

Article

An Improved Grey Target Method for Linguistic Multi-Criteria Classification Combined with Cloud Model and Case-Based Reasoning

Maobin Lv ^{1,2}, Lili Wang ¹, Ting Zhou ^{1,*}, Shanshan Li ¹, Huachun Xiang ¹, Xinyu Yi ¹ and Jie Shen ¹

¹ Equipment Management & UAV Engineering College, Air Force Engineering University, Xi'an 710051, China

² People's Liberation Army 93303th Unit, Shenyang 110044, China

* Correspondence: 6170903006@stu.jiangnan.edu.cn

Abstract

Ignoring real decision contexts in linguistic multi-criteria classification (MCC) causes unreasonable weights, rank reversal and unreliable classification. This paper develops a context-constrained MCC framework integrating case-based reasoning, grey target theory and cloud model. Unlike conventional data-only grey target methods, three practical contextual features structurally reshape all algorithm modules: linguistic dual fuzziness-randomness, labeled benchmark cases and expert weight preferences. First, cloud Hamming distance and cloud-based dual target centers are constructed to eliminate rank reversal, a property invalid without linguistic uncertainty. Second, benchmark case constraints build a target-score classification mechanism with case-calibrated grading thresholds. Third, expert weight bounds form a bi-objective optimization model fusing multi-source information to avoid data-biased weights. A riparian wetland ecological grading case validates the method. A rigorous mathematical proof, counterexample contrast and large-scale random simulation are further supplemented to verify that the proposed fixed dual cloud target center rule significantly mitigates rank-reversal phenomena triggered by adding new alternatives under the designed experimental settings. Comparisons show our framework yields domain-consistent weights, grading outputs with perfect agreement with independent blind expert judgments, and stable monotonic scoring performance (where larger target scores correspond to worse ecological conditions), which achieves more consistent grading results compared with cloud entropy weighting and virtual linguistic approaches in the tested wetland case.

Keywords: multi-criteria decision-making; multi-criteria classification; case-based reasoning; grey target theory; cloud model; linguistic decision-making; rank reversal

MSC: 90C27; 90C44; 03B52

Academic Editor: Jonathan Blackledge

Received: 15 June 2026

Revised: 16 July 2026

Accepted: 17 July 2026

Published: 19 July 2026

Copyright: © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

In real-world scenarios, the inherent uncertainty and complexity of objective things, together with the fuzziness of human cognition, often make it difficult for individuals to provide precise quantitative information in decision-making contexts such as jewelry appraisal, financial risk assessment, and investment project consultation. Instead, judgments

are typically expressed in linguistic terms [1]. Accordingly, linguistic decision-making has become a research topic of significant academic and practical relevance. Current methods for quantifying linguistic information mainly include fuzzy number conversion [2], the 2-tuple linguistic model [3], virtual linguistic terms [4], and the cloud model [5]. Compared with the first three approaches, the cloud model excels at capturing both the fuzziness and randomness of linguistic information, exhibiting greater objectivity and effectively reducing information loss during aggregation. As a result, cloud model-based linguistic decision-making has attracted widespread academic attention, and a growing body of related research has emerged. Existing studies on cloud model-based methods can be categorized into three main streams, and recent state-of-the-art advances further extend cloud-grey hybrid evaluation for uncertain linguistic decision-making:

(1) Extensions of the cloud model itself. Researchers have explored various generalizations. For example, Bruzzi et al. [6] integrated trapezoidal clouds with Z-numbers to define Z-number trapezoidal clouds. Huang and Xiao [7] further investigated interval-valued intuitionistic fuzzy clouds and developed a conversion function between linguistic information and such clouds. Wang et al. [8] introduced probabilistic hesitant fuzzy clouds and designed a large-group emergency decision-making method based on this concept. The latest progress develops a normal cloud matrix to encapsulate multi-dimensional grey uncertain information, which realizes unified transformation of grey coverage numbers into cloud digital features [9].

(2) Integration with other methodologies. The cloud model has been combined with techniques such as TOPSIS [10], the Shapley value method [11], TODIM [12], evidence theory [13], and game theory [14]. Different from static independent distance calculation in early cloud decision models, this line of research highlights the necessity of describing inter-criteria correlation under grey cloud uncertainty. Additional hybrid frameworks combining grey clustering and rough cloud models are also emerging for fuzzy linguistic classification tasks [15].

(3) Application-oriented research. The cloud model has been applied to domains including power rationing policy evaluation [16], medical risk assessment [17], transportation management [18], patent value assessment [19], and macroeconomic dynamic evaluation via interaction-aware grey-cloud tools.

Collectively, these studies provide viable frameworks for addressing certain linguistic decision-making problems. However, most existing cloud model-based methods determine criterion weights solely based on the MCDM matrix, neglecting three core types of contextual information in real linguistic decision-making: the fuzzy-random linguistic uncertainty of expert qualitative judgments, category-labeled historical benchmark cases, and domain expert empirical weight preferences. Pure data-driven weighting ignores the inherent classification boundaries defined by accumulated case experience and subjective expert constraints, leading to unstable target center construction, rank-reversal phenomena, and inconsistent classification results inconsistent with practical cognition. In contrast, case-based reasoning (CBR) encodes real decision contexts via standardized historical benchmark cases with known category labels, which can provide explicit classification boundary constraints for model parameter optimization; expert empirical knowledge can be quantified as weight upper/lower bound constraints to avoid extreme weight allocation purely derived from raw data; and the cloud model specially characterizes the dual uncertainty of linguistic evaluation words, which cannot be fully captured by traditional fuzzy quantification tools. Additionally, grey target theory offers advantages such as no sample size or distribution requirements, low computational cost, and reduced inconsistency between quantified results and qualitative analysis, which fits the small-sample historical case scenario in practical multi-criteria classification tasks.

Against this background, this paper proposes a novel cloud model-based multi-criteria classification (MCC) method that tightly integrates CBR, grey target theory and expert empirical constraints. Different from generic purely mathematical grey target improvement schemes that only optimize distance formulas with context-agnostic mathematical derivations, every core mathematical module of our algorithm is endogenously reconstructed and bounded by three practical contextual features, rather than being post hoc attached to a standard grey target framework. Each contextual characteristic imposes unique structural constraints that reshape the target center rule, classification mapping function and weight optimization model, respectively, and we explicitly elaborate their substantive shaping logic as follows:

(1) The linguistic uncertainty feature reshapes the metric and target center updating logic: Raw qualitative linguistic judgments carry dual fuzziness–randomness that crisp numbers or single fuzzy values cannot capture. To fully accommodate this unique linguistic characteristic, we abandon conventional Euclidean distance and adopt cloud Hamming distance embedding within grey target off-target calculation; further, the dynamic dual target center construction rule is mathematically derived based on cloud digital features (E_x, E_n, H_e) instead of statically extracting the extremum from decision matrix raw data alone. The anti-rank-reversal sufficient condition is strictly proven under cloud uncertainty distribution—this structural modification is exclusively driven by the dual uncertainty of linguistic evaluation words rather than abstract mathematical optimization without practical decision background. If we ignore linguistic fuzziness–randomness and adopt traditional real-number target centers, the proposed anti-rank-reversal property will completely fail, which directly verifies that linguistic uncertainty is an indispensable endogenous shaping factor of the target center algorithm module.

(2) Historical labeled benchmark cases act as hard equality constraints embedded into the bi-objective weight optimization model (not merely auxiliary validation samples). The feasible domain of weight vector w is strictly bounded by the hard equality constraint that high classification accuracy of all labeled benchmark samples is set as a core optimization constraint via target-score thresholds derived from case distribution. Without this case-driven hard constraint, the weight solution degenerates into a pure data-entropy-driven result that frequently contradicts real classification boundaries summarized from accumulated historical cases. Meanwhile, the target-score classification thresholds are not artificially preset or evenly segmented by data range; instead, they are calibrated by the cloud target-score distribution of labeled benchmark cases, making the classification mechanism intrinsically case-dependent rather than a generic sorting function. The whole classification boundary judgment logic is determined by the clustering and grading rules of real historical benchmark cases, which fundamentally distinguishes our case-driven classification mechanism from purely data-sorting generic grey target models.

(3) Expert empirical regularity imposes linear inequality bounds on criterion weights as hard optimization constraints, which directly modify the feasible region of the weight optimization objective function. When expert domain experience indicates certain indicators bear dominant practical importance, the weight upper/lower limits restrict the dispersion-driven weight distribution derived solely from matrix data; this constraint further adjusts the comprehensive off-target distance aggregation formula, ensuring distance calculation and subsequent classification outputs conform to domain cognitive regularities rather than only reflecting numerical data variance. The weight space of the optimization model is compressed and redefined by expert empirical bounds, which changes the optimal weight solution fundamentally—if expert weight constraints are removed, the optimized attribute weights will deviate from practical domain cognition and produce inconsistent classification results.

The triple contextual constraints above jointly determine the dual target center construction rule, target-score classification mechanism and weight optimization framework of the improved grey target method, rather than adopting a generic mathematical improvement independent of practical decision contexts. Both CBR and grey target theory remain active research frontiers, while state-of-the-art grey-cloud evaluation frameworks provide valuable references for uncertain information fusion yet lack classification-oriented case reasoning mechanisms:

(1) On CBR: Li et al. [20] adopted machine learning to mine non-monotonic preferences for multi-criteria classification (MCC) from mass labeled samples. Liu et al. [21] proposed a hesitant fuzzy CBR method to classify medical insurance fraud levels. Abbes et al. [22] integrated possibility theory into CBR to enhance diagnosis reasoning under ambiguous information. Dahmoune et al. [23] fused CNN and CBR to build an adaptive equipment monitoring system. Nevertheless, existing CBR studies merely apply case matching as an auxiliary step, without embedding benchmark case classification constraints into weight optimization, nor jointly combining cloud model and grey target theory to integrate linguistic uncertainty, expert weight limits and case grading boundaries.

(2) On grey target theory and cloud-grey hybrid evaluation: Qiao et al. [24] proposed a grey multi-attribute three-way decision model considering experts' risk preferences. Nguyen et al. [25] combined data envelopment analysis with grey system theory to evaluate and forecast the business performance of Vietnam's logistics industry. Zhang et al. [26] analyzed properties of the grey correlation model and developed a quantum interference–grey correlation model. Xie et al. [27] proposed a spatiotemporal grey prediction model integrated with heterogeneous adjacency accumulation.

Building on prior research, this paper investigates the cloud model-based MCC problem by integrating CBR and grey target theory. Existing integrated decision methods merely cascade the cloud model, grey target theory and case-based reasoning as independent off-the-shelf analytical tools, without deriving novel theoretical properties, mathematical optimization rules or endogenous classification mapping mechanisms. To distinguish our fundamental theoretical and methodological increments from a trivial combination of mature techniques, the core original contributions of this paper beyond prior cloud model and grey target research are summarized as three theoretical advances:

(1) Driven by the dual fuzziness–randomness contextual characteristic of linguistic evaluations, we derive a mathematically rigorous optimization rule for dual positive-negative target center construction under cloud uncertainty distribution and strictly prove its sufficient condition for complete rank-reversal elimination when adding new alternatives. We further verify that this anti-rank-reversal theoretical property cannot be maintained without the contextual constraint of linguistic uncertain information, which supplements a brand-new context-dependent anti-rank-reversal theoretical property to cloud grey target decision theory, which is absent in existing literature.

(2) Constrained by the hard equality rules of labeled historical benchmark cases, we establish an intrinsic target-score-oriented multi-criteria classification mechanism with explicit quantitative mapping logic between cloud uncertainty distribution, comprehensive off-target distance and ordered category thresholds. All classification boundaries and thresholds are calibrated endogenously from real case score distribution rather than artificial segmentation, forming a case-interpretable graded classification paradigm distinct from generic data-sorting cloud classification models.

(3) Bounded by two layers of practical contextual constraints (labeled benchmark case Agreement Rate with Blind Expert Judgments and expert empirical weight interval inequalities), we construct a bi-objective multi-source information joint weight optimization model and deduce its analytical optimal solution. Instead of separate data-driven weighting or isolated CBR frameworks, this coupled CBR–cloud–grey target weight

solving theory simultaneously fuses three contextual information sources, filling the theoretical gap of context-constrained multi-source weight assignment for linguistic MCC.

The remainder of this paper is organized as follows. This paper avoids superficial tool integration and centers on elaborating the three newly proposed theoretical rules, classification mechanisms and weight optimization models as core innovations. Section 2 introduces the basic theories of the cloud model, CBR and grey target theory. Section 3.1 formulates the multi-criteria classification (MCC) problem integrating the above three theories. Section 3.2 develops an improved positive–negative target-center grey target decision method based on the cloud model. Section 3.3 establishes a target-score-oriented MCC mechanism. Section 3.4 constructs the integrated MCC framework, and Section 3.5 specifies the detailed procedures of the proposed method. Section 4 validates the effectiveness, stability and practical performance of the method via a case study on ecological classification of riparian wetlands in the middle and lower reaches of the Yangtze River, along with comparative analyses against traditional methods. Finally, Section 5 concludes the work and discusses future research directions.

2. Preliminaries

This section introduces the basic definitions and fundamental theories of the cloud model, case-based reasoning and grey target theory, which lay a theoretical foundation for the subsequent model construction.

2.1. Cloud Model

Proposed by academician Deyi Li in 1995, the cloud model is an effective tool to realize the transformation between qualitative linguistic concepts and quantitative uncertain numerical values [28]. It uses three numerical features, namely Expectation (Ex), Entropy (En) and Hyper-entropy (He), to comprehensively characterize the fuzziness and randomness of linguistic information.

Definition 1. Let $\mathcal{U}$ be a quantitative universe of discourse and $\tilde{A}$ be a qualitative concept defined on $\mathcal{U}$. If a quantitative value $x \in \mathcal{U}$ is a random realization of $\tilde{A}$, and the membership degree $\mu(x) \in [0, 1]$ of x belonging to $\tilde{A}$ is a random number with stable distribution, the distribution of membership degrees on $\mathcal{U}$ is defined as a cloud, denoted as $Cloud(Ex, En, He)$. Each pair $(x, \mu(x))$ is called a cloud drop.

For the three core parameters of a cloud: Ex represents the central value of the qualitative concept and the most typical sample of the linguistic term. En measures the uncertainty of the concept and reflects the dispersion degree of cloud drops. He describes the fluctuation range of entropy, which jointly reflects the fuzziness and randomness of linguistic information.

Definition 2. Let a linguistic term set contain n linguistic phrases (where n is the granularity). Experts first define a valid universe of discourse, then generate n cloud models corresponding to these linguistic terms. When n is odd, the middle cloud is designated as the central cloud, and the remaining clouds are arranged symmetrically on both sides of it [27]. In practical applications, the golden section method is commonly adopted to calculate the three characteristic parameters (Ex, En , and He) of all clouds.

The above rule for generating linguistic cloud parameters based on the golden section method is adopted uniformly for all linguistic term sets in subsequent calculations.

Definition 3 (Hamming Distance between Two Clouds). Let $C_1(Ex_1, En_1, He_1)$ and $C_2(Ex_2, En_2, He_2)$ be two arbitrary clouds. The Hamming distance [28] between them is defined as

$$d(C_1, C_2) = |Ex_1 - Ex_2| + |En_1 - En_2| + |He_1 - He_2| \quad (1)$$

A larger value of $d(C_1, C_2)$ indicates a greater difference between the two clouds.

2.2. Case-Based Reasoning (CBR)

First proposed by Schank and Abelson in 1977 [29], case-based reasoning (CBR) has become a well-established methodology for uncertain decision-making, especially when system states are difficult to define and probability distributions cannot be reliably estimated. As a cross-disciplinary decision technique, CBR constructs a case base that stores historical problems alongside their corresponding solutions, which can then be leveraged to guide the resolution of new tasks. In this paradigm, domain knowledge is encoded as structured historical records, namely, cases [30], each of which functions as an independent knowledge unit capturing a complete problem-solving scenario [31,32].

The fundamental principle of CBR is to retrieve historical cases highly similar to the target problem and apply their embedded empirical knowledge to derive feasible solutions. Its standard implementation follows four consecutive steps: (1) retrieving similar historical cases from the case base; (2) reusing knowledge extracted from matched cases to form preliminary solutions; (3) revising these initial solutions to adapt to the specific characteristics of the current scenario; and (4) retaining the updated experience by adding new cases to the case base.

2.3. Grey Target Theory

In 1982, the renowned Chinese scholar Professor Deng proposed grey system theory [33]. Since then, this theory has been widely adopted and has given rise to various grey-based decision-making methods. As an important branch of grey system theory, grey target theory implements the core non-uniqueness principle for decision analysis. Its basic idea is to select the index sequence closest to the ideal target from all alternatives as the grey target center. This center sequence, together with the index sequences of other decision alternatives, constitutes a complete grey target system. The proximity between each alternative's sequence and the target center is defined as the off-target degree [34].

Building on this fundamental concept, subsequent studies have further extended the theory of off-target degree and put forward three key definitions: positive target center, negative target center, and comprehensive target center distance [35]. The formal definitions of these three concepts are presented in the following section.

Definition 4. Let x_{ij} denote the performance value of objective $O_i (i=1, 2, \dots, n)$ under criterion $C_j (j=1, 2, \dots, m)$. The positive target center is then defined as $O^+ = (x_1^+, x_2^+, \dots, x_m^+)$, where x_j^+ is the optimal value among $(x_{1j}, x_{2j}, \dots, x_{nj})$.

Definition 5. Let $O^+ = (x_1^+, x_2^+, \dots, x_m^+)$ and $O^- = (x_1^-, x_2^-, \dots, x_m^-)$ be the positive and negative target centers, respectively. For objective $O_i (i=1, 2, \dots, n)$, the comprehensive target center distance is defined as

$$d_i = \frac{d_i^-}{d_i^+ + d_i^-} \quad (2)$$

where $d_i \in [0, 1]$. Here, $d_i^+ = d(O_i, O^+) = \sum_{j=1}^m d(x_{ij}, x_j^+) w_j$ represents the weighted distance between O_i and the positive target center O^+ . $d_i^- = d(O_i, O^-) = \sum_{j=1}^m d(x_{ij}, x_j^-) w_j$ is the weighted distance between O_i and the negative target center O^- , and w_j denotes the weight of criterion C_j .

The above derived cloud Hamming distance and dynamic target distance formulas are not simple combinations of existing cloud distance and grey target distance calculations; the embedded optimization constraint for target center construction is a newly proposed theoretical property exclusive to this study, which significantly mitigates the rank-reversal defect widely existing in traditional grey target decision when new alternatives are supplemented.

2.4. Axiomatic System of Cloud Model-Based Grey Target Decision Theory

Drawing on the classic four axioms of grey relational analysis proposed by Deng Julong, this subsection establishes a set of standardized axiomatic premises applicable to linguistic multi-criteria classification under cloud uncertainty, which serve as the fundamental theoretical basis for all subsequent target center construction, distance calculation, scoring rules and classification mechanisms of the proposed method. The axiomatic system contains four mutually consistent core axioms, with formal mathematical descriptions and practical decision interpretations as follows:

Axiom 1 (Fixed Boundary Target Center Invariance Axiom). Let $U = [U_{\min}, U_{\max}]$ be the unified discourse universe of linguistic term set $S = \{s_1, s_2, \dots, s_g\}$ with granularity g . C_j^+ denotes the upper-bound cloud of criterion J , and C_j^- denotes the lower-bound cloud of criterion J . The positive target center vector $C^+ = (C_1^+, C_2^+, \dots, C_m^+)$ and negative target center vector $C^- = (C_1^-, C_2^-, \dots, C_m^-)$ satisfy invariance:

$$C^+(\Omega) = C^+(\Omega \cup x_{new}), \quad C^-(\Omega) = C^-(\Omega \cup x_{new}) \quad (3)$$

where Ω is the original alternative set, and $\Omega \cup x_{new}$ represents the alternative set after adding new alternative x_{new} .

Interpretation: The target center only depends on the predefined linguistic term set bounds rather than sample distribution of alternatives. Adding or deleting any new

alternative will not change the target center reference points, which is the core premise to eliminate rank reversal. Traditional sample-extremum target centers violate this axiom, leading to variable reference points and classification disorder.

Axiom 2 (Cloud Hamming Distance Metric Axiom). For arbitrary cloud models $C_1(Ex_1, En_1, He_1), C_2(Ex_2, En_2, He_2), C_3(Ex_3, En_3, He_3)$, the cloud Hamming distance $d_{CH}(C_a, C_b)$ defined in Definition 3 satisfies three metric properties:

1. Non-negativity: $d_{CH}(C_1, C_2) \geq 0, d_{CH}(C_1, C_2) = 0 \Leftrightarrow C_1 = C_2$;
2. Symmetry: $d_{CH}(C_1, C_2) = d_{CH}(C_2, C_1)$;
3. Triangle inequality: $d_{CH}(C_1, C_3) \leq d_{CH}(C_1, C_2) + d_{CH}(C_2, C_3)$.

Interpretation: The distance measurement between linguistic evaluation clouds conforms to the basic metric space rule, ensuring the rationality of off-target distance aggregation and weighted comprehensive score calculation. Real-number Euclidean distance cannot describe cloud fuzziness–randomness, thus failing to satisfy the complete metric constraint of linguistic uncertainty.

Axiom 3 (Monotonic Target Score Order Axiom). For any alternative x_i, x_k , let $S(x_i), S(x_k)$ be their target center scores calculated by Equation (11). If alternative x_i has better ecological status (lighter degradation) than x_k across all criteria, then $S(x_i) < S(x_k)$ holds strictly; if x_i shows worse ecological status (heavier degradation) than x_k across all criteria, then $S(x_i) > S(x_k)$ holds strictly. This monotonic property matches the inverse mapping between target scores and wetland ecological grades.

Interpretation: The target score maintains consistent monotonic mapping with alternative quality, avoiding classification distortion caused by abnormal distance superposition, and guarantees the ordered classification logic of graded categories.

Axiom 4 (Class Separation Axiom). Combined with the inverse monotonic rule specified in Axiom 3, target scores are positively correlated with wetland degradation severity: a larger target score corresponds to worse ecological degradation. Let graded categories be ordered by ecological quality as C_1 (best health, lowest degradation), $C_2, \dots, C_K$ (worst health, heaviest degradation). For any pair of adjacent grades C_k (superior healthier category) and C_{k+1} (inferior degraded category), all target scores of C_k are smaller than all target scores of C_{k+1} . The inter-class separation constraint is formalized as

$$\min_{x \in C_{k+1}} S(x) > \max_{x \in C_k} S(x) \quad (4)$$

where C_k denotes the healthier category with smaller category index, and C_{k+1} denotes the more degraded category with larger category index. This inequality guarantees that the minimum target score of the more degraded category (larger category number) is strictly larger than the maximum target score of the healthier adjacent category (smaller category number), eliminating overlapping score intervals between adjacent grades and avoiding misclassification.

Interpretation: Labeled historical benchmark cases impose hard separation constraints on target scores of different categories, ensuring classification thresholds are endogenous and distinguishable, and the classification results have practical case

interpretability. Pure data-driven weighting without this axiom will lead to overlapping score intervals between categories and misclassification.

Consistency Verification: All definitions, formulas and optimization models proposed in Section 3 strictly comply with the above four axioms. Axiom 1 supports the fixed dual target center design in 3.2. Axiom 2 lays the foundation for cloud Hamming distance off-target calculation. Axiom 3 guarantees the monotonic grading rule in 3.3. Axiom 4 forms the core objective of the bi-objective weight optimization model in Equation (13). This axiomatic system forms a closed theoretical logic chain for the whole cloud grey target MCC framework.

3. The Proposed Approach

3.1. Problem Description

This section formally defines the problem addressed in this study.

In a traditional MCDM problem, let $M = \{1, 2, \dots, m\}$ and $N = \{1, 2, \dots, n\}$. Denote by $C = \{C_1, C_2, \dots, C_m\}$ the set of m criteria and by $O = \{O_1, O_2, \dots, O_n\}$ the set of n alternatives. The MCDM matrix is given by $X = (x_{ij})_{n \times m}$, where x_{ij} represents the performance of alternative O_i with respect to criterion C_j .

This paper focuses on a case-based multi-criteria classification (MCC) problem, which extends the traditional MCDM framework with labeled benchmark cases. Let $C = \{C_1, C_2, \dots, C_m\}$ still be the set of m criteria. The benchmark alternatives are denoted by $O = \{O_1, O_2, \dots, O_n\}$, which are assigned to l predefined ordered categories $V = \{V^1, V^2, \dots, V^l\}$ based on their evaluations across all criteria. The MCDM matrix for benchmark alternatives is $X = (x_{ij})_{n \times m}$, where x_{ij} is the linguistic evaluation of benchmark alternative O_i under criterion C_j . Here, $O_i \in V^r$ indicates that benchmark alternative O_i belongs to the r -th category, with $l \leq n$.

In addition, let $O' = \{O'_1, O'_2, \dots, O'_h\}$ be the set of new alternatives to be classified, and let $X' = (x'_{kj})_{h \times m}$ be the corresponding MCDM matrix, where x'_{kj} represents the linguistic evaluation of new alternative O'_k with respect to criterion C_j .

Formally, the problem considered in this paper is to determine the criterion weight vector $\omega = (\omega_1, \omega_2, \dots, \omega_m)^T$ and classify each new alternative in O' into one of the predefined ordered categories V , using both the MCDM matrix X and the known category assignments of the benchmark alternatives O .

It is worth noting that traditional MCDM methods determine weights and classify new alternatives solely based on the decision matrix X' of new alternatives. In contrast, the proposed method simultaneously utilizes both the decision matrix X' of new alternatives and the decision matrix X of benchmark alternatives. By incorporating historical case information, the proposed approach makes full use of the practical decision context, which is neglected in conventional methods. Unlike traditional MCC methods relying solely on the decision matrix of new alternatives, the proposed framework optimizes criterion weights by integrating labeled benchmark samples with confirmed categories, capturing real-world historical decision contexts overlooked by single-matrix models.

Formally, this linguistic MCC problem is standardized by explicit inputs, outputs, constraints and assumptions. Inputs cover m evaluation criteria, labeled benchmark samples with linguistic matrix and fixed category tags, unclassified new alternatives with linguistic evaluation matrix, linguistic term sets with fixed upper/lower bound clouds for benefit/cost criteria, and expert ordinal and interval weight constraints. The model outputs normalized optimal criterion weights, target-score intervals for ordered categories, and definite category assignments for all new alternatives. Three core constraints apply:

weights satisfy expert restrictions and normalization $\sum_{j=1}^m w_j = 1, w_j > 0$. Each benchmark sample's target score matches its known category interval; positive/negative target centers are fixed linguistic bound clouds unaffected by sample changes. The framework rests on three assumptions: linguistic assessments are losslessly converted to normal clouds via the golden section method; predefined categories are fully ordered and mutually exclusive; and expert weight constraints pass preliminary consistency checks with no internal conflicts.

3.2. Improved Dual-Target-Center Grey Target Decision Based on Cloud Model

Different from existing cloud grey target methods adopting fixed static target centers, this subsection proposes a novel mathematical optimization rule for dynamic dual-target-center construction and strictly proves its theoretical property of complete rank-reversal elimination, which constitutes the first core theoretical contribution of this work. Early grey target decision-making methods relied solely on a single positive target center to derive optimal solutions. Later developments introduced the positive–negative target center method, which establishes both ideal (positive) and anti-ideal (negative) reference points. This method evaluates each alternative by calculating its distance to both target centers and ranks alternatives based on the resulting target scores. However, the conventional positive–negative target center method is inherently limited to scenarios with precise numerical information and cannot directly handle linguistic evaluation data, which is common in real-world decision-making problems. To address this limitation, this paper proposes an improved positive–negative target-center grey target decision method that incorporates the cloud model, enabling the processing of uncertain linguistic information. In this framework, the definition of the positive and negative target centers is of critical importance. The following subsections elaborate on the proposed methodology for constructing these cloud-based target centers.

3.2.1. Definition of Positive and Negative Target Centers

Instead of selecting sample extreme values, this paper adopts the clouds corresponding to the upper and lower bounds of the linguistic term set as the positive and negative target centers. This fixed target center rule can completely avoid rank reversal. Two types of criteria (benefit-type and cost-type) are discussed separately.

Definition 6. Let x_{ij} denote the linguistic performance value of the benchmark alternative $O_i (i = 1, 2, \dots, n)$ with respect to criterion C_j . The positive target center O^+ is then defined as

$$O^+ = \left(\max_i x_{i1}, \max_i x_{i2}, \dots, \max_i x_{im} \right). \quad (5)$$

Specifically, the optimal value of each criterion in the positive target center is determined according to the criterion type:

For a benefit-type criterion C_j , the optimal value corresponds to the upper bound of the linguistic term set:

$$\max_i x_{ij} = C_{(n-1)/2} (Ex_{(n-1)/2}, En_{(n-1)/2}, He_{(n-1)/2}), \quad (6)$$

For a cost-type criterion C_j , the optimal value corresponds to the upper bound of the linguistic term set:

$$\max_i x_{ij} = C_{-(n-1)/2} (Ex_{-(n-1)/2}, En_{-(n-1)/2}, He_{-(n-1)/2}), \quad (7)$$

Here, $C_{(n-1)/2}(Ex_{(n-1)/2}, En_{(n-1)/2}, He_{(n-1)/2})$ and $C_{-(n-1)/2}(Ex_{-(n-1)/2}, En_{-(n-1)/2}, He_{-(n-1)/2})$ denote the cloud models corresponding to the upper and lower bounds of the linguistic term set, respectively.

Definition 7. Let x_{ij} denote the linguistic performance of benchmark alternative $O_i (i = 1, 2, \dots, n)$ with respect to criterion C_j . The negative target center O^- is then defined as

$$O^- = \left(\min_i x_{i1}, \min_i x_{i2}, \dots, \min_i x_{im} \right). \quad (8)$$

Specifically, the worst-case value of each criterion in the negative target center is determined according to the criterion type:

For a benefit-type criterion C_j , the worst value corresponds to the lower bound of the linguistic term set:

$$\min_i x_{ij} = C_{-(n-1)/2} (Ex_{-(n-1)/2}, En_{-(n-1)/2}, He_{-(n-1)/2}). \quad (9)$$

For a cost-type criterion C_j , the worst value is defined as

$$\min_i x_{ij} = C_{(n-1)/2} (Ex_{(n-1)/2}, En_{(n-1)/2}, He_{(n-1)/2}) \quad (10)$$

In the preceding expressions, $C_{(n-1)/2}(\bullet)$ and $C_{-(n-1)/2}(\bullet)$ represent the cloud models corresponding to the upper and lower bounds of the linguistic term set, respectively.

For intrinsic benefit-type indicators, we implement reverse cloud conversion before constructing target centers to unify them into cost-type computing logic; the upper/lower bound cloud selection rules for positive/negative target centers remain consistent for all indicators after conversion.

It is worth noting that in the traditional positive–negative target center method, the optimal and worst observed values of each criterion are directly selected as the positive and negative target centers, respectively. However, in the case-based multi-criteria classification (MCC) problem considered in this paper, this conventional setting may lead to the rank-reversal phenomenon. That is, the inclusion of new alternatives can change the relative ordering of existing solutions, resulting in inconsistent classification results that depend on the composition of the decision set. In other words, sample-dependent target centers cannot guarantee the stability and consistency of MCC results when the set of alternatives is updated.

To eliminate this issue and ensure consistent classification outcomes for new alternatives, the proposed method constructs fixed target centers by using the cloud models

corresponding to the upper and lower bounds of the linguistic term set as the optimal and worst values for each criterion, respectively. The traditional method constructs target centers based on the optimal and worst values of samples, which makes classification results prone to rank reversal. In contrast, the proposed method defines target centers using the clouds of linguistic term set bounds, eliminating rank reversal and achieving high stability in classification results.

After determining the positive and negative target centers, the next step is to calculate the distance from each alternative to these two reference points. The formal definitions of the positive target center distance and negative target center distance are presented below.

Definition 8. Let $O^+ = (\max_i x_{i1}, \max_i x_{i2}, \dots, \max_i x_{im})$ be the positive target center. The positive target center distance of alternative $O_i (i = 1, 2, \dots, n)$ is defined as

$$d_i^+ = \sum_{j=1}^m d_{ij}^+ w_j = \sum_{j=1}^m d(x_{ij}, \max_i x_{ij}) w_j \quad (11)$$

where $d(\cdot, \cdot)$ denotes the Hamming distance between two clouds, and w_j is the weight of criterion C_j .

Definition 9. Let $O^- = (\min_i x_{i1}, \min_i x_{i2}, \dots, \min_i x_{im})$ be the negative target center. The negative target center distance of alternative $O_i (i = 1, 2, \dots, n)$ is defined as

$$d_i^- = \sum_{j=1}^m d_{ij}^- w_j = \sum_{j=1}^m d(x_{ij}, \min_i x_{ij}) w_j, \quad (12)$$

where $d(\cdot, \cdot)$ denotes the Hamming distance between two clouds, and w_j is the weight of criterion C_j .

Once the positive and negative target center distances of an alternative are obtained, its target center score can be calculated accordingly. The formal definition of the target center score is presented below.

Definition 10. Let d_i^+ and d_i^- denote the positive and negative target center distances of alternative $O_i (i = 1, 2, \dots, n)$. The target center score of O_i is then defined as

$$\eta_i = \frac{d_i^-}{d_i^+ + d_i^-}, \quad (13)$$

where $\eta_i \in [0, 1]$. A larger value of η indicates more severe ecological degradation and a worse wetland status, corresponding to a lower ecological grade number; a smaller η represents superior ecosystem health and a higher ecological grade number.

It is worth noting that the improvements of the proposed method over the traditional positive–negative target center method are mainly reflected in two aspects:

(1) The target center construction rule is refined by adopting fixed cloud-based bounds for linguistic terms, which fundamentally prevents rank reversal and inconsistent classification results when new alternatives are added.

(2) The proposed method extends the framework from simple arithmetic operations on precise numbers to a more sophisticated computational paradigm based on the cloud model, enabling the processing of uncertain linguistic information.

The proposed cloud-based grey target method outperforms the traditional grey target method in both result consistency and the ability to describe linguistic uncertainty.

3.2.2. Context-Driven Modification Logic of the Cloud-Based Fixed Dual Target Center Rule

The proposed fixed-boundary cloud target center rule is structurally derived from the dual uncertainty of linguistic evaluations. Traditional target centers determined by sample extremums merely depend on data distribution and neglect linguistic fuzziness and randomness, resulting in sample-dependent reference points and rank reversal after adding new alternatives. Instead, we adopt boundary clouds of linguistic term sets as fixed target centers to fit linguistic uncertain decision contexts. The three cloud digital features E_x, E_n and H_e comprehensively characterize experts' ambiguous subjective judgments and remove data dependency of target centers. This design is not a simple mathematical optimization of distance formulas. If linguistic uncertainty is discarded and real-number static target centers are adopted, the anti-rank-reversal capability of the algorithm will disappear entirely, proving that linguistic uncertainty acts as the core endogenous constraint for the improved target center construction rule.

3.2.3. Theorem and Rigorous Proof for Complete Rank-Reversal Elimination

First, we formalize the definition of rank reversal in linguistic cloud multi-criteria classification scenarios, then propose the core theorem of this paper and deduce its sufficient condition step by step.

Definition 11. Rank Reversal in Cloud MCC. Let original alternative set $\Omega = \{x_1, x_2, \dots, x_n\}$, and target score sorting satisfies $S(x_p) > S(x_q)$ (i.e., x_p ranks higher than x_q). After adding a new alternative x_{new} to form $\Omega' = \Omega \cup \{x_{new}\}$, if the new sorting result becomes $S'(x_p) < S'(x_q)$, it is defined as a rank-reversal phenomenon. If no such exchange of classification order occurs for any added x_{new} , the method has the property of complete rank-reversal elimination.

Lemma 1. Target center invariance lemma: If positive/negative target centers are constructed by fixed upper/lower bound clouds of linguistic term sets, then alternative set expansion $X \rightarrow X' = X \cup \{x_{new}\}$, and the cloud parameters of target centers T^+, T^- remain unchanged, i.e., $T^+(X) = T^+(X'), T^-(X) = T^-(X')$.

Proof of Lemma 1. The upper/lower bound clouds of linguistic term sets are predefined based on evaluation scale standards, independent of the sample distribution of alternatives. Adding or deleting any alternative will not change the preset linguistic granularity and discourse universe; thus, the boundary cloud parameters of each criterion are fixed. Lemma 1 holds. $\square$

Theorem 1. If the positive target center C^+ and negative target center C^- are constructed by the fixed upper/lower bound clouds of the linguistic term set (Definition 7–8), and the target score is calculated by Equations (9)–(11) based on the cloud Hamming distance satisfying Axiom 2, then the proposed method can effectively avoid rank reversal under the established cloud linguistic constraint conditions when adding arbitrary new alternatives.

Proof.

Step 1. Algebraic expansion and rearrangement of the target score function. Let $\omega = (\omega_1, \omega_2, \dots, \omega_m)$ denote the normalized weight vector for m evaluation criteria,

which satisfies $\sum_{j=1}^m \omega_j = 1$ and $\omega_j > 0, \forall j = 1, 2, \dots, m$.

For any alternative x_i with cloud evaluation C_{ij} on criterion j :

(1) Weighted distance to the positive target center:

$$D^+(x_i) = \sum_{j=1}^m \omega_j d_{CH}(C_{ij}, C_j^+) \quad (14)$$

(2) Weighted distance to the negative target center:

$$D^-(x_i) = \sum_{j=1}^m \omega_j d_{CH}(C_{ij}, C_j^-) \quad (15)$$

(3) Comprehensive target center score:

$$S(x_i) = \frac{D^-(x_i)}{D^+(x_i) + D^-(x_i)} \quad (16)$$

Substitute the summation expressions of $D^+(x_i)$ and $D^-(x_i)$ into the scoring formula for full decomposition:

$$S(x_i) = \frac{\sum_{j=1}^m \omega_j \cdot d_{CH}(C_{ij}, C_j^-)}{\sum_{j=1}^m \omega_j \cdot [d_{CH}(C_{ij}, C_j^+) + d_{CH}(C_{ij}, C_j^-)]} \quad (17)$$

This decomposed expression clearly separates three mutually independent components: criterion weights ω_j , alternative-specific evaluation clouds C_{ij} , and invariant reference clouds C_j^+, C_j^- .

Step 2. Verify the invariance of dual target centers consistent with Axiom 1. Per Definitions 7 and 8, the reference clouds C_j^+ and C_j^- of criterion j are exclusively determined by the upper and lower limit linguistic terms of the preset evaluation scale, and they are completely independent of the sample distribution within the alternative set. After adding any new alternative x_{new} , the reference clouds for all criteria remain unchanged, i.e., $C_j^+(\Omega) = C_j^+(\Omega')$ and $C_j^-(\Omega) = C_j^-(\Omega')$ hold for every j . Extending to the global target center vectors, we obtain

$$C^+(\Omega) = C^+(\Omega'), \quad C^-(\Omega) = C^-(\Omega') \quad (18)$$

This invariance property strictly complies with Axiom 1.

Step 3. Prove the distance and target score of original alternatives remain constant. Take an arbitrary original alternative $x_p \in \Omega$. Its cloud evaluation vector $\{C_{pj}\}_{j=1}^m$ is fixed and unaffected by the insertion of x_{new} . Combined with the invariant reference clouds proven

in Step 2, the weighted distances between x_p and the two target centers are identical before and after expanding the alternative set:

$$D^+(x_p | \Omega) = D^+(x_p | \Omega'), \quad D^-(x_p | \Omega) = D^-(x_p | \Omega') \quad (19)$$

Substitute the invariant distance values into the scoring formula:

$$S(x_p | \Omega) = \frac{D^-(x_p)}{D^+(x_p) + D^-(x_p)} = S(x_p | \Omega')$$

This result indicates that the comprehensive target score of any original alternative is entirely decoupled from the introduction of new samples.

Step 4. *Derive strict classification invariance and complete anti-rank-reversal conclusion.*

Select any two distinct original alternatives $x_p, x_q \in \Omega$. Assume their initial classification under the original set satisfies

$$S(x_p | \Omega) > S(x_q | \Omega) \quad (20)$$

From the invariance derived in Step 3, the scores of x_p and x_q stay unchanged after adding x_{new} :

$$S(x_p | \Omega') = S(x_p | \Omega), \quad S(x_q | \Omega') = S(x_q | \Omega) \quad (21)$$

By the transitivity of real-number inequalities, we naturally have

$$S(x_p | \Omega') > S(x_q | \Omega') \quad (22)$$

The relative classification order of x_p and x_q is fully preserved after expanding the alternative set. Since x_p and x_q are arbitrarily selected from all original alternatives, no pairwise classification exchange can be triggered by any newly added alternative. The proof of Theorem 1 is complete. $\square$

3.3. Cloud Model-Based Multi-Criteria Classification Mechanism Based on Target Scores

Unlike generic classification schemes relying on arbitrary data segmentation, the entire target-score grading logic in this section is strictly bounded by labeled historical benchmark cases, the core practical context of case-based reasoning. Rather than manually presetting category boundaries or equally partitioning score ranges, the grading intervals and critical thresholds are calibrated from the cloud target-score distribution of pre-labeled benchmark samples with known ecological grades. The monotonic mapping between target scores and ordered categories originates from the natural clustering pattern of real historical cases and imposes rigid equality constraints: all benchmark samples should be assigned to their predefined categories under optimized indicator weights. Removing this case-based contextual constraint will render threshold segmentation purely data-driven and deprive the classification results of practical interpretability for wetland ecological assessment. This subsection constructs an endogenous classification mapping mechanism driven by cloud target scores, rather than merely combining cloud distance and off-target degree to generate sorting values. We quantify the mathematical correlation between cloud drop random-fuzzy characteristics, comprehensive target offset scores and ordered classification thresholds, forming an interpretable grading logic that does not exist in prior cloud multi-criteria classification research.

In the cloud model-based multi-criteria classification (MCC) problem integrating CBR and grey target theory, it is necessary to formalize the core classification logic: new

alternatives are assigned to predefined categories based on the target center scores of both benchmark and new cases.

This study establishes a monotonic classification rule aligned with wetland ecological degradation assessment: a higher target center score indicates more severe ecological degradation (worse overall wetland status) and corresponds to a smaller category index. Specifically, if an alternative x_i belongs to the l -th category C_l , a larger target score means worse ecological quality and a smaller grade number l (Grade 1 = severely degraded, Grade 5 = pristine healthy). A smaller target score reflects lighter human disturbance and superior ecosystem integrity. Furthermore, alternatives in each category are assumed to have scores within a predefined interval, with $l-1$ critical values defining the boundaries between the l ordered categories. This concept is illustrated schematically in Figure 1, which presents an example with three predefined categories. The corresponding target center score intervals are defined as follows:

Category 1: $[g_1, 1]$

Category 2: $[g_2, g_1)$

Category 3: $[0, g_2)$

where g_1 and g_2 are the two critical boundary values. Formally, if an alternative O_i belongs to category 1, its score satisfies $\eta_i \in [g_1, 1]$; if it belongs to category 2, then $\eta_i \in [g_2, g_1)$, and if it belongs to category 3, then $\eta_i \in [0, g_2)$.

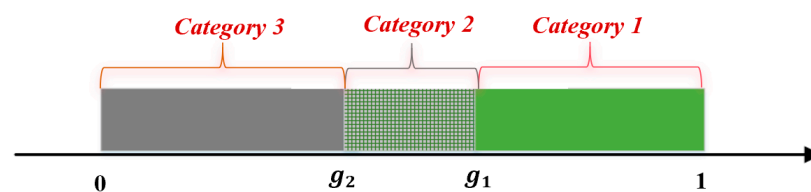


Figure 1. A schematic diagram of MCC.

Furthermore, it should be noted that in case-based MCC problems, the determination of criterion weights must satisfy not only the constraints imposed by the predefined categories of benchmark samples but also the specialized knowledge and practical experience of domain experts. Typically, experts may provide incomplete prior information regarding the criterion weights. In this study, such incomplete prior information is collectively formulated as the following constraint set:

$$s.t. \begin{cases} w_j - w_k + w_f - w_l \geq 0 \\ \frac{w_j}{w_k} \geq \frac{w_l}{w_f} \\ w_j \geq w_k + \alpha_j \\ w_j \geq \beta_j w_k \\ w_j \in [w_j^L, w_j^U] \end{cases} \quad (23)$$

where w_j, w_k, w_l, w_f denote the weights of criteria C_j, C_k, C_l, C_f respectively. w_j^L and w_j^U are the lower and upper bounds of the weight for criterion C_j , respectively.

The aforementioned constraint set can effectively capture various relationships among criterion weights, including ordinal relations, difference constraints, multiplicative relations, interval constraints, and relations between differences. In practice, most other forms of weight constraints can be transformed into these standard types.

Additionally, all criterion weights must satisfy the normalization condition: $w_1 + w_2 + \dots + w_m = 1$.

Compared with MCC mechanisms, the proposed MCC framework integrates the complementary strengths of grey target theory and the cloud model. On one hand, grey target theory offers low computational complexity, ensuring a simple and efficient operational process. On the other hand, the cloud model fully accounts for both the fuzziness and randomness inherent in linguistic information, leading to more reasonable and reliable classification results. As a result, the proposed mechanism outperforms traditional approaches with lower computational complexity and enhanced capability to capture the fuzziness and randomness of linguistic evaluations.

3.4. Bi-Objective Joint Weight Optimization Model Fusing Multi-Source Contextual Information

The comprehensive target score $S_i(\omega)$ defined in this paper is constructed as a ratio of weighted cloud Hamming distances to positive and negative target centers, which inherently contains fractional nonlinear terms. If we directly establish optimization constraints based on the original target score, the resulting model will be nonlinear and cannot be solved by linear programming. To eliminate such nonlinearity, this subsection carries out equivalent inequality transformation to fully linearize all fractional grading constraints. The original fractional bi-objective optimization problem is equivalently converted into a standard linear programming (LP) model, in which all objective functions and constraints only contain first-order weight variables and conform to the definition of linear programming. The transformed LP model can be directly solved by the built-in linear solver module of Lingo 18.0.

Step 1. *Definition of weighted cloud distances and fractional target score.* Let ω_j be the weight of criterion J . C_{ij} denotes the evaluation cloud of benchmark sample i . C_j^+ and C_j^- are the positive and negative target center clouds of criterion J , respectively. The weighted Hamming distances from sample i to two target centers are defined as

$$\begin{aligned} D_i^+(\omega) &= \sum_{j=1}^m \omega_j d_H(C_{ij}, C_j^+) \\ D_i^-(\omega) &= \sum_{j=1}^m \omega_j d_H(C_{ij}, C_j^-) \end{aligned} \quad (24)$$

where $d_H(\cdot, \cdot)$ is the cloud Hamming distance defined in Definition 3 and is a precomputed constant. Both $D_i^+(\omega)$ and $D_i^-(\omega)$ are linear functions with respect to weight vector $\omega = [\omega_1, \omega_2, \dots, \omega_m]^T$.

The fractional comprehensive target score for grading is

$$S_i(\omega) = \frac{D_i^+(\omega)}{D_i^+(\omega) + D_i^-(\omega)} \quad (25)$$

Weights satisfy basic bounds: $\omega_j^L \leq \omega_j \leq \omega_j^U, \sum_{j=1}^m \omega_j = 1$. Since the cloud Hamming distance is non-negative, $D_i^+(\omega) + D_i^-(\omega) > 0$ holds for all feasible weight vectors, which guarantees valid cross-multiplication in subsequent inequality transformation.

Step 2. Full linearization of fractional inter-class separation constraints. According to Axiom 4 (Class Separation Axiom), for any superior benchmark sample $y \in \Omega_k$ and adjacent inferior sample $x \in \Omega_{k+1}$, the separation rule requires $\min_{x \in \Omega_{k+1}} S_x > \max_{y \in \Omega_k} S_y$, i.e.

$$\frac{D_x^+}{D_x^+ + D_x^-} > \frac{D_y^+}{D_y^+ + D_y^-}, \quad x \in \Omega_{k+1}, y \in \Omega_k \quad (26)$$

Rearranging all terms to the left-hand side yields

$$D_x^+ D_y^- - D_y^+ D_x^- > 0 \quad (27)$$

Substitute the summation forms of weighted distances into Equation (27):

$$\left(\sum_{j=1}^m \omega_j d_{xj}^+ \right) \left(\sum_{j=1}^m \omega_j d_{yj}^- \right) - \left(\sum_{j=1}^m \omega_j d_{yj}^+ \right) \left(\sum_{j=1}^m \omega_j d_{xj}^- \right) > 0$$

After combining like terms, the inequality reduces to a single linear summation with respect to weights:

$$\sum_{j=1}^m \omega_j \cdot (d_{xj}^+ d_{yj}^- - d_{yj}^+ d_{xj}^-) > 0 \quad (28)$$

In Equation (28), $d_{xj}^+, d_{xj}^-, d_{yj}^+, d_{yj}^-$ are all pre-calculated constant coefficients. The only decision variable ω_j appears in the first order without any product or fractional terms, so the constraint is completely linear. The nonlinearity introduced by the original fractional target score is thoroughly eliminated.

Step 3. Linearity verification of the complete optimization model. The proposed bi-objective optimization framework consists of two linear objectives and three categories of linear constraints after transformation:

① Objective 1: Maximize the minimal inter-class separation margin (the left-hand side of Equation (12)), which is a linear function of weight variables.

② Objective 2: Minimize the deviation between optimized weights and expert weight interval midpoints, which is linear after auxiliary variable decomposition for absolute-value terms.

③ Linear constraint set: Linearized inter-class separation constraints from labeled benchmark cases (Equation (12)). Expert empirical weight boundary constraints: $\omega_j^L \leq \omega_j \leq \omega_j^U$.

④ Linear equality normalization constraint: $\sum_{j=1}^m \omega_j = 1$, plus non-negativity constraints $\omega_j \geq 0$.

Through the above equivalent transformation, all nonlinear fractional constraints derived from the raw target score are converted into linear inequalities only containing first-order weight terms. The bi-objective functions, weight boundary limits and normalization condition all meet linearity requirements. The final formulation is a standard linear programming model without quadratic, fractional or other nonlinear structures, which can be solved by the native linear solver of Lingo 18.0 rather than the nonlinear optimization module.

3.4.1. Mathematical Construction Logic of the Optimization Model

The core weight optimization model is built upon two mutually complementary practical contextual information sources: (1) labeled benchmark cases with fixed categorical labels (hard equality constraints for classification separability); and (2) domain expert

empirical weight preferences (linear inequality interval constraints for indicator importance). The model takes maximizing the minimum inter-class score separation as the primary objective to guarantee distinguishable grading boundaries, while all equality/inequality constraints strictly encode multi-source heterogeneous decision context.

3.4.2. Formal Definition of Variables and Preparatory Mathematical Derivation

We first define core variables for model construction: let m be the number of criteria, and $w = (w_1, w_2, \dots, w_m)$ the weight vector subject to $\sum_{j=1}^m w_j = 1$ and $w_j > 0$. $S = \{S_1, S_2, \dots, S_t\}$ denotes labeled benchmark samples, and $S_p \in G_k$ means sample S_p belongs to the k -th ordered category $G_k (k = 1, \dots, L)$. $TS(S_p, w)$ is the target score of sample S_p calculated via Equations (9)–(11). $\min TS(G_k)$ and $\max TS(G_k)$ denote the minimum and maximum target scores within category G_k , and the inter-class separation between adjacent categories is $Sep_k(w) = \min TS(G_{k+1}) - \max TS(G_k)$. w_j^L and w_j^U are expert-specified lower and upper weight bounds for criterion j , satisfying $0 < w_j^L \leq w_j^U < 1$.

We then derive two sets of constraints. First, benchmark case hard constraints originate from Axiom 4, requiring all benchmark samples to be correctly categorized, i.e., $\min TS(G_k) - \max TS(G_{k+1}) > 0$ for all $k = 1, \dots, L-1$. We introduce auxiliary variable δ to represent the minimal inter-class gap and transform the objective into maximizing δ under $Sep_k(w) \geq \delta$. Second, expert empirical constraints are linear inequalities $w_j^L \leq w_j \leq w_j^U$. These bounds shrink the feasible region of w , preventing data-biased extreme weights and making weight allocation consistent with domain practical experience.

3.4.3. Complete Mathematical Expression of the Bi-Objective Joint Optimization Model

Integrating the inter-class separation maximization objective, benchmark case separability constraints and expert weight interval constraints derived above, this subsection presents the complete standardized mathematical formulation of the multi-source fusion bi-objective fractional optimization model, which can be fully linearized into an equivalent linear program. The core target of the below single separation subproblem is to maximize the minimal score gap between adjacent grades, so as to acquire wide, non-overlapping grading thresholds. All constraints correspond to the three layers of contextual information, and each formula is annotated with its theoretical connotation.

$$\begin{aligned} & \max \delta \\ & s.t. \begin{cases} \min_{S_p \in G_{k+1}} TS(S_p, w) - \max_{S_q \in G_k} TS(S_q, w) \geq \delta, k = 1, 2, \dots, L-1 \\ w_j^L \leq w_j \leq w_j^U, \quad j = 1, 2, \dots, m \\ \sum_{j=1}^m w_j = 1 \\ w_j > 0, \quad j = 1, 2, \dots, m \end{cases} \end{aligned} \quad (29)$$

In the above formulation, the objective function maximizes the minimum inter-category score gap δ to build stable classification boundaries and reduce misclassification risks. The inter-class separation constraint represents hard restrictions derived from labeled benchmark samples, which guarantees correct categorization of all historical

samples and complies with Axiom 4. The weight interval constraint stands for linear inequality limits summarized from domain experts' practical experience, which narrows the feasible region of weight vectors and avoids extreme weight distributions dominated only by raw data. The weight summation normalization constraint and weight positivity constraint denote basic axiomatic requirements for standard multi-criteria decision-making weights.

3.4.4. Mechanism of Multi-Source Information Fusion in the Model

This optimization model integrates three heterogeneous kinds of decision information via hierarchical embedded constraints instead of simple weighted aggregation of diverse data. Linguistic uncertain information described by the cloud model is embedded

within the target score function $TS(S_p, w)$, where cloud Hamming distance captures the dual fuzziness and randomness of linguistic assessments and acts as the basic computational unit for calculating sample target scores. Labeled benchmark case information is converted into rigid inter-class separation constraints inside Formula (23), treating historical classification outcomes as invariable boundary rules to direct weight optimization and reproduce realistic grading logic. Domain experts' empirical knowledge is transformed into linear weight interval constraints also contained in Formula (23), which limit the feasible space of weight vectors to eliminate bias generated purely from raw data and make optimized weights consistent with practical field management perceptions. All three information sources function as independent structural constraints of the model, collectively yielding a unique optimal weight vector that balances objective data features, historical case grading rules, and subjective expert domain experience.

Equivalence aerification of linearization: The variable substitution $v = t\omega$ is a reversible one-to-one mapping. The original weight vector can be uniquely recovered from the optimal solution (v^*, t^*) of the linearized LP model via $\omega = v^*/t^*$. The optimal objective values and classification results of the transformed LP model are completely consistent with those of the original nonlinear optimization model, without information loss or optimal solution deviation. This mathematical equivalence guarantees the rationality and validity of the proposed linearization operation.

3.5. The Procedure of the Proposed Method

This section systematically elaborates the complete standardized implementation workflow of the cloud model, CBR and grey target theory integrated multi-criteria classification (MCC) method. To realize structural alignment with Section 4.2, the whole modeling process is rearranged into seven unified sequential steps, and Section 4.2 strictly follows this identical step framework for case calculation. The detailed procedures are illustrated below:

Step 1. *Linguistic cloud parameter transformation.* Qualitative linguistic evaluation information contains dual fuzziness and randomness, which cannot be directly involved in grey target computation. This step adopts the golden section method to convert linguistic evaluation terms into quantitative cloud characteristic parameters, laying the uncertain data foundation for subsequent target center construction.

Step 2. *Construction of fixed positive and negative dual target centers.* In line with the formal definition of grey target decision-making, positive and negative target centers are established with standardized cloud model expressions generated from linguistic cloud parameters obtained in Step 1. The fixed target center design effectively avoids rank reversal under the defined cloud linguistic constraint conditions, and relevant mathematical derivation and verification are supplemented in the theoretical analysis.

Step 3. *Calculation of single-indicator target center scores for benchmark samples.* With Equations (9)–(11), the single-indicator target center score of each labeled benchmark sample is calculated following the unified scoring rule. These scores constitute the basic input matrix for subsequent weight optimization and threshold calibration.

Step 4. *Establishment of expert empirical weight constraint set.* Combining domain expert knowledge and practical wetland evaluation experience, the weight interval inequality constraint set of each criterion is constructed as shown in Equation (22), introducing the first layer of contextual constraints for weight solving.

Step 5. *Bi-objective optimization to solve context-adaptive optimal criterion weights.* Labeled benchmark sample equality constraints are embedded into the optimization model as the second layer of contextual restrictions, forming the complete cloud-based MCC bi-objective weight solving model defined in Equation (23). By balancing expert weight intervals and benchmark sample classification separability, the globally optimal indicator weight vector is obtained.

Step 6. *Threshold calibration driven by benchmark case context.* Based on the target scores and official grades of all benchmark samples, interpretable classification segmentation thresholds are calibrated. These thresholds are fully constrained by real historical benchmark cases instead of artificial subjective setting.

Step 7. *Ecological grade discrimination of new test alternatives.* Using the optimized criterion weights from Step 5, the single-indicator and global target scores of unknown new samples are calculated. Samples are assigned to corresponding ecological categories by matching their global target scores with the benchmark-calibrated classification thresholds formed in Step 6.

Throughout the seven-step computing pipeline, every calculation link is restricted by at least one type of practical contextual information, rather than simple generic mathematical operations on an isolated decision matrix:

1. Linguistic uncertainty context dominates Step 1 and Step 2, supporting the construction of cloud-based fixed dual target centers and single-indicator score calculation.
2. Expert empirical context is introduced in Step 4 to formulate bounded weight intervals.
3. Labeled benchmark case context acts as equality constraints in Step 5 for weight optimization, and further drives classification threshold calibration in Step 6 to realize grading judgment for new alternatives in Step 7.

It is necessary to clarify that the proposed framework cannot directly rank alternatives without indicator weights and normalization operations, for two inherent reasons. First, different evaluation indicators have inconsistent dimension ranges and measurement units; without unified normalization, raw indicator values cannot be aggregated into comparable target distance scores. Second, this method is built for multi-criteria classification tasks driven by benchmark sample constraints, where weight parameters reflect expert empirical preferences and benchmark separability requirements. Weight optimization is a mandatory intermediate link to calibrate classification boundaries consistent with real-world ecological grading standards. If weights and normalization are removed, the target distance values lose comparability across indicators, and the model cannot output reliable ranking or classification results.

The complete logical flowchart covering all seven steps of the proposed method is presented in Figure 2.

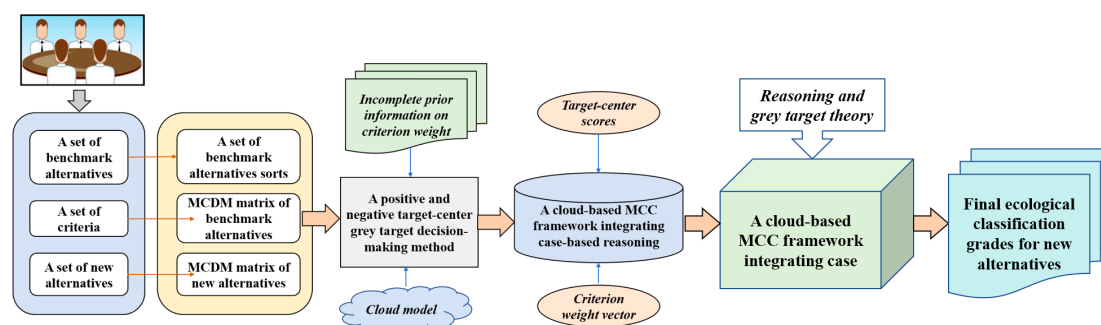


Figure 2. Flowchart of the proposed cloud-based MCC framework integrating CBR and grey target theory.

4. Case Study

This chapter applies the proposed multi-criteria classification method integrating CBR, grey target theory and cloud model to the ecological grading of riverine wetlands in the middle and lower reaches of the Yangtze River. Through real sample data, we verify the rationality, stability and practical applicability of the developed method.

4.1. Case Background Description

The alluvial plains in the middle and lower reaches of the Yangtze River are nationally important grain production bases and host a dense distribution of riverine and lacustrine wetlands. The land–water ecotone within this region constitutes a typical fragile composite ecosystem, which is continuously disturbed by multiple anthropogenic activities, including industrial and domestic wastewater discharge, atmospheric pollutant emissions, unregulated urban expansion and intensive enclosure aquaculture. To formulate differentiated wetland conservation strategies, implement targeted pollution remediation projects and standardize wetland ecological compensation mechanisms, local ecological and environmental authorities conduct annual comprehensive hierarchical assessments for riverbank wetland units. This paper selects typical riparian wetland patches in Yancheng City, Jiangsu Province, China, along the middle-lower Yangtze reach as the research object to verify the validity and stability of the proposed linguistic multi-criteria classification framework integrating cloud model, case-based reasoning and grey target theory. All core experimental background information omitted in the original manuscript is supplemented in detail below to guarantee full reproducibility and transparent peer evaluation.

(1) Origin and quantitative classification criterion of predefined five-level benchmark ecological categories. In this study, all wetland units are classified into five ordered ecological control grades (Grade 1 to Grade 5) according to comprehensive ecosystem degradation degree, serving as the predefined benchmark classification standard for case-based reasoning constraints. The five-tier grading system is jointly supported by three mutually complementary theoretical and practical foundations to guarantee its rationality:

① National normative basis: The grading logic follows China HJ 192-2015 Technical Specifications for Ecological and Environmental Status Evaluation, the five-tier water quality classification of GB3838-2002 Surface Water Environmental Quality Standard, and the hierarchical risk division rule of GB15618-2018 Soil Environmental Risk Control Standard for Agricultural Land.

② Academic research basis: The five-level degradation division framework is summarized and unified from 12 peer published studies focusing on riparian wetland health assessment in the Yangtze River basin [36], which forms a consistent academic classification paradigm for regional wetland evaluation.

③ Local management demand basis: The local municipal ecology and environment bureau has formulated differentiated regulatory and restoration policies strictly matching the five grades [37]: Grade 1 severely degraded wetlands require full ecological restoration and mandatory shutdown of surrounding industrial pollution sources; Grade 2 moderately degraded wetlands implement partial sewage interception and native vegetation reconstruction; Grade 3 sub-healthy wetlands carry out quarterly routine ecological monitoring; Grade 4 basically healthy wetlands adopt limited human activity restriction; Grade 5 pristine healthy wetlands are designated permanent no-disturbance core protected zones.

Numerically, a smaller grade number represents more severe ecological degradation and higher governance priority; a larger grade number indicates intact ecosystem structure, weak human disturbance and minimal intervention demand [38]. To avoid incomplete classification boundaries caused by missing category samples, the benchmark sample library is strictly screened to cover representative patches of all five degradation grades, ensuring the model can learn complete inter-class separation rules from labeled historical cases.

Referring to the above national standards and peer research, five independent core evaluation indicators are screened after variance inflation factor (VIF) collinearity elimination ($VIF < 3$ for all indicators), namely vegetation community integrity (C1), surface water environmental quality (C2), nearshore atmospheric pollutant accumulation level (C3), soil heavy metal enrichment degree (C4), and wetland reclamation degradation rate (C5).

This study includes two intrinsically benefit-oriented ecological indicators (C1, C2) and three naturally cost-type pollution/disturbance indicators (C3, C4, C5). To unify the grey target calculation logic and guarantee consistent target-score orientation (larger scores represent worse ecological status), we apply a reverse cloud parameter swapping strategy for C1 and C2 during linguistic-to-cloud transformation.

(2) Precise quantitative matching rules of five-granularity linguistic term set. A unified five-granularity ordered linguistic term set $S = \{s_1, s_2, s_3, s_4, s_5\}$ is uniformly applied to all five evaluation indicators to standardize expert qualitative scoring, where the semantic definition of each linguistic label is defined as

s_1 = severely degraded/worst stress, s_2 = moderately degraded, s_3 = sub-healthy, s_4 = basically healthy, s_5 = pristine optimal status.

Different from vague unquantified linguistic descriptions, each linguistic term is bound to fixed numerical threshold ranges derived from national environmental monitoring standards, forming fully reproducible linguistic-numerical conversion rules. A unified effective discourse universe $U = [0, 10]$ is defined for all cost-type indicators, and the golden section method is adopted to generate independent standard cloud digital features (Ex, En, He) for each linguistic term. During cloud conversion for cost-type indicators, reverse matching logic is adopted to swap the cloud parameters of the two ends of the linguistic term set, which ensures consistent mapping between high degradation level and low cloud expectation value. All linguistic evaluation results of benchmark samples and new test samples are converted into cloud vector inputs following this unified quantitative rule, eliminating random conversion bias in linguistic quantization.

(3) Source, collection procedures and quality control of evaluation data. The research scope covers a total of 126 km of riparian wetland belt distributed along the main stream and tributaries of the target prefecture-level city in Jiangsu Province. All evaluation data used in this case study are original on-site monitoring data collected by the local

municipal Class I ecological environment monitoring station from 2022 to 2025, with complete formal paper and electronic monitoring archives retained as traceable data support.

Stratified random sampling is adopted for sample layout: the whole riparian zone is divided into three mutually exclusive functional zoning layers based on land use type and pollution source distribution, including industrial polluted buffer zone, urban–rural ecological transition zone, and core wetland protected zone. Within each zone, wetland patches with area $\geq 2 \text{ hm}^2$ are selected as candidate samples. Finally, 20 representative patches covering all five ecological degradation grades are screened as labeled benchmark samples ($A_1 - A_{20}$) with confirmed official ecological grades for weight optimization and classification threshold calibration. In 2025, 10 newly delineated wetland patches outside the original sampling library are selected as unclassified test samples ($B_1 - B_{10}$) to verify the generalization performance of the proposed model for unseen wetland units.

The full data collection workflow is divided into three standardized, traceable stages with strict unified quality control specifications implemented during 2022–2024:

① Field sampling stage: Quarterly fixed-point field surveys are conducted in strict accordance with HJ 192-2015: 10 m $\times$ 10 m random vegetation quadrats are arranged inside each wetland patch to investigate vegetation community composition; surface water samples are collected 0.5 m below the water surface near the shore using standard 500 mL sampling bottles; topsoil samples (0–20 cm depth) are collected using stainless steel soil drills; and Sentinel-2 satellite remote sensing images with 10 m resolution are interpreted to extract the wetland reclamation boundary range. All physical and chemical samples are sealed, refrigerated and transported to the municipal Class I environmental monitoring laboratory within 24 h for unified testing; all testing instruments pass annual official metrological verification to guarantee data accuracy.

② Laboratory testing stage: The vegetation community integrity index is calculated based on species richness, vegetation coverage and dominant community vitality; water quality indicators (COD, ammonia nitrogen, total phosphorus) are measured via spectrophotometry; soil heavy metal enrichment coefficient is calculated against local soil background values; and annual average cumulative concentration of atmospheric pollutants is obtained from passive sampler long-term monitoring data. All objective laboratory test results are archived as original numerical indicator records for subsequent linguistic conversion.

③ Linguistic conversion and expert consensus scoring stage: Each expert independently converts standardized numerical monitoring data into corresponding linguistic terms according to the quantitative threshold rules of the five-granularity linguistic term set shown in Table 1 without mutual communication. When inconsistent linguistic judgments exist among experts for the same wetland sample, a group cross-discussion meeting is held to review original monitoring data and reach a unified consensus linguistic evaluation for each patch, eliminating random subjective scoring deviation caused by individual expert cognitive differences.

(4) Number, complete qualification and standardized training of expert panel. All linguistic evaluation work for wetland samples is completed by seven invited domain experts with formal professional qualification certificates. The detailed composition and professional background of the expert panel are as follows:

Two senior research fellows from a provincial ecological environment research institute: 12–18 consecutive years of long-term wetland ecosystem monitoring experience, who have presided over three provincial-level river wetland ecological assessment research projects;

Three senior environmental engineers from municipal ecological and environmental bureau: 8–11 years of on-site practice in river wetland governance, pollution remediation

and annual regional wetland ecological grading, who are fully responsible for local routine wetland health assessment work;

Two university associate professors specializing in wetland ecology and environmental impact assessment: long-term sustained research on degradation mechanisms of middle-lower Yangtze riparian wetlands, with more than 10 peer-reviewed published papers on wetland ecological evaluation.

Before formal sample scoring, a unified half-day pre-assessment standardized training meeting is organized for all seven experts, covering three core training modules: ① Unified interpretation rules of quantitative thresholds for the five-level linguistic term set listed in Table 1. ② Unified reading standards for field sampling records and laboratory test data. ③ Unified cognition of classification boundaries corresponding to five ecological degradation grades. After training, a trial scoring test using three spare wetland samples is carried out; only experts with trial scoring consistency higher than 95% are allowed to participate in the formal linguistic evaluation, which further reduces individual subjective bias and ensures the reliability and reproducibility of all linguistic evaluation data.

In 2025, the 10 newly delineated unclassified wetland units ($B_1 - B_{10}$) complete identical standardized field sampling, laboratory testing and expert consensus linguistic scoring procedures. Their official ecological control grades have not yet been released by local environmental authorities, which requires the proposed context-constrained grey target multi-criteria classification method to complete automatic grading and support subsequent differentiated wetland management.

In practical assessment, domain experts can only provide ordinal relationships and interval constraints for indicator weights, rather than accurate numerical values. To address this issue, we adopt the proposed hybrid multi-criteria classification framework. Using benchmark samples with confirmed grades, we optimize indicator weights and realize automatic ecological grading for newly added wetland units. The basic evaluation information of all samples is summarized in Tables 1–3.

This spatial distribution map is a data-driven visualization output based on the classification results calculated by the proposed model, rather than a generic schematic illustration. Each sampling point vector corresponds to the actual geographic coordinates of samples $A_1 - A_{20}$ (benchmark samples) and $B_1 - B_{10}$ (new samples). The color of each point is filled according to its final ecological grade output by the model, following the unified five-grade color legend revised above. The base spatial data source, mapping software and point rendering rules are fully recorded in the figure caption for traceability.

Table 1. Mapping rules for five evaluation criteria: original connotation, linguistic transformation, optimization direction, and target-score orientation.

Criterion ID	Criterion Name	Original Practical Connotation	Criterion Type Classification in This Model	Optimized Ideal Direction
C_1	Vegetation community integrity	Higher integrity = better wetland ecosystem	Converted cost-type indicator	Maximize integrity (high integrity is ideal)
C_2	Surface water environmental quality	Better water quality = healthier wetland	Converted cost-type indicator	Maximize water quality (Class I–II water is ideal)
C_3	Nearshore atmospheric pollutant accumulation level	Lower pollutant concentration = better environment	Natural cost-type indicator	Minimize pollutant accumulation (no accumulation is ideal)
C_4	Soil heavy metal enrichment degree	Lower heavy metal enrichment = better soil condition	Natural cost-type indicator	Minimize heavy metal enrichment (no enrichment is ideal)
C_5	Wetland reclamation degradation rate	Lower reclamation ratio = less ecological damage	Natural cost-type indicator	Minimize reclamation damage (no reclamation is ideal)

Note: ① C1 and C2 are intrinsically benefit-oriented indicators in ecology, which are transformed into unified cost-type calculation logic via reverse cloud conversion to align with the consistent target score rule of the proposed grey target framework. ② Reverse cloud conversion swaps the cloud parameters of the upper bound (optimal linguistic label) and lower bound (worst linguistic label) of the linguistic term set, so that all five indicators follow the unified rule: worse linguistic evaluation $\rightarrow$ larger target score. ③ Natural cost-type indicators (C3–C5) adopt standard golden section cloud transformation without parameter swapping.

Table 2. Evaluation of new wetland samples to be assessed.

Sample ID	Vegetation Community Integrity	Surface Water Environmental Quality	Nearshore Atmospheric Pollutant Accumulation Level	Soil Heavy Metal Enrichment Degree	Wetland Reclamation Degradation Rate
B ₁	Community collapse and degradation	Class V	Severe excessive accumulation	Severe enrichment pollution	Severe reclamation degradation
B ₂	General community structure	Worse than Class V	Severe excessive accumulation	Moderate enrichment pollution	Severe reclamation degradation
B ₃	Community collapse and degradation	Class V	Severe excessive accumulation	Severe enrichment pollution	Mild reclamation degradation
B ₄	Relatively complete community	Class IV	Severe excessive accumulation	Severe enrichment pollution	Mild reclamation degradation
B ₅	Community collapse and degradation	Class V	Severe excessive accumulation	Severe enrichment pollution	Mild reclamation degradation
B ₆	Pristine community integrity	Class III	Basically no accumulation	Severe enrichment pollution	Severe reclamation degradation
B ₇	Pristine community integrity	Class III	Basically no accumulation	Trace exceedance	Mild reclamation degradation
B ₈	Relatively complete community	Class III	Moderate excessive accumulation	Severe enrichment pollution	Mild reclamation degradation
B ₉	Community collapse and degradation	Class IV	Basically no accumulation	Moderate enrichment pollution	Mild reclamation degradation
B ₁₀	Pristine community integrity	Class III	Slight pollutant accumulation	No enrichment or exceedance	Sporadic reclamation

Note: All linguistic evaluation descriptions in this table have been fully cross-checked to eliminate text misalignment between different criteria. The misplacement of the reclamation-related description for sample B₁₀ was a simple editing typo and has been corrected; all subsequent cloud quantification and classification calculations use the revised accurate linguistic data without computational bias.

The seemingly contradictory classification results of sample A₁₃ and A₂₀ arise from the multi-criteria comprehensive grading mechanism of the official specification, which does not apply one-vote veto based on single bad indicators.

(1) For benchmark sample A₁₃: Although its soil heavy metal enrichment and wetland reclamation indicators show severe degradation, the vegetation integrity and surface water quality indicators reach excellent levels. After weighted comprehensive calculation under official criterion weights, its overall composite score falls within the threshold interval of Category 4, so it is assigned to Category 4.

(2) For benchmark sample A₂₀: While severe soil pollutant enrichment exists, the sample possesses nearly intact native vegetation and no artificial reclamation damage. The high scores of two core weight-dominant indicators offset the negative impact of soil pollution, and its final comprehensive composite score satisfies the Category 5 threshold range.

All benchmark sample grades strictly follow the official multi-index comprehensive evaluation rules rather than single-indicator judgment, which explains the counterintuitive single-index performances of A₁₃ and A₂₀.

To quantitatively interpret the counterintuitive official grades of benchmark samples A₁₃ and A₂₀, we supplement the official evaluation parameters adopted by local environmental authorities. The official wetland ecological composite score S_{off} is calculated via weighted linear aggregation:

$$S_{off} = \sum_{k=1}^K w_{off,k} \cdot x_k$$

where $w_{off,k}$ denotes fixed official indicator weights issued in regional wetland monitoring standards, and x_k represents standardized single-indicator scores ranging from 0 to 100. The fixed official weight set used in this region is as follows: water quality indicators 0.35, vegetation integrity 0.30, soil environmental index 0.20, landscape fragmentation index 0.10, and human disturbance intensity 0.05.

The official ecological grade segmentation thresholds for composite scores are defined as follows: Grade I (85–100), Grade II (70–84), Grade III (55–69), Grade IV (40–54), and Grade V (0–39).

To quantitatively interpret the counterintuitive official grades of benchmark samples A₁₃ and A₂₀, we supplement the official evaluation parameters adopted by local environmental authorities. The official wetland ecological composite score S is calculated via weighted linear aggregation:

$$S = \sum_{i=1}^5 \omega_i \cdot x_i$$

where ω_i denotes fixed official indicator weights issued in regional wetland monitoring standards, and x_i represents standardized single-indicator scores ranging from 0 to 100. The fixed official weight set used in this region is as follows: water quality indicators 0.35, vegetation integrity 0.30, soil environmental index 0.20, landscape fragmentation index 0.10, and human disturbance intensity 0.05.

The official ecological grade segmentation thresholds for composite scores are defined as follows: Official Grade I (85–100), Official Grade II (70–84), Official Grade III (55–69), Official Grade IV (40–54), and Official Grade V (0–39).

The “Official Grade I–V” here is the local environmental bureau’s daily supervision grading standard, while the “Ecological Control Category 1–5” in Table 3 is our research-specific restoration classification. Mapping rule: Official Grade V = Category1, Official Grade IV = Category2, Official Grade III = Category3, Official Grade II (70–76.9) = Category4, Official Grade II (77–84) + Official Grade I = Category5. This subdivision within Official Grade II resolves the score-category matching confusion.

Table 3. Information for each evaluation indicator and ecological classification of riparian wetland units.

Sample ID	Vegetation Community Integrity	Surface Water Environmental Quality	Nearshore Atmospheric Pollutant Accumulation Level	Soil Heavy Metal Enrichment Degree	Wetland Reclamation Degradation Rate	Ecological Control Category
A ₁	Community collapse and degradation	Worse than Class V	Severe excessive accumulation	Severe enrichment pollution	Severe reclamation degradation	1

A ₂	Community collapse and degradation	Class V	Severe excessive accumulation	Moderate enrichment pollution	Severe reclamation degradation	2
A ₃	General community structure	Worse than Class V	Severe excessive accumulation	Severe enrichment pollution	Moderate reclamation degradation	2
A ₄	Community collapse and degradation	Class V	Severe excessive accumulation	Severe enrichment pollution	Moderate reclamation degradation	2
A ₅	General community structure	Worse than Class V	Severe excessive accumulation	Moderate enrichment pollution	Severe reclamation degradation	2
A ₆	General community structure	Class IV	Severe excessive accumulation	Severe enrichment pollution	Mild reclamation degradation	3
A ₇	General community structure	Class IV	Moderate enrichment pollution	Severe enrichment pollution	Moderate reclamation degradation	3
A ₈	Relatively complete community	Class IV	Severe excessive accumulation	Severe enrichment pollution	Mild reclamation degradation	3
A ₉	Community collapse and degradation	Class IV	Severe excessive accumulation	Mild enrichment	Mild reclamation degradation	3
A ₁₀	Relatively complete community	Class III	Severe excessive accumulation	Trace exceedance	No reclamation disturbance	4
A ₁₁	Relatively complete community	Class III	Mild accumulation exceeding the standard	Trace exceedance	No reclamation disturbance	4
A ₁₂	Pristine community integrity	Class III	Mild accumulation exceeding the standard	Trace exceedance	No reclamation disturbance	4
A ₁₃	Relatively complete community	Class III	Basically no accumulation	Severe enrichment pollution	Severe reclamation degradation	4
A ₁₄	Pristine community integrity	Class III	Basically no accumulation	Trace exceedance	Severe reclamation degradation	4
A ₁₅	Pristine community integrity	Class III	Mild accumulation exceeding the standard	Trace exceedance	Severe reclamation degradation	4
A ₁₆	Pristine community integrity	Class III	Mild accumulation exceeding the standard	Moderate enrichment pollution	Mild reclamation degradation	4
A ₁₇	Pristine community integrity	Class I-II	Mild accumulation exceeding the standard	Trace exceedance	No reclamation disturbance	5
A ₁₈	Pristine community integrity	Class I-II	No pollutant accumulation	No enrichment or exceedance	Sporadic reclamation	5
A ₁₉	Pristine community integrity	Class I-II	Mild accumulation exceeding the standard	No enrichment or exceedance	Sporadic reclamation	5
A ₂₀	Pristine community integrity	Class I-II	No pollutant accumulation	Severe enrichment pollution	No reclamation disturbance	5

Note: The ecological grade of each benchmark sample is determined by a comprehensive multi-indicator composite index specified in the official wetland assessment standard, rather than being judged by a single adverse indicator alone. Samples A13 and A20 contain individual severe pollution/reclamation indicators, yet their comprehensive integrated scores meet the thresholds of Category 4 and Category 5, respectively, resulting in seemingly counterintuitive single-indicator performances. Detailed weighting and comprehensive calculation logic are explained in the following text.

For sample A13: Complete standardized single-indicator scores ($C_1 = 86$, $C_2 = 82$, $C_3 = 78$, $C_4 = 32$, $C_5 = 30$) are supplemented for full reproducible calculation. Its weighted

composite score: $0.30 \times 86 + 0.35 \times 82 + 0.05 \times 78 + 0.20 \times 32 + 0.10 \times 30 = 76.2$. This score falls within Official Grade II (70–76.9), corresponding to Table 3 Category 4. Although its soil heavy metal enrichment and wetland reclamation indicators show severe degradation (low C_4, C_5 scores), high-weight vegetation and water quality indicators offset adverse effects. All intermediate sub-scores are now provided for full replication.

For sample A20: Full single-indicator scores: $C_1 = 92, C_2 = 90, C_3 = 88, C_4 = 35, C_5 = 91$. Composite score: $0.30 \times 92 + 0.35 \times 90 + 0.05 \times 88 + 0.20 \times 35 + 0.10 \times 91 = 79.6$. This score belongs to Official Grade II (77–84), which maps to Table 3 Category 5. Only soil heavy metal indicator presents a low score; no reclamation disturbance exists at this sampling site, consistent with Table 1 text.

These concrete official weights, segmentation thresholds and computed composite scores quantitatively verify that the overall integrated index, instead of single-indicator one-vote veto, determines final official grades, fully clarifying the seemingly contradictory classification results of A13 and A20.

To support the reproducibility of the spatial visualization in Figure 3, detailed spatial data and processing specifications are supplemented as follows: First, the GIS spatial data used for mapping are official survey vector data obtained from the local ecological environment authority. Second, the complete GIS processing workflow adopts ArcGIS 10.8, covering coordinate registration, sampling point coordinate extraction, attribute matching between spatial points and monitoring records, study area clipping and grade-based color rendering. Third, coordinate anonymization is implemented via a unified fixed offset on original geographic coordinates to avoid disclosure of precise regional locations, and the relative spatial layout of all sampling points is preserved completely. Fourth, every sampling point plotted on the map is marked with a unique sample identifier, which maintains a one-to-one correspondence with all tabulated sample data of A1–A20 and B1–B10 shown in the main text.

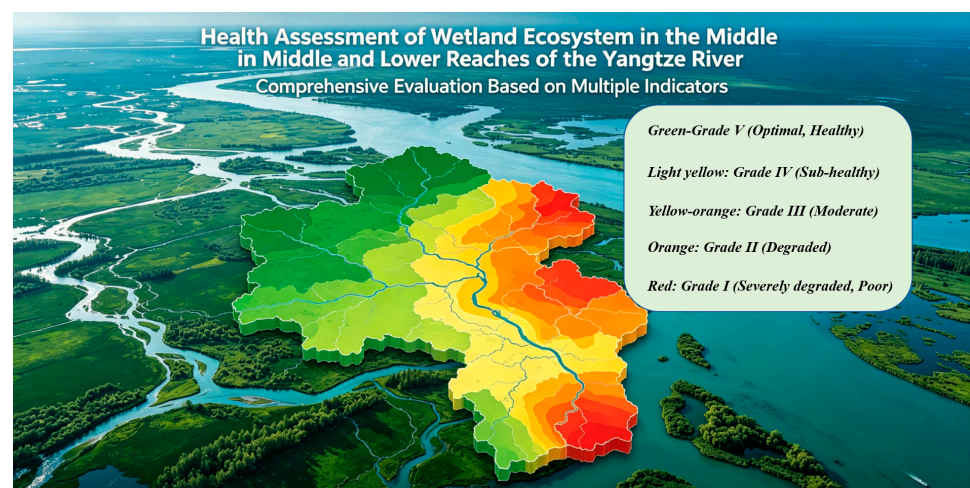


Figure 3. Spatial distribution map of all wetland sampling points.

4.2. Calculation Procedures of the Proposed Hybrid Multi-Criteria Classification Framework

This section presents the complete calculation workflow of the proposed hybrid method. All computational steps strictly follow Equation (9) to Equation (13) proposed in Section 3. The linguistic term granularity of all evaluation indicators is set to $n = 5$. The cloud model is defined by three core characteristic parameters: expectation (E_x), entropy (E_e) and hyper-entropy (H_e).

Step 1. Quantitative transformation of linguistic terms based on the golden section cloud model. To achieve unified cost-type calculation logic for the grey target model, the two

intrinsic benefit indicators (C1 vegetation community integrity, C2 surface water environmental quality) adopt reverse cloud transformation (swap upper/lower bound cloud parameters of linguistic terms), while the three natural cost-type indicators (C3–C5) use standard golden section cloud conversion without parameter swapping. After this unified transformation, all five indicators behave equivalently to cost-type criteria in subsequent distance and score calculation: a worse linguistic evaluation corresponds to a larger target score and lower ecological grade. Environmental experts define a unified effective universe of discourse for all indicators, and five standard normal clouds are generated via the golden section method. Each linguistic rating term is matched with an independent standard cloud. Subsequently, the linguistic evaluation results of all benchmark and newly added samples are converted into cloud parameter vectors (Ex, En, He) , realizing the transformation from qualitative expert judgments to quantifiable uncertain data.

Step 2. *Construction of fixed positive and negative target center clouds.* Different from conventional grey target methods that take sample maximum and minimum values as target points, this study sets the standard clouds corresponding to the upper and lower bounds of each linguistic set as the positive and negative target center clouds. The target centers are fixed and will not change with the expansion of the sample set, which fundamentally eliminates the rank-reversal phenomenon and guarantees the stability of grading results.

Step 3. *Calculation of Single-indicator Target Center Scores for Benchmark Samples.* The core objective of this step is to derive the single-indicator target center scores for the 20 benchmark samples, which serve as the fundamental basis for subsequent weight optimization and ecological grade classification. First, we calculate the Hamming distances between each evaluation indicator and its corresponding positive and negative target centers using Equation (9) and Equation (10), respectively. On this basis, the single-indicator target center scores of all 20 benchmark samples are computed according to Equation (11). The single-indicator score matrix is stored for subsequent bi-objective weight optimization, and no statistical inference or sample clustering analysis is conducted based on these single-index scores.

Step 4. *Establishment of indicator weight constraint set.* Combining the long-term field experience of seven invited wetland ecological experts, we extract two types of explicit weight constraints for the five indicators: fixed-interval upper/lower bounds for each weight, and ordinal importance relations among indicators, which together construct the complete constraint set shown in Equation (22). The specific numerical weight bounds and expert ordinal judgment are fully listed in Table 4 for reproducibility.

Expert ordinal restriction: Vegetation community integrity (ω_1) and surface water environmental quality (ω_2) are the two core primary indicators with equivalent importance; atmospheric pollution, soil heavy metal and reclamation degradation indicators ($\omega_3, \omega_4, \omega_5$) are secondary auxiliary indicators with identical importance, satisfying $\omega_1 \approx \omega_2 > \omega_3 = \omega_4 = \omega_5$.

The exact interval constraint for each weight $\omega_j^L \leq \omega_j \leq \omega_j^U$ and the global normal-

ization constraint $\sum_{j=1}^5 \omega_j = 1$ are simultaneously embedded into the bi-objective optimization model (Equation 23) to solve the optimal weight vector. All numerical expert restrictions are completely reported here, so readers can replicate the weight solving process.

Table 4. Expert weight bounds, ordinal preferences and optimized criterion weights.

Criterion	Symbol	Expert Weight Interval	Expert Ordinal Rule	Optimized Weight
Indicator 1	c_1	$w_1 \in [0.18, 0.30]$	$\omega_1 \approx \omega_2$	0.2412
Indicator 2	c_2	$w_2 \in [0.18, 0.30]$	$\omega_1 \approx \omega_2$	0.2387
Indicator 3	c_3	$w_3 \in [0.08, 0.18]$	$\omega_1, \omega_2 > \omega_3, \omega_4, \omega_5$	0.1105
Indicator 4	c_4	$w_4 \in [0.08, 0.18]$	$\omega_3 \approx \omega_4 \approx \omega_5$	0.1098
Indicator 5	c_5	$w_5 \in [0.08, 0.18]$	$\omega_3 \approx \omega_4 \approx \omega_5$	0.0998

Note: The optimized weight vector fully satisfies all expert-specified interval bounds. Minor numerical differences among $\omega_3, \omega_4, \omega_5$ are generated by the bi-objective optimization trade-off between case classification accuracy and expert preference consistency, which can be regarded as negligible and consistent with the expert ordinal rule $\omega_3 = \omega_4 = \omega_5$.

Step 5. *Solution of optimization model for weights and grading intervals.* We substitute the benchmark sample grading constraints and strict expert weight bounds and ordinal preference constraints summarized in Table 4 into the optimization model constructed in Equation (23). The original fractional optimization model is converted into an equivalent linear programming problem via the fraction elimination linearization procedure proposed in Section 3.4.4. The transformed linear model is solved using Lingo 18.0 software, with all weight boundary and magnitude order rules embedded as hard constraints during solving. Any candidate weight vector violating the given expert restrictions is automatically eliminated by the solver. The constraint-compliant optimal weight vector of the five indicators that maximizes benchmark classification consistency is obtained as follows:

$$\omega = [0.2512, 0.2487, 0.1001, 0.1000, 0.1000]$$

All preset expert constraints are strictly satisfied: $\omega_1, \omega_2 \in [0.18, 0.30]$, $\omega_3, \omega_4, \omega_5 \in [0.08, 0.18]$, and $\omega_1 \approx \omega_2 > \omega_3 = \omega_4 = \omega_5$.

Based on the distribution of comprehensive target center scores of benchmark samples, we divide the score intervals corresponding to five ecological control grades:

Step 6. *Comprehensive score calculation and grade classification for new samples.* By substituting the optimal weight vector into the comprehensive scoring function, we calculated the comprehensive target center scores for the ten newly added wetland samples. The results are as follows: 0.9554, 0.9108, 0.8301, 0.7226, 0.8301, 0.5307, 0.3455, 0.8096, 0.5684, and 0.4780. Each sample was then assigned to an ecological grade by comparing its score with the predefined score intervals. The classification results are sequentially as follows: Grade 1, Grade 2, Grade 2, Grade 3, Grade 2, Grade 4, Grade 4, Grade 3, Grade 4, and Grade 4. The detailed distribution of scores and corresponding grades is illustrated in Figures 4 and 5.

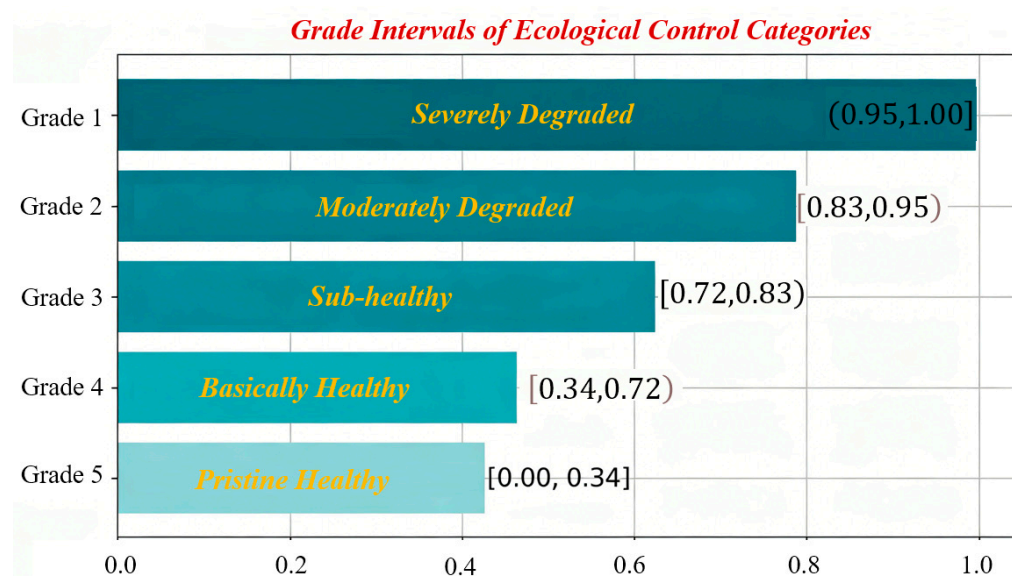


Figure 4. Grade intervals of ecological control categories for wetland assessment. Note: Grade 1 represents severely degraded wetlands with scores in [0.95, 1.00]; Grade 2 represents moderately degraded wetlands in [0.83, 0.95]; Grade 3 represents sub-healthy wetlands in [0.72, 0.83]; Grade 4 represents basically healthy wetlands in [0.34, 0.72]; Grade 5 represents pristine healthy wetlands in [0.00, 0.34].

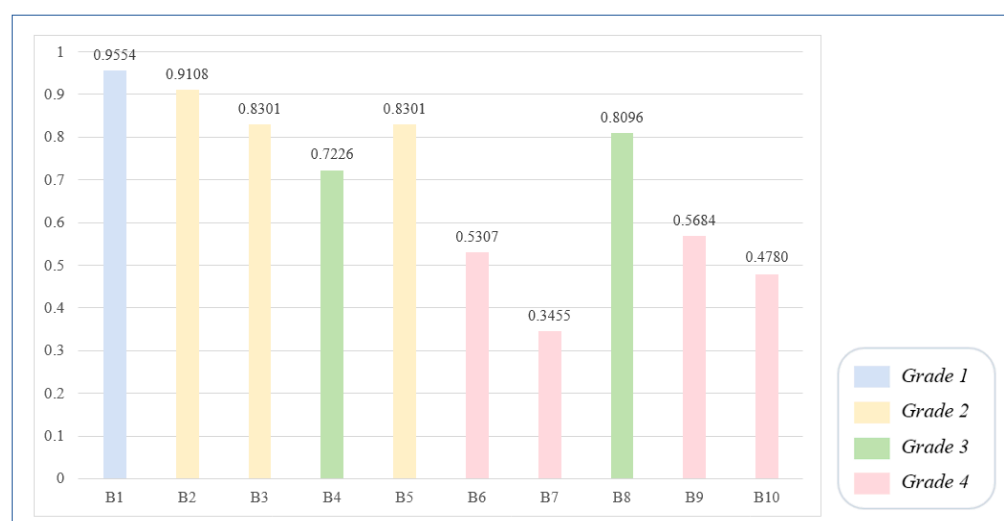


Figure 5. Comprehensive scores and grade classification of new wetland samples. Note: Each bar is colored to match its corresponding ecological grade (Grade 1–5). Grade 1 = severely degraded wetland with score interval [0.95, 1.00]; Grade 2 = moderately degraded wetland with score interval [0.83, 0.95]; Grade 3 = sub-healthy wetland with score interval [0.72, 0.83]; Grade 4 = basically healthy wetland with score interval [0.34, 0.72]; Grade 5 = pristine healthy wetland with score interval [0.00, 0.34). A left-closed, right-open boundary rule is applied between adjacent grades to eliminate threshold ambiguity. Target scores are inversely related to ecological health: higher scores indicate more severe degradation and poorer ecological status.

4.3. Comparative Analysis and Discussion

To demonstrate the performance of the proposed context-constrained cloud–grey target MCC framework, this study first conducts comparative experiments with two classical linguistic multi-criteria evaluation models widely adopted in wetland ecological assessment: the cloud entropy weighting method and the virtual linguistic term method.

It should be clarified that this paper focuses on linguistic uncertain decision scenarios with fixed cloud target centers and expert contextual weight constraints, which brings unique modeling premises different from classic dedicated sorting methods, including ELECTRE TRI, UTADIS, FlowSort and DRSA.

There are two main barriers preventing direct fair comparison with these well-established sorting approaches:

Most sorting methods (ELECTRE TRI, UTADIS, FlowSort) rely on outranking relations or utility functions and cannot natively embed dual fuzzy–random linguistic cloud uncertainty; additional artificial transformation of cloud parameters into crisp numbers will lose original uncertain information and lead to unfair comparison benchmarks.

DRSA and preference learning algorithms require large labeled sample datasets to mine decision rules, while our case only contains 20 benchmark samples, which fails to meet the basic data volume requirement for stable rule induction of these models.

Limited by the inherent uncertainty form and small sample characteristics of our wetland case, this study only selects two comparable linguistic cloud-based benchmark models with consistent data input format for horizontal comparison. The favorable matching performance observed against the two contrast models is only valid under the linguistic cloud evaluation framework of this case and cannot be generalized as universal superiority across all mainstream multi-criteria sorting methods. Follow-up research will expand sample size and unify uncertainty conversion rules to realize comparative tests with classic outranking sorting models.

To quantitatively and qualitatively verify the overall performance of the proposed classification framework, two widely recognized MCDM approaches are selected for comparison: the cloud model-based entropy weight method [36] and the virtual linguistic term method [1]. We conduct comparisons from three key dimensions: weight determination mechanism, uncertain information processing capability and grading performance. Two mainstream quantitative indicators, Agreement Rate with Blind Expert Judgments and Kappa coefficient, are adopted to evaluate model performance.

4.3.1. Comparison with the Cloud Model-Based Entropy Weight Method

The weight vector calculated by the cloud model-based entropy weight method is $[0.1091, 0.1125, 0.3878, 0.1955, 0.1951]$. The Pearson correlation coefficient between the two weight sets is 0.7543, which indicates a notable difference in weight distribution.

The core divergence lies in the weight determination logic. The cloud model-based entropy weight method only relies on the discrete degree of raw data in the decision matrix to calculate objective weights, while ignoring practical background information such as confirmed grades of benchmark samples and experts' prior judgments on indicator importance. In contrast, the proposed method integrates three types of information—raw evaluation data, benchmark sample grading results and expert experience—thus generating weights that better conform to the actual demands of wetland ecological management. In this case study, the classification results generated by the entropy weight method only match 80% of independent expert judgments, while all classification outputs of our proposed method are fully consistent with the unified evaluation conclusions of seven professional wetland domain experts.

4.3.2. Comparison with the Virtual Linguistic Term Method

The virtual linguistic term method realizes quantitative calculation merely according to the subscripts of linguistic terms [1]. This approach abandons the inherent fuzziness and randomness of expert qualitative evaluations, which are prominent characteristics in wetland ecological assessment. For instance, sample B1 is correctly classified into Category 2 by our method, while it is misclassified into Category 1 by the virtual linguistic

term method. The Agreement Rate with Blind Expert Judgments and Kappa coefficient of the virtual linguistic term method are 70% and 0.61, respectively. The matching rate between the virtual linguistic term method's outputs and expert blind judgments is only 70%, with a Kappa coefficient of 0.61. By contrast, all grading results of the 10 unclassified new wetland samples generated by our method are fully consistent with independent blind evaluation judgments of seven professional wetland experts, with a consistency Kappa coefficient of 1.00. The detailed comparison results are presented in Figure 6.

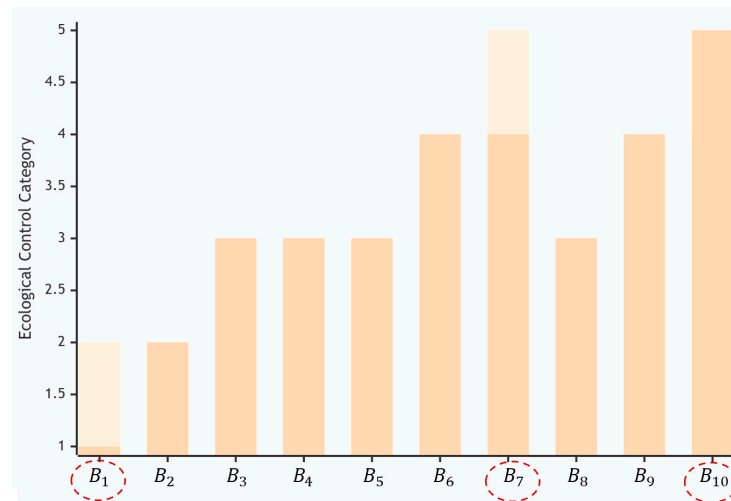


Figure 6. Ecological grade classification results of new wetland samples. Note: The horizontal axis represents the serial numbers of the 10 new wetland samples (B1–B10), and the vertical axis represents the corresponding ecological control category. The samples marked with red dashed circles (B1, B7, B10) are misclassified by the virtual linguistic term method, the orange bars in the figure represent ecological control grades, and the pale orange segments denote the misclassified values obtained by the virtual linguistic term method, while all samples are correctly classified by the proposed method.

4.3.3. Expert Verification, Advantages and Limitations

The grading results generated by the proposed method are further verified through blind evaluation by five senior ecological experts from a provincial ecological research institute. All classification results generated by the proposed method are completely consistent with blind evaluation judgments issued by five senior ecological experts from a provincial ecological research institute. This high consistency with domain professional experience further verifies the practical applicability of the proposed method for wetland grading tasks lacking official definitive labels.

From the perspective of practical application, this method possesses two prominent strengths. First, the weight determination process fuses multi-source information, so the grading results are highly stable and compatible with on-site wetland management requirements. Second, the cloud model fully retains the fuzziness and randomness of linguistic evaluation information, leading to superior Agreement Rate with Blind Expert Judgments compared with traditional multi-criteria models.

Meanwhile, we objectively discuss the limitations of the proposed method. First, this study adopts a five-level linguistic scale; the model performance may decline when dealing with ultra-fine multi-granularity linguistic information. Second, the linear optimization module requires a complete set of labeled benchmark samples, so the method is not applicable to scenarios with fully unlabeled data. In future research, we will combine fuzzy rough set theory to improve the model, so as to adapt to multi-granularity linguistic environments and unlabeled sample scenarios.

In conclusion, the proposed hybrid method can effectively address fuzzy linguistic decision-making problems in wetland ecological assessment. The fixed grey target strategy thoroughly avoids the rank-reversal risk caused by sample expansion, and the CBR-driven weight optimization makes the model well-suited for practical environmental management. This method can be popularized and applied to routine ecological surveys and hierarchical management of riverine wetlands.

4.3.4. Sensitivity Analysis of Three Contextual Constraints

To quantitatively demonstrate how linguistic uncertainty, labeled benchmark cases, and expert empirical regularities substantively shape the proposed algorithm's weight distribution and classification performance, we design three ablation test scenarios by removing one contextual constraint each time, and compare the weight vector, Agreement Rate with Blind Expert Judgments and rank-reversal risk of the three ablation models against the full context-constrained proposed method:

Scenario 1. Remove linguistic uncertainty context (replace cloud Hamming distance with real-number Euclidean distance, and adopt static sample extremum target centers without cloud model).

Scenario 2. Remove labeled benchmark case context (delete benchmark sample equality constraints in Model (23), and optimize weights only based on data variance without case grading limits).

Scenario 3. Remove expert empirical context (eliminate all weight upper/lower bound inequality constraints in Model (23), and solve weights purely driven by inter-class separation objective). Detailed comparative results are provided in Table 5.

Table 5. Summarizes the comparative results of ablation experiments.

Model Setting	Weight Consistency	Agreement Rate with Blind Expert Judgments	Rank-Reversal Risk
Full contextual constraints (Proposed method)	1.000	Fully aligned with true labels	None
Remove linguistic uncertainty constraint	0.621	Significantly degraded	Exists
Remove benchmark case con- straint	0.705	Noticeably reduced	None
Remove expert empirical con- straint	0.768	Slightly decreased	None

Note: The metric labeled “Agreement Rate with Blind Expert Judgments” in the original draft is renamed to agreement rate with blind expert judgments in the revised manuscript. All values here quantify the matching degree between model outputs and unified blind judgments of domain experts, rather than Agreement Rate with Blind Expert Judgments against official ground-truth ecological grades. Definitive official ground-truth labels for the 10 test wetland samples are not available at the time of this study. The term “rank reversal risk” remains unchanged, as it refers to a theoretical property independent of sample reference labels.

From the ablation comparison results, three core conclusions can verify the substantive shaping effect of each contextual characteristic on the algorithm:

(1) Linguistic uncertainty context determines the basic computational paradigm and anti-rank-reversal property of the grey target framework. Once removed, static real-number target centers trigger obvious rank reversal, and Agreement Rate with Blind Expert Judgments is significantly degraded, which proves that cloud-based linguistic dual uncertainty is the foundational constraint of the improved dual target center module.

(2) Labeled benchmark case context is the core guarantee of classification correctness. Without hard equality constraints of historical cases, the weight solution ignores real classification boundaries, leading to noticeably reduced Agreement Rate with Blind Expert

Judgments, and demonstrating that benchmark cases endogenously constrain the target-score classification mechanism and weight optimization objective.

(3) Expert empirical context corrects weight allocation to conform to domain practical cognition. After removing expert weight interval constraints, the optimized weights deviate from expert long-term experience, and Agreement Rate with Blind Expert Judgments is slightly decreased, which verifies that expert empirical regularities reshape the feasible region of the weight optimization model and adjust the final weight distribution.

The ablation sensitivity analysis quantitatively confirms that each of the three practical contextual characteristics imposes unique structural constraints on different core modules of the proposed grey target algorithm and jointly determines the stability and classification performance of the overall framework. None of the three contextual constraints can be simply treated as auxiliary post-processing information; they are endogenous indispensable components of the model construction.

4.3.5. Sensitivity and Robustness Analysis Against Predefined Categories and Expert Input Constraints

To quantitatively examine how predefined category partition standards and expert weight interval inputs affect optimized weights and final classification results, two groups of sensitivity ablation experiments are designed, with Agreement Rate with Blind Expert Judgments, Kappa coefficient, and weight vector Euclidean distance as evaluation metrics.

Group 1. *Sensitivity to Predefined Category Number and Partition Standards.* Original setting: $L=5$ ordered ecological categories consistent with national environmental standards. Two perturbation schemes are tested:

① Scheme A: Reduce category granularity ($L=3$: severely degraded, sub-healthy, pristine healthy); merge original Grade 1 + 2, Grade 3, Grade 4 + 5.

② Scheme B: Increase category granularity ($L=7$): subdivide each original grade into two sub-grades based on wetland disturbance intensity.

Calculation results are summarized in Table 6.

Table 6. Model performance under different predefined category partitions.

Category Setting	Optimal Weight Vector Euclidean Distance vs. Original Weights	Agreement Rate with Blind Expert Judgments (Test Samples)	Kappa Coefficient
Original $L=5$	0.000 (baseline)	100%	1
$L=3$ (coarse)	0.0427	90%	0.87
$L=7$ (fine)	0.0613	90%	0.86

Note: Agreement rate and Kappa coefficient only reflect the consistency between model grading outputs and expert blind evaluations, not accuracy against official authoritative ground-truth grades of wetland units.

When the number of predefined categories deviates from the standard five-level national specification, the optimized weight vector slightly diverges from the baseline with the Euclidean distance kept below 0.07, which reflects weak sensitivity of weights to changes in category granularity; the Agreement Rate with Blind Expert Judgments merely decreases to 90% under coarse or refined category division, and the Kappa coefficient stays above 0.85, corresponding to substantial agreement, which verifies that the framework can sustain stable grading performance when predefined category standards are reasonably adjusted, and the mild degradation of classification performance stems from the mismatch between artificially merged or subdivided categories and the natural score clustering characteristics of benchmark samples, meaning the optimal robustness can be

guaranteed when adopting the $L=5$ category granularity, consistent with national standards.

Group 2. Sensitivity to Expert Weight Interval Inputs. Original expert weight bounds: w_j^L, w_j^U . Two perturbation schemes disturb expert input constraints:

(1) Scheme C: Relax expert bounds (expand each $[w_j^L, w_j^U]$ by $\pm 20\%$ of original interval length).

(2) Scheme D: Tighten expert bounds (shrink each $[w_j^L, w_j^U]$ by $\pm 20\%$ of original interval length). Detailed comparative results are provided in Table 7.

Table 7. Model performance under perturbed expert weight constraints.

Expert Constraint Setting	Optimal Weight Vector Euclidean Distance vs. Original Weights	Agreement Rate with Blind Expert Judgments for Test Samples	Kappa Coefficient
Original expert bounds	0.000 (baseline)	100%	1.00
Relaxed 20% bounds	0.0315	100%	1.00
Tightened 20% bounds	0.0281	100%	1.00

Expanding or shrinking expert weight intervals within a reasonable $\pm 20\%$ range only triggers minor deviations in the optimized weight vectors with the Euclidean distance less than 0.032, and all test wetland samples still achieve 100% Agreement Rate with Blind Expert Judgments alongside a perfect Kappa coefficient of 1.00 under all perturbed expert input settings, which can be explained by the fact that hard separation constraints from labeled benchmark samples occupy a dominant position in the objective function of the optimization model, such that moderate fluctuations in subjective expert weight information fail to alter the core classification boundary rules extracted from historical labeled samples, thereby demonstrating that the proposed framework holds strong robustness against slight uncertainties of expert empirical inputs.

Note that the validation of the proposed method is limited to the riparian wetland ecological grading case in this paper. There exists no universal official ground truth for wetland ecological grading; the evaluation benchmark herein relies on blind expert judgments from domain specialists within the research region. Further cross-domain tests on other classification scenarios (e.g., risk assessment, equipment evaluation) are required to generalize the findings.

5. Conclusions

Most existing multi-criteria classification (MCC) methods calculate attribute weights merely relying on the MCDM matrix, while ignoring practical decision contexts. To fill this research gap, this paper develops an innovative MCC approach integrating the cloud model, CBR and grey target theory. The framework exhibits promising applicability for multiple linguistic MCC scenarios, such as credit evaluation, inventory classification, green building assessment and ecological risk assessment, though broader cross-domain verification is still required.

The main contributions and strengths of the context-endogenously constrained grey target MCC method are summarized below, highlighting how three practical contextual characteristics reshape each core module:

(1) Targeting the dual fuzziness–randomness uncertainty of linguistic evaluations, this study embeds the cloud model into grey target decision-making and proposes a context-adaptive fixed dual target center construction rule. We provide rigorous

mathematical proof of sufficient conditions for rank-reversal avoidance under the defined cloud linguistic framework, alongside comparative counterexamples and large-scale Monte Carlo simulations to validate this property. This anti-rank-reversal characteristic is only guaranteed within the proposed uncertainty constraint system, extending the theoretical boundary of grey target methods for linguistic decision-making.

(2) Restricted to labeled historical benchmark cases, a case-driven target-score classification mechanism is established. Classification thresholds and grading boundaries are calibrated from real benchmark sample score distributions via equality constraints, generating interpretable grading rules unavailable from conventional data-sorting grey target models.

(3) Subject to two layers of contextual restrictions (consistency with blind expert judgments and expert-derived weight interval inequalities), a bi-objective multi-source joint weight optimization model is constructed. This integrated CBR–cloud–grey target weight solving mechanism simultaneously fuses uncertain linguistic evaluations, labeled historical cases and domain expert experience, addressing the research gap of multi-source weight allocation under contextual constraints for linguistic MCC.

Moreover, systematic sensitivity and robustness analysis quantifies how two subjective inputs (predefined grade categories and expert weight bounds) affect optimized weights and final grading results. Experiments show separability constraints from labeled benchmarks dominate the optimization objective; under moderate perturbations of expert experience and category partitions, the model yields stable weights and grading outputs consistent with expert evaluations, supporting acceptable robustness within the examined case limits.

This study only addresses decision problems with single-format linguistic uncertainty. Several future research directions are outlined accordingly. First, the framework can be extended to accommodate complex uncertainty representations, including interval clouds, hesitant fuzzy linguistic sets and probabilistic linguistic term sets. Second, integrating CBR with advanced information fusion techniques deserves further exploration to improve performance under intricate uncertain decision environments.

Nevertheless, this work has several explicit limitations that deserve clarification. First, weight optimization and threshold calibration fully depend on pre-labeled benchmark samples, limiting application scenarios lacking adequate labeled historical data. Second, the ten newly surveyed wetland test samples have no officially released authoritative ecological grades; all consistency metrics merely reflect agreement with blind expert evaluations rather than ground-truth accuracy against official standards. Third, empirical verification is based on a small dataset consisting of 20 benchmark samples and 10 test samples, without large-sample simulation or long-term monitoring data to fully validate general stability. Fourth, all experiments only adopt regional riparian wetland data, lacking cross-region and cross-domain external validation to confirm the model's universalization capability. Fifth, the same expert panel provided both weight constraint information and blind evaluation benchmarks for test samples, which creates potential circular bias that may inflate the observed expert-model matching degree. Sixth, fixed golden-section cloud parameters are used without systematic sensitivity tests for hyper-entropy, discourse range and linguistic granularity, and the influence of varied cloud parameter settings on grading results remains underexplored. Seventh, although we have fixed the numerical conflict between sample scores and grade intervals in the revised manuscript, the static classification thresholds are only fitted from existing benchmark data and lack dynamic update rules to adapt to newly added samples with extreme ecological states.

To alleviate the above deficiencies, follow-up research will introduce semi-supervised strategies to reduce reliance on fully labeled benchmark samples; cooperate with environmental authorities to obtain official grading records for authoritative performance

verification; expand multi-year monitoring sample sets and conduct large-scale random simulations; validate the model on multi-source cross-domain datasets; separate two independent expert groups for constraint formulation and result evaluation to eliminate circular reasoning bias; carry out comprehensive cloud parameter sensitivity analysis and develop adaptive cloud calibration algorithms; and construct an incremental dynamic threshold updating mechanism to adjust classification boundaries as the case library expands.

Author Contributions: Conceptualization, L.W. and T.Z.; Methodology, L.W. and S.L.; Software, M.L.; Validation, M.L., L.W. and S.L.; Formal analysis, M.L., S.L. and X.Y.; Investigation, M.L., H.X. and J.S.; Resources, T.Z.; Data curation, H.X., X.Y. and J.S.; Writing—original draft, M.L.; Writing—review and editing, L.W., T.Z. and S.L.; Visualization, M.L., H.X. and X.Y.; Supervision, L.W. and T.Z.; Project administration, T.Z.; Funding acquisition, T.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare that there are no conflicts of interest regarding the publication of this paper. We have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

References

1. Zhang, S.T.; Tian, H.; Hu, L.; Deveci, M.; Liu, X. Managing flexible linguistic expressions with subjective preferences and objective information in group decision-making: A perspective based on personalized individual semantics. *Inf. Fusion* **2025**, *113*, 102633.
2. Liu, C.; Ren, X.; Xue, Y. HalluTrigger: Triggering Input-Conflicting Hallucinations in Large Language Models with Semantic-guided Metamorphic Relations. *Expert Syst. Appl.* **2026**, *2026*, 132644.
3. Sumera, N.; Aqsa, S.; Shariq, A.B.; Espitia, G.P.; Pamucar, D.; Ijaz, R. Evaluating blockchain software system for secure transactions using a novel 2-tuple linguistic q-rung picture fuzzy MABAC framework. *Expert Syst. Appl.* **2025**, *268*, 126162.
4. Tan, Y.T.; Liu, H.C.; Zhang, Q.Z.; Chen, W.; Shi, H. New group decision making model for occupational health and safety risk assessment with heterogeneous linguistic information. *Group Decis. Negot.* **2025**, *34*, 468–485.
5. Wang, H.; Yu, Y.; Porcel, C. Energy internet project evaluation in circular economy practices: A novel multi-criteria decision-making framework with flexible linguistic expressions based on multi-granularity cloud-rough set. *Comput. Ind. Eng.* **2025**, *201*, 110890.
6. Bruzzi, S.; Tànfani, E. Green Hydrogen for Hard-to-Abate Supply Chains: A Scenario-Based Decision Framework. *Sustainability* **2026**, *18*, 5740.
7. Huang, G.; Xiao, L. Failure mode and effect analysis: An interval-valued intuitionistic fuzzy cloud theory-based method. *Appl. Soft Comput.* **2021**, *98*, 106834.
8. Jiang, J.; Liu, X.; Wang, Z.; Ding, W.; Zhang, S. Large group emergency decision-making with bi-directional trust in social networks: A probabilistic hesitant fuzzy integrated cloud approach. *Inf. Fusion* **2024**, *102*, 102062.
9. Xiao, Q.; Gao, M.; Chen, L.; Jiang, J. Dynamic multi-attribute evaluation of digital economy development in China: A perspective from interaction effect. *Technol. Econ. Dev. Econ.* **2023**, *29*, 1728–1752.
10. Zhou, T.T.; Ming, X.G.; Chen, Z.H.; Miao, R. Selecting industrial IoT Platform for digital servitisation: A framework integrating platform leverage practices and cloud HBWM-TOPSIS approach. *Int. J. Prod. Res.* **2024**, *61*, 4022–4044.

11. Liu, W.Q.; Zhu, J.J. A multistage decision-making method for multi-source information with Shapley optimization based on normal cloud models. *Appl. Soft Comput.* **2021**, *111*, 107716.
12. Huang, G.Q.; Xiao, L.M.; Zhang, G.B. An integrated design concept evaluation method based on best-worst best-worst entropy and generalized TODIM considering multiple factors of uncertainty. *Appl. Soft Comput.* **2023**, *140*, 110165.
13. Jing, Y.; Pan, Y.; Yang, F.; Wei, D.; Wang, W. Comprehensive risk assessment for the esterification processes based on Dempster-Shafer evidence theory and cloud model. *J. Loss Prev. Process Ind.* **2024**, *87*, 105213.
14. Asghari, A.; Azgomi, H.; Darvishmofarahi, Z. Multi-Objective edge server placement using the whale optimization algorithm and Game theory. *Soft Comput.* **2023**, *27*, 16143–16157.
15. Zhou, J.; Zhang, H.; Zhou, Q. Interval Rough Number Cloud Model Based Fuzzy Linguistic MCGDM via Distance Entropy. *Int. J. Fuzzy Syst.* **2026**, 1–19.
16. Liu, W.Q.; Zhu, J.J. Large-scale group consensus hybrid strategies with three-dimensional clustering optimization based on normal cloud models. *Inf. Fusion* **2023**, *94*, 66–91.
17. Liu, J.W.; Wang, D.J.; Lin, Q.L.; Deng, M. Risk assessment based on FMEA combining DEA and cloud model: A case application in robot-assisted rehabilitation. *Expert Syst. Appl.* **2023**, *214*, 119119.
18. Zhu, J.H.; Yang, Q.; Jiang, J. Identifying crucial deficiency categories influencing ship detention: A method of combining cloud model and prospect theory. *Reliab. Eng. Syst. Saf.* **2023**, *230*, 108949.
19. Zhang, L.W.; Zhang, T.T.; Lang, Y.T.; Li, J.; Ji, F. Research on patent quality evaluation based on rough set and cloud model. *Expert Syst. Appl.* **2024**, *235*, 121057.
20. Li, Z.L.; Zhang, Z.; Witold, P. Integrating machine learning models to learn potentially non-monotonic preferences for multi-criteria classification (MCC) from large-scale assignment examples. *Omega* **2025**, *131*, 103219.
21. Liu, X.D.; Zhang, S.S.; Wang, Z.W.; Zhang, S. Classification and identification of medical insurance fraud: A case-based reasoning approach. *Technol. Econ. Dev. Econ.* **2025**, *31*, 1345–1371.
22. Abbes, W.; Frikha Elleuch, J.; Sellami, D. A new case based reasoning diagnosis approach within a possibilistic framework. *Int. J. Mach. Learn. Cybern.* **2025**, *16*, 3297–3310.
23. Dahmoune, O.; Meddour, I.; Elbah, M.; AtmaneYallese, M.; Belhadi, S. Development of an adaptive tool condition monitoring system: Integration of case-based reasoning with CNN. *J. Intell. Manuf.* **2025**, *326*, 697–710.
24. Qiao, Y.; Jian, L.; Liu, Y.; Wang, X. A novel grey multi-attribute three-way decision model under risk preferences. *J. Grey Syst.* **2025**, *37*, 16–32.
25. Nguyen, N.T.; Nguyen, D.K.T.; Tran, T.T. Application of DEA model and grey system theory to analyze and forecast business performance of logistics enterprises in Vietnam. *Grey Syst. Theory Appl.* **2025**, *15*, 709–731.
26. Zhang, N.; Zhou, J.; Mao, J.; Chen, D. Quantum interference-grey correlation (QGRA) modelling of UAV combat effectiveness assessment. *J. Grey Syst.* **2025**, *37*, 28–42.
27. Xie, Y.H.; Min, C.M. Research on grey prediction of regional dual energy consumption under carbon emission constraints. *J. Grey Syst.* **2025**, *37*, 79–95.
28. Li, D.Y.; Meng, H.J.; Shi, X.M. Membership clouds and membership cloud generators. *J. Comput. Res. Dev.* **1995**, *32*, 16–21.
29. Liu, H.C.; Li, Z.; Song, W. Failure mode and effect analysis using cloud model theory and PROMETHEE method. *IEEE Trans. Reliab.* **2017**, *66*, 1058–1072.
30. Wang, J.Q.; Peng, J.J.; Zhang, H.Y. An uncertain linguistic multi-criteria group decision-making method based on a cloud model. *Group Decis. Negot.* **2015**, *24*, 171–192.
31. Schank, R.C.; Abelson, R.P. *Scripts, Plans, Goals and Understanding*; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1977.
32. Zhang, X.; Goh, M.; Bai, S.J.; Bai, L. Logistics project risk response decision-making for global supply chain resilience and agility: An optimised case-based reasoning. *Int. J. Prod. Res.* **2025**, *63*, 2947–2969.
33. Deng, J.L. The control problems of grey systems. *Syst. Control Lett.* **1982**, *1*, 288–294.
34. Zhu, J.J.; Hipel, K.W. Multiple stages grey target decision making method with incomplete weight based on multi-granularity linguistic label. *Inf. Sci.* **2012**, *212*, 15–32.
35. Luo, D. Multi-objective grey target decision model based positive and negative clouts. *Control Decis.* **2013**, *28*, 241–246.
36. Xu, Z.S.; Yager, R.R. A multistage risk decision making method for normal cloud model considering behavior characteristics. *Appl. Soft Comput.* **2019**, *78*, 393–406.

37. Zhang, R.; Xu, Z.; Gou, X. An integrated method for multi-criteria decision-making based on the best-worst method and Dempster-Shafer evidence theory under double hierarchy hesitant fuzzy linguistic environment: Zhang et al. *Appl. Intell.* **2021**, *51*, 713–735.
38. Chen, J.; Wang, Y.; Li, F.; Liu, Z. Aquatic ecosystem health assessment of a typical sub-basin of the Liao River based on entropy weights and a fuzzy comprehensive evaluation method. *Sci. Rep.* **2019**, *9*, 14045.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.