preference information and the associated weight drift. In the ranking stage, a probability–membership interaction term is added to the PHFE dissimilarity measure and embedded into TODIM. The published case involving five alternatives and seven criteria produces the baseline order TLP5 $>$ TLP2 $>$ TLP1 $>$ TLP4 $>$ TLP3, but the first two providers are very close.

The validation results provide cautious support for the two methodological adjustments. The inconsistency stress test shows that the coordinated adjustment can reduce weight drift compared with a fixed linguistic-anchor repair. The dissimilarity-measure ablation shows that the interaction term changes the top-three ordering, indicating that probability allocation has a material influence on the result. The sensitivity analyses also reveal important qualifications: the first position can switch between TLP5 and TLP2 under different theta or weight settings, probability perturbation affects the top order, and deletion of an alternative may produce rank reversal. The scenario-based numerical experiments further indicate that the procedure remains stable under several synthetic settings with different numbers of alternatives, criteria, and probability distributions, although exact full-order retention becomes weaker when the scale and probability dispersion increase.

Future work should collect direct PHF-BWM judgments, test more real cases, compare the adjustment with additional consistency-repair algorithms, and further study ranking procedures with greater robustness to rank reversal.

Acknowledgments

None.

Funding

This work was funded by the Science Research Project of Hebei Education Department (No. BJK2024132).

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Fei Wang
Formal analysis: Baoyi Zhang
Investigation: Baoyi Zhang
Methodology: Fei Wang, Baoyi Zhang
Writing–original draft: Baoyi Zhang, Fei Wang
Writing–review & editing: Fei Wang, Siyuan Liu

Availability of data

The data used in this study are derived from a published case study by Krishankumar *et al.*[34] The relevant data are publicly available in the cited reference and are also reported in the present manuscript. No new datasets were generated in this work. Further information can be obtained from the corresponding author upon reasonable request.

AI tools statement

AI tools were used to assist with language polishing and reference formatting during the preparation of this manuscript. All authors reviewed and edited the content and take full responsibility for the final version of the manuscript.

References

1. Garg H, Krishankumar R, Ravichandran KS. Decision framework with integrated methods for group decision-making under probabilistic hesitant fuzzy context and unknown weights. *Expert Syst Appl.* 2022;200:117082.

   https://doi.org/10.1016/j.eswa.2022.117082

2. Fang B. Some uncertainty measures for probabilistic hesitant fuzzy information. *Inf Sci.* 2023;625:255-276.

   https://doi.org/10.1016/j.ins.2022.12.101

3. Aydoğan H, Özkır V. A Fermatean fuzzy MCDM method for selection and ranking problems: Case studies. *Expert Syst Appl.* 2024;237:121628.

   https://doi.org/10.1016/j.eswa.2023.121628

4. Zadeh LA. Fuzzy sets. *Inf Control.* 1965;8(3):338-353.

https://doi.org/10.1016/S0019-9958(65)90241-X

5. Torra V. Hesitant fuzzy sets. *Int J Intell Syst.* 2010;25(6):529-539.

   https://doi.org/10.1002/int.20418

6. Xu Z, Zhou W. Consensus building with a group of decision makers under the hesitant probabilistic fuzzy environment. *Fuzzy Optim Decis Mak.* 2017;16(4):481-503.

   https://doi.org/10.1007/s10700-016-9257-5

7. Greco S, Słowiński R, Wallenius J. Fifty years of multiple criteria decision analysis: From classical methods to robust ordinal regression. *Eur J Oper Res.* 2025;323(2):351-377.

   https://doi.org/10.1016/j.ejor.2024.07.038

8. Das S, Dutta R, De S, De S. Review of multi-criteria decision-making for sustainable decentralized hybrid energy systems. *Renew Sustain Energy Rev.* 2024;202:114676.

   https://doi.org/10.1016/j.rser.2024.114676

9. Ortiz-Barrios M, Cabarcas-Reyes J, Ishizaka A, Barbati M, Jaramillo-Rueda N, Carrascal-Zambrano GDJ. A hybrid fuzzy multi-criteria decision making model for selecting a sustainable supplier of forklift filters: A case study from the mining industry. *Ann Oper Res.* 2021;307(1-2):443-481.

   https://doi.org/10.1007/s10479-020-03737-y

10. Ding Q, Zhou Y, Cheng TCE, Ji M. A decision framework for selecting emergency logistics suppliers based on an extended PHFS-TODIM method. *Int J Fuzzy Syst.* 2025;27(5):1498-1513.

    https://doi.org/10.1007/s40815-024-01848-3

11. Liu Y, Guan X, Wu B. Research on decision-making method based on probabilistic hesitant fuzzy comprehensive distance measurement. *J Northwest Polytech Univ.* 2023;41(6):1209-1220.

    https://doi.org/10.1051/jnwpu/20234161209

12. Liao N, Wei G, Chen X. TODIM method based on cumulative prospect theory for multiple attributes group decision making under probabilistic hesitant fuzzy setting. *Int J Fuzzy Syst.* 2022;24(1):322-339.

    https://doi.org/10.1007/s40815-021-01138-2

13. Zhou Y, Zhou L, Lin Z, Xu X. Probabilistic q-rung hesitant fuzzy TODIM method and its application. *J Shandong Univ Nat Sci.* 2023;58(6):9-17.

    https://doi.org/10.6040/j.issn.1671-9352.0.2022.260

14. Leoneti AB, Gomes LFAM. A novel version of the TODIM method based on the exponential model of prospect theory: The ExpTODIM method. *Eur J Oper Res.* 2021;295(3):1042-1055.

    https://doi.org/10.1016/j.ejor.2021.03.055

15. Divsalar M, Ahmadi M, Ghaedi M, Ishizaka A. An extended TODIM method for hyperbolic fuzzy environments. *Comput Ind Eng.* 2023;185:109655.

    https://doi.org/10.1016/j.cie.2023.109655

16. Liu Y, Qin Y, Liu H, Abdullah S, Rong Y. Prospect theory-based q-rung orthopair fuzzy TODIM method for risk assessment of renewable energy projects. *Int J Fuzzy Syst.* 2024;26(3):1046-1068.

    https://doi.org/10.1007/s40815-023-01652-5

17. Więckowski J, Kizielewicz B, Shekhovtsov A, Sałabun W. RANCOM: A novel approach to identifying criteria relevance based on inaccuracy expert judgments. *Eng Appl Artif Intell.* 2023;122:106114.

    https://doi.org/10.1016/j.engappai.2023.106114

18. Więckowski J, Kizielewicz B, Sałabun W. Fuzzy RANCOM: A novel approach for modeling uncertainty in decision-making processes. *Inf Sci.* 2025;694:121716.

    https://doi.org/10.1016/j.ins.2024.121716

19. Zakeri S, Konstantas D, Chatterjee P, Zavadskas EK. Soft cluster-rectangle method for eliciting criteria weights in multi-criteria decision-making. *Sci Rep.* 2025;15:284.

    https://doi.org/10.1038/s41598-024-81027-4

20. Yang W, Liang C. A large-scale consensus decision-making model for non-cooperative behavior based on incomplete probabilistic hesitant fuzzy information in social trust networks. *Inf Sci.* 2025;714:122196.

    https://doi.org/10.1016/j.ins.2025.122196

21. Yang G, Ren M, Hao X. Multi-criteria decision-making problem based on the novel probabilistic hesitant fuzzy entropy and TODIM method. *Alex Eng J.* 2023;68:437-451.

    https://doi.org/10.1016/j.aej.2023.01.014

22. Saha A, Debnath BK, Chatterjee P, Panaiyappan AK, Das S, Anusha G. Generalized Dombi-based probabilistic hesitant fuzzy consensus reaching model for supplier selection under healthcare supply chain framework. *Eng Appl Artif Intell.* 2024;133:107966.

    https://doi.org/10.1016/j.engappai.2024.107966

23. Wang C, Zhou J, Yang Y, Zhou L. Probabilistic q-rung hesitant fuzzy TOPSIS method and its application. *Fuzzy Syst Math.* 2025;39(4):129-139.

24. Rezaei J. Best-worst multi-criteria decision-making method. *Omega.* 2015;53:49-57.

    https://doi.org/10.1016/j.omega.2014.11.009

25. Guo S, Zhao H. Fuzzy best-worst multi-criteria

decision-making method and its applications. *Knowl Based Syst.* 2017;121:23-31.

https://doi.org/10.1016/j.knosys.2017.01.010

26. Ali A, Rashid T. Hesitant fuzzy best-worst multi-criteria decision-making method and its applications. *Int J Intell Syst.* 2019;34(8):1953-1967.

https://doi.org/10.1002/int.22131

27. Xu D, Li W, Ren X, Shen W, Dong L. Technology selection for sustainable hydrogen production: A multi-criteria assessment framework under uncertainties based on the combined weights and interval best-worst projection method. *Int J Hydrogen Energy.* 2020;45(59):34396-34411.

https://doi.org/10.1016/j.ijhydene.2019.09.030

28. Wan S, Dong J, Chen SM. Fuzzy best-worst method based on generalized interval-valued trapezoidal fuzzy numbers for multi-criteria decision-making. *Inf Sci.* 2021;573:493-518.

https://doi.org/10.1016/j.ins.2021.03.038

29. Ali A, Rashid T. Generalized interval-valued trapezoidal fuzzy best-worst multiple criteria decision-making method with applications. *J Intell Fuzzy Syst.* 2020;38(2):1705-1719.

https://doi.org/10.3233/JIFS-182932

30. Mou Q, Xu Z, Liao H. An intuitionistic fuzzy multiplicative best-worst method for multi-criteria group decision making. *Inf Sci.* 2016;374:224-239.

https://doi.org/10.1016/j.ins.2016.08.074

31. Deveci M, Özcan E, John R, Pamucar D, Karaman H. Offshore wind farm site selection using interval rough numbers based Best-Worst Method and MARCOS. *Appl Soft Comput.* 2021;109:107532.

https://doi.org/10.1016/j.asoc.2021.107532

32. Li J, Wang J, Hu J. Multi-criteria decision-making method based on dominance degree and BWM with probabilistic hesitant fuzzy information. *Int J Mach Learn Cybern.* 2019;10(7):1671-1685.

https://doi.org/10.1007/s13042-018-0845-2

33. Liu M, Zhang X, Mo Z. Probabilistic hesitant fuzzy MEREC-TODIM decision-making based on improved distance measures. *Int J Fuzzy Syst.* 2024;26(7):2370-2393.

https://doi.org/10.1007/s40815-024-01741-z

34. Amjad H, Stoklasa J, Bhat SA. Rank reversal free VIKOR and fuzzy VIKOR—Heading toward absolute-type fuzzy multiple-criteria evaluation methods. *IEEE Trans Fuzzy Syst.* 2026;34(9):2978-2992.

https://doi.org/10.1109/TFUZZ.2026.3706988

35. Krishankumar R, Mishra AR, Rani P, *et al.* Two-stage EDAS decision approach with probabilistic hesitant fuzzy information. *Informatica.* 2025;36(1):65-97.

https://doi.org/10.15388/24-INFOR577

36. Ding Q, Wang YM, Goh M. TODIM dynamic emergency decision-making method based on hybrid weighted distance under probabilistic hesitant fuzzy information. *Int J Fuzzy Syst.* 2021;23(2):474-491.

https://doi.org/10.1007/s40815-020-00978-8

37. Bhowmik C, Zindani D, Chatterjee P, Marinković D, Šliogerienė J. Evaluation of green energy sources: An extended fuzzy-TODIM approach based on Schweizer-Sklar and power averaging operators. *Facta Univ Ser Mech Eng.* 2025;23(3):627-648.

https://doi.org/10.22190/FUME240711042B

38. Liu Y, Guan X. Probabilistic hesitant fuzzy recognition method based on comprehensive characteristic distance measure. *Math Probl Eng.* 2021;2021:1738026.

https://doi.org/10.1155/2021/1738026

39. Su Z, Xu ZS, Zhao H, Hao Z, Chen B. Entropy measures for probabilistic hesitant fuzzy information. *IEEE Access.* 2019;7:65714-65727.

https://doi.org/10.1109/ACCESS.2019.2916564

40. Fang B, Han B, Wen C. [Probabilistic hesitant fuzzy multi-attribute group decision-making based on new distance measure]. *Control Decis.* 2022;37(3):729-736. [In Chinese]

https://doi.org/10.13195/j.kzyjc.2020.1118

Fei Wang *Fei Wang received his Ph.D. degree from the School of Economics and Management, Yanshan University, in 2023. He is currently a lecturer at the School of Management Science and Engineering, Hebei University of Economics and Business. His research focuses on intelligent decision-making methods and their applications, including supply chain operation optimization, engineering bid evaluation, risk management in the low-altitude economy, and productivity evaluation. He has published over ten papers as first or corresponding author in high-impact domestic and international journals, including four sole-author top-tier SCI Q1 papers in journals such as Applied Soft Computing, Information Sciences, and International Journal of Intelligent Systems. He has led projects funded by the Hebei Social Science Foundation (Youth Program), the Hebei Education Department (Top Young Talent Program), and a university teaching research project. He was honored as an Excellent Teacher and Research Rising Star of Hebei University of Economics and Business in 2025.*
https://orcid.org/0000-0002-1917-820X

Baoyi Zhang *Baoyi Zhang is currently a master's student*

at the School of Management Science and Engineering, Hebei University of Economics and Business.
https://orcid.org/0009-0004-3334-7657

Siyuan Liu *Siyuan Liu is currently a master's student at the School of Management Science and Engineering, Hebei University of Economics and Business.*
https://orcid.org/0009-0008-6408-5062

An International Journal of Optimization and Control: Theories & Applications (https://accscience.com/journal/ijocta)



This work is licensed under a Creative Commons Attribution 4.0 International License. The authors retain ownership of the copyright for their article, but they allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in IJOCTA, so long as the original authors and source are credited. To see the complete license contents, please visit http://creativecommons.org/licenses/by/4.0/.