Bulletin of Electrical Engineering and Informatics
Vol. 15, No. 6, December 2026, pp. 5584~5594
ISSN: 2302-9285, DOI: 10.11591/eei.v15i6.13862

Stability and robustness of decision rankings across objective weighting methods in multi-criteria decision-making

Itishom Al Khoiry[1], Dhea Rizky Amelia[2]
[1]Department of Digital Business, Universitas Persatuan Guru Republik Indonesia Semarang, Semarang, Indonesia
[2]Department of Development Economics, Universitas Negeri Semarang, Semarang, Indonesia

Article Info

Article history:

Received Jul 2, 2026
Revised Sep 10, 2026
Accepted Oct 4, 2026

Keywords:

Decision support systems
Multi-criteria decision-making
Objective weighting methods
Ranking stability
Robustness analysis
Weight perturbation

ABSTRACT

Objective weighting methods derive criterion importance directly from data, but a method producing consistent rankings across aggregation models may not be robust to weight changes. This study jointly evaluates cross-model ranking stability and perturbation robustness in multi-criteria decision-making (MCDM) across five objective weighting methods: entropy, criteria importance through intercriteria correlation (CRITIC), method based on the removal effects of criteria (MEREC), principal component analysis (PCA), and standard deviation (SD). Two datasets were used: a renewable energy case (five alternatives and ten criteria) and a supplier-selection case (ten alternatives, eleven criteria) for external validation. Each weight vector was applied to five MCDM aggregation models: visekriterijumska optimizacija i kompromisno resenje (VIKOR), complex proportional assessment (COPRAS), multi-objective optimization on the basis of ratio analysis (MOORA), evaluation based on distance from average solution (EDAS), and technique for order preference by similarity to ideal solution (TOPSIS). Stability was measured using rank correlation, rank shift, and rank reversal; robustness through criterion-weight perturbation. The two diagnostics are positively associated but not interchangeable. Entropy and MEREC achieved the highest stability and robustness in both cases, whereas PCA was stable across models yet highly sensitive to weight perturbation, indicating that reliability cannot be judged from cross-model stability alone.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Itishom Al Khoiry
Department of Digital Business, Universitas Persatuan Guru Republik Indonesia Semarang
Sidodadi Timur Street No. 24, Karangtempel, Semarang, Central Java, Indonesia
Email: itishom@upgris.ac.id

1. INTRODUCTION

Decision problems in engineering, environmental, economic, and managerial settings rarely reduce to a single objective, requiring alternatives to be evaluated across multiple dimensions that differ in units, preference directions, and trade-offs. This has motivated multi-criteria decision-making (MCDM), a family of techniques for ranking alternatives against multiple competing criteria [1]. By transforming heterogeneous decision information into a structured preference order, MCDM provides a transparent basis for decisions where intuitive judgment alone is insufficient [2].

Within an MCDM model, criteria weighting plays a decisive role, as even small changes in the weight vector can alter rankings and the final recommendation [3]. Objective weighting methods have gained increasing attention because they derive criteria importance directly from the decision matrix, providing reproducible and transparent weighting with reduced reliance on expert judgment [4]. However, objective weighting methods rely on different statistical principles, so the same decision matrix may generate different

weight distributions and final rankings [3]–[6]. Similar variability has been reported in domains such as renewable energy selection [6] and environmental sustainability assessment [7]. The choice of objective weighting method should therefore be treated as an active factor in the decision process, not as a neutral preprocessing step.

Ranking outcomes are also affected by the aggregation model. MCDM models differ in their ranking philosophies, normalization procedures, and reference solutions, so identical inputs may produce systematically different rankings [8]. The reliability of a weighting method therefore cannot be judged from a single aggregation model, since the same weight vector may behave differently under different aggregation logics [9]. Objective weighting methods should thus be evaluated not only by the weights they produce, but by the consistency of the rankings they generate across aggregation contexts.

Despite the rapid adoption of objective weighting methods, two gaps limit confidence in the rankings they produce. First, aggregation models are frequently selected on pragmatic grounds, such as familiarity or software availability, rather than verified across alternative models [10], even though different aggregation models and normalization procedures can alter a substantial share of ranking positions, with reported changes of 20 to 40% in real applications [8]. Second, robustness to weight variation is rarely examined: a bibliometric analysis of 1,374 MCDM articles published between 2000 and 2023 found that only 9.46% incorporated sensitivity analysis to assess result robustness [11]. As a result, the reliability of an objective weighting method is often judged within a single aggregation framework and without a perturbation check, leaving uncertain whether the ranking would survive a change of aggregation model or a small change in the weights. While some studies have compared ranking stability across several aggregation models [12], the joint assessment of cross-model stability and perturbation robustness as a basis for selecting an objective weighting method has received limited attention. Prior studies typically stop at rank correlation across models, leaving open the question of whether a stable ranking also resists weight changes.

This study investigates how five objective weighting methods, namely entropy, criteria importance through intercriteria correlation (CRITIC), method based on the removal effects of criteria (MEREC), principal component analysis (PCA), and standard deviation (SD), behave when implemented across five widely used MCDM aggregation models: visekriterijumska optimizacija i kompromisno resenje (VIKOR), complex proportional assessment (COPRAS), multi-objective optimization on the basis of ratio analysis (MOORA), evaluation based on distance from average solution (EDAS), and technique for order preference by similarity to ideal solution (TOPSIS) [13]–[17]. The evaluation separates two reliability properties often treated as one. Cross-model stability measures the consistency of rankings across aggregation logics, using Spearman rank correlation, maximum rank shift, and rank reversal [18]. Perturbation robustness measures the resistance of rankings to systematic weight changes, through a test of the minimum perturbation that alters the complete ranking [19], [20]. These need not coincide: a method may be highly consistent across models yet collapse under minor weight changes, especially when its weights are nearly uniform. Treating them as separate diagnostics provides a more comprehensive evaluation of reliability. Unlike rank-correlation-based comparisons, this study tests whether stable rankings also resist weight perturbation, treating stability and robustness as complementary diagnostics. This joint assessment is the study's central contribution, intended to guide practitioners in renewable energy planning and supplier selection toward weighting methods that yield defensible rankings.

2. METHOD

2.1. Research framework and data

The analysis, summarized in Figure 1, applies five objective weighting methods across five MCDM aggregation models to produce twenty-five complete rankings, which are then evaluated for cross-model stability and perturbation robustness. All computations were implemented in Python.

The empirical problem is a renewable energy source selection task, a class widely used in MCDM studies [21], [22]. The data are secondary values adapted from Lee and Chang [23]. Five alternatives are evaluated, namely solar PV, wind, hydro, biomass, and geothermal, against ten criteria spanning economic, technical, environmental, and social aspects. Each criterion is either a benefit or a cost criterion, as listed in Table 1, and the performance values that serve as the input for all procedures are given in Table 2.

2.2. Objective weighting procedures

Five objective weighting methods generate alternative sets of criteria weights from the same decision matrix; all infer importance from the internal data structure but differ in what they reward. The entropy method assigns higher weight to criteria whose values are more dispersed across alternatives, as greater dispersion carries more information [24]. CRITIC combines the contrast intensity of a criterion with its conflict against the others, favoring criteria that are both variable and weakly correlated with the rest [4]. MEREC measures each criterion by the change its removal causes in aggregate performance, so that more

influential criteria receive more weight [5]. PCA derives weights from the contribution of each criterion to the principal components extracted from the z-score-standardized transformed matrix. In this study, components were obtained from the correlation matrix without rotation, retained using the Kaiser criterion, and converted into criterion weights through explained-variance-weighted absolute loadings before normalization [25]. The SD method weights criteria in proportion to the dispersion of their values, on the same discriminative logic as entropy but without the information-theoretic transformation [3], [4]. Because these methods conceptualize criterion importance differently, they yield weight vectors ranging from concentrated to near-uniform, which is central to the analysis that follows.

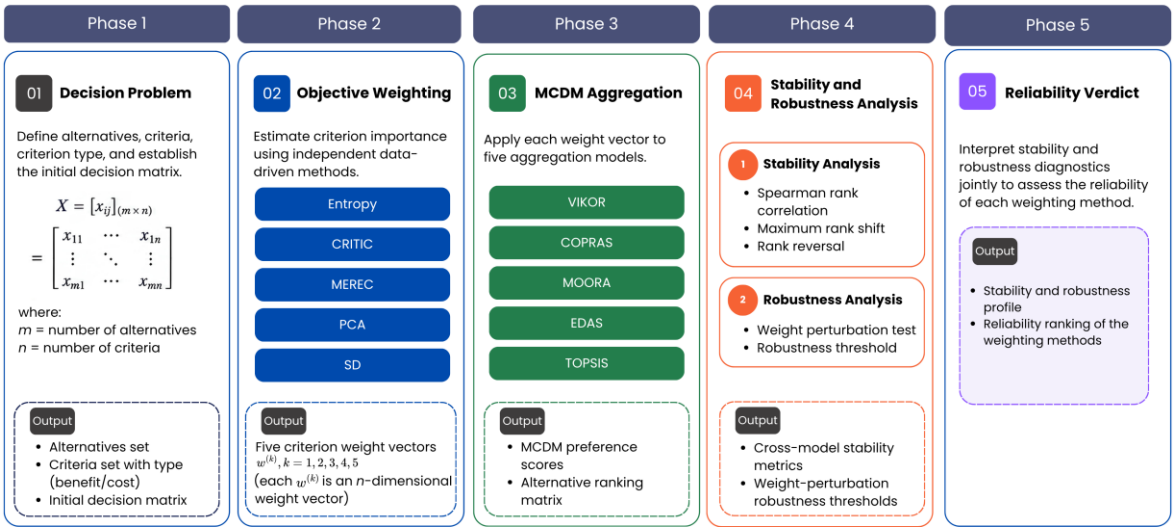


Figure 1. Research framework

Table 1. Evaluation criteria

Code	Criterion	Unit	Type
C1	Investment cost	USD/kW	Cost
C2	O&M cost	USD/kW/yr	Cost
C3	Electric cost	USD/kWh	Cost
C4	Efficiency	%	Benefit
C5	Capacity factor	%	Benefit
C6	Technical maturity	Scale 1 to 5	Benefit
C7	Greenhouse gas (GHG) emission	g CO_2eq/kWh	Cost
C8	Land use	m^2/kW	Cost
C9	Job creation	person/kWh	Benefit
C10	Social acceptance	Scale 1 to 5	Benefit

Table 2. Decision matrix

Alternative	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
A1 solar	4550	30	6.74	20	15	5	85	150	0.87	4.76
A2 wind	3005	60.86	2.4	35	27	4	26	200	0.17	4.51
A3 hydro	2040	14.85	1.7	90	25	5	26	500	0.27	4.19
A4 biomass	3370	99.4	3.25	25.3	54	3	45	222	0.21	3.78
A5 geothermal	3920	112.6	4.93	11.4	71.7	2	50	100	0.25	4.11

2.3. Multi-criteria decision-making aggregation procedures

Each weight vector is then used to rank the alternatives with five MCDM models that embody distinct decision logics, giving broad methodological coverage and reducing reliance on any single ranking mechanism. VIKOR determines a compromise solution from the closeness of each alternative to the ideal, balancing group utility against individual regret [13]. COPRAS evaluates alternatives proportionally by treating benefit and cost criteria separately [14]. MOORA combines normalized benefit and cost contributions into a single preference score [15]. EDAS scores each alternative by its positive and negative

distances from the average solution rather than from an ideal point [16]. TOPSIS ranks alternatives by their relative closeness to the positive ideal and distance from the negative ideal solution [17]. Applying all five to each weight vector produces the twenty-five rankings analyzed below.

2.4. Stability evaluation

Cross-model stability is assessed as the consistency of the five rankings a single weighting method produces across VIKOR, COPRAS, MOORA, EDAS, and TOPSIS. The five ranking lists are compared pairwise, giving ten comparisons summarized by the mean and minimum Spearman rank correlation, the maximum rank shift, and the average number of rank reversals. The Spearman coefficient is:

$$\rho = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} \quad (1)$$

Where d_i is the difference between the ranks of alternative i and n is the number of criteria. The maximum rank shift is the largest change in the position of any alternative across the models:

$$MRS = max_i\ (R_{max,i} - R_{min,i}) \quad (2)$$

Where $R_{max,i}$ and $R_{min,i}$ denote the worst and best ranks of alternative i, respectively. Rank reversal occurs when the relative order of two alternatives changes between ranking methods:

$$(R_i^a - R_j^a)(R_i^b - R_j^b) < 0 \quad (3)$$

Where a smaller number of rank reversals indicates more stable ranking results.

2.5. Robustness evaluation

Ranking robustness is assessed with a weight perturbation test following established practice [20], [26]. Each weight is perturbed from -50% to +50% in steps of 5%, and the remaining weights are rescaled proportionally so the vector sums to one. This procedure is deterministic, with no random sampling or seed dependence. The ±50% range was selected as it represents a substantial and practically interpretable change in criterion importance; perturbations beyond this range would imply that a criterion's weight more than doubles or approaches zero, which falls outside plausible expert disagreement in most decision contexts [27]. When criterion j is perturbed by factor δ, its new weight is computed as:

$$w_j' = max\left(0, w_j\,.\,(1+\delta)\right), \delta \in \{-0.50, -0.45, \ldots, +0.45, +0.50\} \quad (4)$$

The remaining weights are redistributed proportionally to all other criteria $k \neq j$:

$$w_k' = w_k\,.\frac{1-w_j'}{\sum_{k\neq j} w_k},\ \forall k \neq j \quad (5)$$

This ensures that $\sum_{k=1}^{m} w_k' = 1$ is always satisfied. At each perturbation level δ, all five MCDM methods (VIKOR, COPRAS, MOORA, EDAS, and TOPSIS) are re-executed using the perturbed weight vector w′. A trigger is recorded when the complete ranking of the alternative's changes in at least one of the five methods relative to the baseline ranking:

$$Trigger\ (j,\delta) = \begin{cases} 1, if\ \exists s \in S:\ R_s(w') \neq R_s(\mathrm{w}) \\ 0, otherwise \end{cases} \quad (6)$$

Where $S = \{$VIKOR, COPRAS, MOORA, EDAS, TOPSIS$\}$ denotes the set of MCDM methods used, and $R_s(\mathrm{w})$ denotes the complete ranking vector of all alternatives under method s with weight vector w.

The robustness threshold for criterion j is defined as the minimum absolute perturbation, in either direction, that first causes a trigger across any of the five MCDM methods. This operationalizes the notion of weight stability intervals in multi-criteria decision aid [19]:

$$\tau_j = \min(|\,\delta_j^+|, |\delta_j^-|) \quad (7)$$

Where δ_j^+ and δ_j^- are the first positive and negative perturbation levels triggering a ranking change. Criteria with no trigger (NT) within ±50% are classified as NT). The average robustness threshold is computed exclusively over triggered criteria Ω:

$$\bar{\tau} = \frac{1}{|\Omega|} \sum_{j \in \Omega} \tau_j \quad (8)$$

Where | Ω | is the number of triggered criteria. A higher $\bar{\tau}$ reflects greater robustness. The number of triggered criteria | Ω | is reported alongside $\bar{\tau}$.

To characterize how each objective weighting method distributes importance across the criteria, three dispersion measures are computed from each weight vector w=(w_1, …, w_m), where m is the number of criteria. The normalized Shannon entropy (9) equals unity for a perfectly uniform weight vector and decreases as the weights become more concentrated. The coefficient of variation (10) is the ratio of the SD of the weights to their mean. The Gini coefficient (11) equals zero when all weights are equal and approaches unity as the weight becomes concentrated on a single criterion. A lower normalized entropy together with a higher coefficient of variation and Gini coefficient therefore indicates a more concentrated weight vector.

$$H_{norm} = -\frac{1}{ln\, m} \sum_{j=1}^{m} w_j \, ln\, w_j \quad (9)$$

$$CV = \frac{\sigma_w}{\bar{w}}, \ \bar{w} = \frac{1}{m}\sum_{j=1}^{m} w_j, \ \sigma_w = \sqrt{\frac{1}{m}\sum_{j=1}^{m}(w_j - \bar{w})^2} \quad (10)$$

$$G = \frac{\sum_{j=1}^{m}\sum_{k=1}^{m} |w_j - w_k|}{2m \sum_{j=1}^{m} w_j} \quad (11)$$

3. RESULTS AND DISCUSSION

3.1. Criteria weight and concentration

Table 3 reports the criterion weights from the five methods, which assign markedly different importance to the ten criteria. Entropy concentrates weight on operation and maintenance (O&M) cost, efficiency, and job creation, and MEREC shows a similar pattern. CRITIC instead emphasizes capacity factor, land use, and job creation, while PCA and SD spread weight far more evenly across all criteria. Table 4 quantifies this contrast through the concentration measures. Entropy is the most concentrated vector, followed by MEREC, CRITIC, PCA, and SD, with SD almost perfectly uniform and PCA nearly so. This ordering is visualized by the Lorenz curves in Figure 2, where a curve farther from the diagonal denotes greater concentration.

Table 3. Criteria weights from the five objective weighting methods

Criteria	Weighting method				
	Entropy	CRITIC	MEREC	PCA	SD
C1	0.031	0.079	0.067	0.112	0.093
C2	0.222	0.087	0.172	0.103	0.107
C3	0.085	0.090	0.098	0.114	0.099
C4	0.193	0.079	0.165	0.098	0.098
C5	0.111	0.138	0.111	0.110	0.102
C6	0.038	0.090	0.047	0.117	0.107
C7	0.065	0.094	0.068	0.105	0.101
C8	0.086	0.133	0.100	0.086	0.096
C9	0.167	0.116	0.156	0.088	0.102
C10	0.002	0.094	0.016	0.066	0.095

Table 4. Weight concentration profile of the objective weighting methods

Weighting method	H_norm	CV	Gini	w_max	w_min
Entropy	0.885	0.693	0.390	0.222	0.002
MEREC	0.940	0.496	0.281	0.172	0.016
CRITIC	0.991	0.203	0.107	0.138	0.079
PCA	0.995	0.151	0.081	0.117	0.066
SD	1.000	0.045	0.025	0.107	0.093

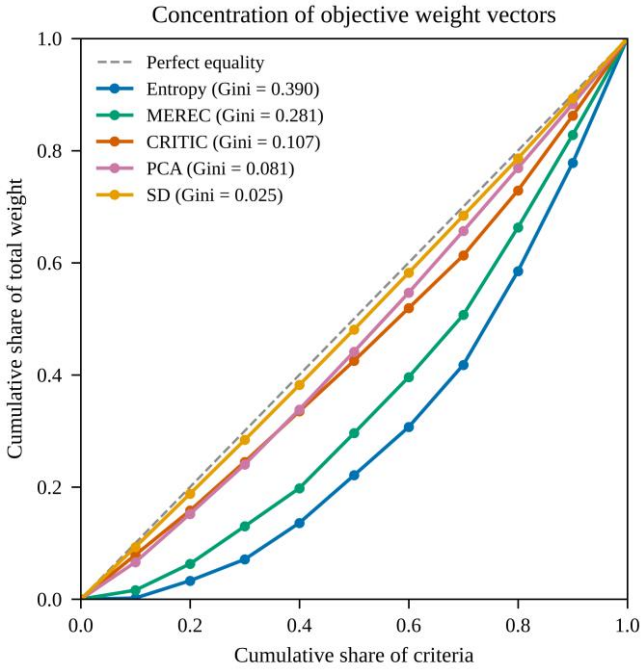


Figure 2. Lorenz curves of the five objective weight vectors

3.2. Alternative ranking across models

Table 5 shows the rank of each alternative across all weighting method and aggregation model combinations. Hydro is the dominant alternative, holding first place in twenty-three of the twenty-five combinations, so the leading recommendation is largely invariant to the weighting method. Variation appears in the intermediate ranks, mainly in the alternation of solar PV and wind between second and third, and in the lower ranks between biomass and geothermal. CRITIC is the clear outlier: it is the only method that displaces hydro from first place, moving wind to the top under VIKOR and Geothermal to the top under TOPSIS. Entropy and MEREC yield near-identical orderings, and PCA and SD often agree. This pattern shows that stability is not uniform across ranking positions, so a stable leading recommendation can coexist with substantial uncertainty among the remaining alternatives, and stability should be read position by position rather than from the top rank alone.

Table 5. Alternative rankings by objective weighting method and MCDM aggregation model

Alternative	VIKOR					COPRAS					MOORA					EDAS					TOPSIS				
	E	C	M	P	S	E	C	M	P	S	E	C	M	P	S	E	C	M	P	S	E	C	M	P	S
Solar PV	2	5	2	4	4	2	3	2	3	3	2	3	2	3	3	2	3	2	4	3	2	2	2	3	2
Wind	3	1	3	2	2	3	2	3	2	2	3	2	3	2	2	3	2	3	2	2	3	4	3	2	3
Hydro	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3	1	1	1
Biomass	4	4	4	3	3	4	5	4	4	4	4	5	4	4	4	4	5	4	3	4	5	5	5	4	5
Geothermal	5	3	5	5	5	5	4	5	5	5	5	4	5	5	5	5	4	5	5	5	4	1	4	5	4

Abbreviations: E=entropy, C=CRITIC, M=MEREC, P=PCA, and S=SD.

3.3. Cross-model ranking stability

Table 6 summarizes cross-model stability across the five aggregation models. Entropy and MEREC are the most stable, each with a mean Spearman correlation of 0.960, a minimum of 0.900, a single-position maximum shift, and only 0.40 rank reversals on average, so their rankings are almost invariant to the choice of model. PCA is comparably stable at 0.940 and SD moderately stable at 0.860. CRITIC is by a wide margin the least stable, with a mean correlation of 0.490, a negative minimum of -0.200, a three-position maximum shift, and three reversals on average. CRITIC is the only method whose rankings change qualitatively across models; the mechanism is examined in subsection 3.6.

Table 6. Cross-model ranking stability by objective weighting method

Weighting method	Mean ρ	Min ρ	Max rank shift	Avg rank reversal
Entropy	0.960	0.900	1	0.40
MEREC	0.960	0.900	1	0.40
PCA	0.940	0.900	1	0.60
SD	0.860	0.500	2	1.20
CRITIC	0.490	−0.200	3	3.00

3.4. Robustness analysis

Robustness was evaluated with the weight perturbation test defined in (4)-(8); the criterion-level thresholds are shown in Figure 3 and the method-level summary in Table 7. No threshold (NT) denotes no ranking change within the tested ±50% interval. The results reveal two distinct groups. MEREC and entropy are markedly more robust, each leaving three criteria with NT within ±50% and yielding average thresholds of 20.0% and 17.9% respectively; their NT criteria are the three with the smallest baseline weights: investment cost (C1), technical maturity (C6), and social acceptance (C10). SD, CRITIC, and PCA form a more sensitive group in which every criterion alters the complete ranking within ±50%, with average thresholds of 9.5%, 8.0%, and 8.0%.

At the criterion level, job creation (C9) is the single most sensitive criterion, altering the complete ranking at 5% under all five weighting methods. Efficiency (C4) is similarly sensitive, triggering at 5% under entropy, MEREC, and CRITIC. For SD, CRITIC, and PCA the near-uniform weight allocation leaves most criteria able to overturn the ranking under a 5% change. A secondary distinction appears within each group: MEREC's average threshold (20.0%) modestly exceeds entropy's (17.9%) despite both triggering the same seven criteria, and SD's threshold (9.5%) is likewise somewhat higher than CRITIC's and PCA's (8.0% each), so full triggering does not make the three methods equally fragile. The robustness grouping does not coincide with the stability ordering. PCA is the clearest case: among the most stable methods across aggregation models, yet in the most sensitive robustness group.

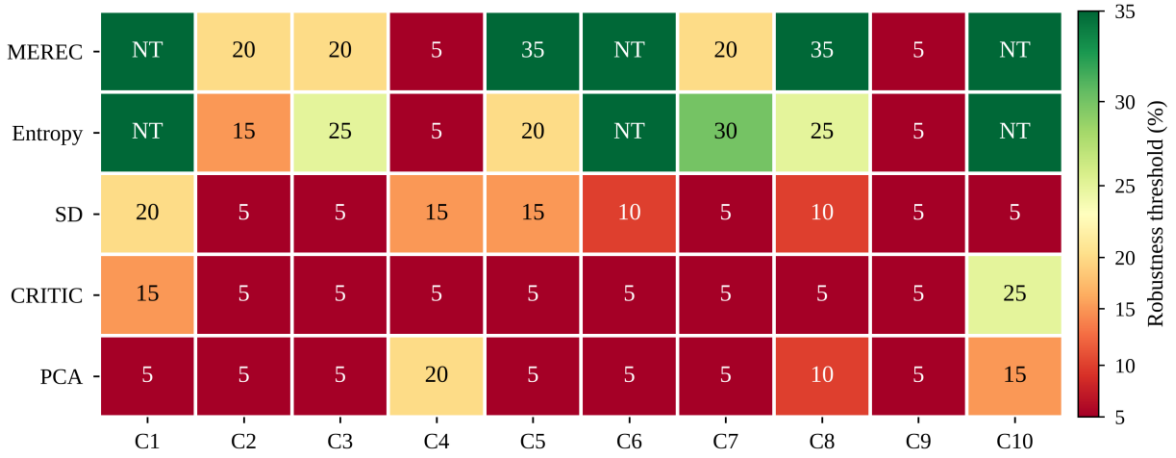


Figure 3. Robustness threshold per criterion and weighting method

Table 7. Robustness performance by objective weighting method

Weighting method	Triggered criteria	NT criteria	Avg threshold (%)
MEREC	7	3	20.0
Entropy	7	3	17.9
SD	10	0	9.5
CRITIC	10	0	8.0
PCA	10	0	8.0

3.5. External validation

To test whether these patterns are specific to the renewable energy problem, the procedure was repeated on an independent supplier-selection dataset from Subramanian *et al.* [28], evaluating ten manufacturing subcontractors against eleven benefit criteria, compared in Table 8.

Table 8. Method comparison across the primary and external datasets

Weighting method	Primary dataset (5×10)			External dataset (10×11)		
	Stab. ρ	Robust (%)	NT	Stab. ρ	Robust (%)	NT
Entropy	0.960	17.9	3	0.928	12.5	3
MEREC	0.960	20.0	3	0.918	7.8	2
PCA	0.940	8.0	0	0.880	5.0	0
SD	0.860	9.5	0	0.838	7.7	0
CRITIC	0.490	8.0	0	0.779	5.9	0

The two datasets yield consistent results. The stability ordering is preserved, and the robustness ordering is preserved with a single adjacent exchange at the top. The same structural pattern reappears: methods that resist perturbation across all criteria in the primary dataset do so externally as well, while a method stable across aggregation models can still sit close to the boundary between competing orderings under weight perturbation. The external dataset's criterion count is close to that of the primary dataset, so this validation speaks to generalizability across problem domains rather than to how the framework behaves as the number of criteria grows substantially larger, which we leave to future work. Within that scope, the results confirm the central observation that cross-model stability and perturbation robustness are distinct properties, not an artifact of the original problem.

3.6. Findings discussion

The cross-model design addresses three issues a single aggregation model cannot: ranking stability across MCDM models, the stability-robustness relationship, and the weight-vector properties governing robustness. MCDM instability is usually framed as rank reversal from adding, removing, or reweighting an alternative [29]; the present results show that even with fixed alternatives and weights, rankings still vary across aggregation models depending on the weighting method used, so stability depends on both the aggregation model and the weighting method applied beforehand.

However, cross-model stability does not guarantee robustness to weight changes. Sensitivity analysis, formalized by Triantaphyllou and Sánchez [20] as the smallest weight change needed to reverse a ranking, is extended here to the complete ranking: robustness is positively associated with stability but does not follow from it. PCA illustrates this, stable across models yet highly sensitive to perturbation. Stability and robustness are therefore distinct properties that should be assessed jointly, as visualized in Figure 4.

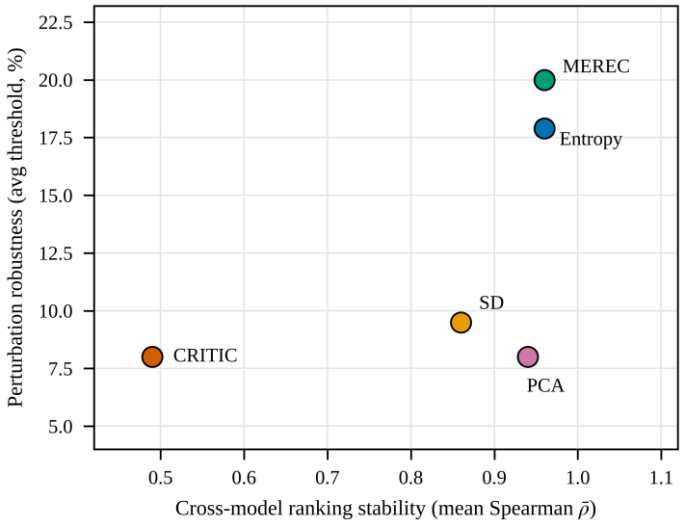


Figure 4. Cross-model stability versus perturbation robustness for each weighting method

What governs robustness is the weight vector's structure, not its evenness. A near-uniform vector might seem a neutral, safe basis for ranking, but the results show otherwise: PCA, among the most evenly distributed (Gini=0.081), is in the most sensitive group, while the more concentrated entropy and MEREC are the most robust. Concentration correlates with robustness (Spearman ρ≈0.56) but does not determine it, since SD (Gini=0.025), the most uniform method, is more robust than CRITIC and PCA. What reconciles this is the score margin between adjacent alternatives: a vector is robust when it separates alternatives widely,

sensitive when nearly tied. Concentrated weighting widens these margins and uniform weighting compresses them, but the outcome depends on where the weight falls, why SD escapes the general tendency and CRITIC stays sensitive despite moderate concentration. Criterion correlation also shapes these results: CRITIC and PCA are correlation-aware by construction, discounting or absorbing redundant criteria, while entropy and SD evaluate each criterion in isolation and can let correlated criteria dominate the weighting. This is why a stable-looking ranking can mislead, driven by one factor counted several times, and shift once the redundancy is recognized; reporting the correlation structure alongside the weights is a useful safeguard.

Together, these findings show that objective weighting methods should be evaluated by the stability and robustness of the rankings they induce, not by their weights alone, since the two diagnostics are distinct and can diverge. For practitioners, they form a two-stage check: first apply the weight vector across multiple aggregation models, where high agreement rules out an aggregation artifact; then run the perturbation test to flag criteria sensitive enough to warrant scrutiny before adoption, a bar that entropy and MEREC clear in this study. For a decision-maker choosing energy technologies or suppliers, such a ranking is a safer recommendation basis than one that is merely stable, such as PCA, which can still flip its ranking under a small weight change despite looking consistent across aggregation models. These diagnostics measure methodological robustness, not engineering superiority: a consistently top-ranked alternative is the most defensible choice given the criteria and data at hand, but only the technically best option if those criteria adequately capture the engineering objectives. More broadly, energy technology choices feed sustainability agendas and supplier choices feed supply chain resilience agendas, both increasingly shaped by rankings that inform public policy, so a ranking's dependence on methodology rather than substance becomes a matter of public accountability, not just academic interest.

4. CONCLUSION

This study evaluated five objective weighting methods, entropy, CRITIC, MEREC, PCA, and SD, across five MCDM aggregation models (VIKOR, COPRAS, MOORA, EDAS, and TOPSIS) on a renewable energy selection problem. Cross-model stability is attainable, with entropy and MEREC most stable and CRITIC least stable, but stability does not imply robustness: PCA is among the most stable methods yet overturns the complete ranking under a weight change as small as 5%. What governs robustness is not weight evenness or concentration but the score margin the weights open between adjacent alternatives.

The key contribution is treating cross-model stability and perturbation robustness as two distinct but jointly necessary diagnostics. Both should be reported: a method stable across models may still be sensitive to weight changes, and a recommendation is most defensible when corroborated across aggregation models. These findings were confirmed on an independent supplier-selection dataset (ten alternatives and eleven criteria), where the stability and robustness orderings were preserved and the divergence between the two diagnostics reappeared. Across both datasets, entropy and MEREC were the only methods consistently both stable and robust, making them the more dependable default when a new application cannot run the full diagnostic procedure. The framework is applicable to any domain where objective weighting methods are used in data-driven selection problems, though limitations remain. The primary dataset has few alternatives and relies on secondary data adapted from a published source, and rankings may be more sensitive to individual criteria than with a larger set, where each criterion's weight share is more diluted; the perturbation test also varies one criterion at a time rather than jointly. Further study should extend the framework to larger criteria sets, perturb several of them simultaneously, add further aggregation models, integrate hybrid weighting approaches, and examine robustness under correlated criteria.

FUNDING INFORMATION

This study was conducted without external funding and received no specific grant from any public, commercial, or not-for-profit funding agency.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Itishom Al Khoiry	✓	✓	✓	✓		✓	✓	✓	✓		✓	✓	✓	✓
Dhea Rizky Amelia	✓	✓		✓	✓		✓	✓		✓			✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY

The data that support the findings of this study are publicly available in the article by Lee and Chang [23], and the supplier-selection dataset used for external validation in the article by Subramanian *et al*. [28]. Code is available from the corresponding author upon reasonable request.

REFERENCES

[1] H. Taherdoost and M. Madanchian, "Multi-Criteria Decision Making (MCDM) Methods and Concepts," *Encyclopedia*, vol. 3, no. 1, pp. 77–87, Jan. 2023, doi: 10.3390/encyclopedia3010006.

[2] S. Chakraborty, R. D. Raut, T. M. Rofin, and S. Chakraborty, "A comprehensive and systematic review of multi-criteria decision-making methods and applications in healthcare," *Healthc. Anal.*, vol. 4, p. 100232, Dec. 2023, doi: 10.1016/j.health.2023.100232.

[3] I. Z. Mukhametzyanov, "Specific character of objective methods for determining weights of criteria in MCDM problems: Entropy, CRITIC, SD," *Decis. Mak. Appl. Manag. Eng.*, vol. 4, no. 2, pp. 76–105, Oct. 2021, doi: 10.31181/DMAME210402076I.

[4] D. Diakoulaki, G. Mavrotas, and L. Papayannakis, "Determining objective weights in multiple criteria problems: The critic method," *Comput. Oper. Res.*, vol. 22, no. 7, pp. 763–770, Aug. 1995, doi: 10.1016/0305-0548(94)00059-H.

[5] M. Keshavarz-Ghorabaee, M. Amiri, E. K. Zavadskas, Z. Turskis, and J. Antucheviciene, "Determination of objective weights using a new method based on the removal effects of criteria (MEREC)," *Symmetry*, vol. 13, no. 4, p. 525, Mar. 2021, doi: 10.3390/sym13040525.

[6] N. T. D. Linh, N. H. Son, and D. X. Thao, "Evaluating the Impact of Weighting Methods on the Stability of Scores for Alternatives in Multi-Criteria Decision-Making Problems," *Eng. Technol. Appl. Sci. Res.*, vol. 15, no. 1, pp. 19998–20004, Feb. 2025, doi: 10.48084/etasr.9518.

[7] A. Anagnostopoulou, D. Sotiropoulos, and K. Tserpes, "A Robust Sustainability Assessment Methodology for Aircraft Parts: Application to a Fuselage Panel," *Sustain.*, vol. 17, no. 8, 2025, doi: 10.3390/su17083299.

[8] A. Aytekin, "Comparative analysis of normalization techniques in the context of MCDM problems," *Decis. Mak. Appl. Manag. Eng.*, vol. 4, no. 2, pp. 1–25, Mar. 2021, doi: 10.31181/dmame210402001a.

[9] D. D. Trung, H. T. Dung, N. C. Bao, and D. V. Duc, "A novel approach for determining criteria weights: application in ranking materials for mechanical manufacturing processes," *Manuf. Rev.*, vol. 12, p. 16, Jul. 2025, doi: 10.1051/mfreview/2025014.

[10] M. Danielson, "Widely Used Multi-Criteria Decision Analysis Methods—A Comprehensive Ranking," *Appl. Sci.*, vol. 16, no. 14, p. 7269, Jul. 2026, doi: 10.3390/app16147269.

[11] G. Demir, P. Chatterjee, and D. Pamucar, "Sensitivity analysis in multi-criteria decision making: A state-of-the-art research perspective using bibliometric analysis," *Expert Syst. Appl.*, vol. 237, p. 121660, Mar. 2024, doi: 10.1016/j.eswa.2023.121660.

[12] B. Paradowski, A. Shekhovtsov, A. Bączkiewicz, B. Kizielewicz, and W. Sałabun, "Similarity analysis of methods for objective determination of weights in multi-criteria decision support systems," *Symmetry*, vol. 13, no. 10, p. 1874, Oct. 2021, doi: 10.3390/sym13101874.

[13] S. Opricovic and G. H. Tzeng, "Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS," *Eur. J. Oper. Res.*, vol. 156, no. 2, pp. 445–455, Jul. 2004, doi: 10.1016/S0377-2217(03)00020-1.

[14] E. K. Zavadskas and A. Kaklauskas, "Multiple criteria evaluation of buildings," *Vilnius, Lith.*, 1996.

[15] W. K. M. Brauers and E. K. Zavadskas, "The MOORA method and its application to privatization in a transition economy," *Control Cybern.*, vol. 35, no. 2, pp. 445–469, 2006.

[16] M. K. Ghorabaee, E. K. Zavadskas, L. Olfat, and Z. Turskis, "Multi-Criteria Inventory Classification Using a New Method of Evaluation Based on Distance from Average Solution (EDAS)," *Informatica*, vol. 26, no. 3, pp. 435–451, Jan. 2015, doi: 10.15388/Informatica.2015.57.

[17] C.-L. Hwang and K. Yoon, "Methods for Multiple Attribute Decision Making," *Multiple Attribute Decision Making*, pp. 58–191, 1981, doi: 10.1007/978-3-642-48318-9_3.

[18] W. Sałabun and K. Urbaniak, "A new coefficient of rankings similarity in decision-making problems," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12138 LNCS, 2020, pp. 632–645, doi: 10.1007/978-3-030-50417-5_47.

[19] B. Mareschal, "Weight stability intervals in multicriteria decision aid," *Eur. J. Oper. Res.*, vol. 33, no. 1, pp. 54–64, Jan. 1988, doi: 10.1016/0377-2217(88)90254-8.

[20] E. Triantaphyllou and A. Sánchez, "A sensitivity analysis approach for some deterministic multi-criteria decision-making methods," *Decis. Sci.*, vol. 28, no. 1, pp. 151–194, Jan. 1997, doi: 10.1111/j.1540-5915.1997.tb01306.x.

[21] Z. Li, Z. Luo, Y. Wang, G. Fan, and J. Zhang, "Suitability evaluation system for the shallow geothermal energy implementation in region by Entropy Weight Method and TOPSIS method," *Renew. Energy*, vol. 184, pp. 564–576, Jan. 2022, doi: 10.1016/j.renene.2021.11.112.

[22] S. S. Goswami, S. K. Mohanty, and D. K. Behera, "Selection of a green renewable energy source in India with the help of MEREC integrated PIV MCDM tool," *Mater. Today Proc.*, vol. 52, pp. 1153–1160, 2022, doi: 10.1016/j.matpr.2021.11.019.

[23] H. C. Lee and C. T. Chang, "Comparative analysis of MCDM methods for ranking renewable energy sources in Taiwan," *Renew.*

Sustain. Energy Rev., vol. 92, pp. 883–896, Sep. 2018, doi: 10.1016/j.rser.2018.05.007.
[24] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948, doi: 10.1002/j.1538-7305.1948.tb01338.x.
[25] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *J. Educ. Psychol.*, vol. 24, no. 6, pp. 417–441, 1933, doi: 10.1037/h0071325.
[26] J. Więckowski and W. Sałabun, "Sensitivity analysis approaches in multi-criteria decision analysis: A systematic review," *Appl. Soft Comput.*, vol. 148, p. 110915, Nov. 2023, doi: 10.1016/j.asoc.2023.110915.
[27] J. Si and L. Marjanovic-Halburd, "Criteria weighting for green technology selection as part of retrofit decision making process for existing non-domestic buildings," *Sustain. Cities Soc.*, vol. 41, pp. 625–638, 2018, doi: 10.1016/j.scs.2018.05.051.
[28] S. P. Subramanian, P. Pandian, R. Sivaprakasam, and J. Kadarkarai, "A Novel Machine Learning Framework for Optimized Supplier Selection Using the Weights by ENvelope and SLOpe (WENSLO) Technique," *Appl. Oper. Anal.*, vol. 1, no. 1, pp. 1–10, Dec. 2025, doi: 10.1080/29966892.2025.2474824.
[29] R. F. de F. Aires and L. Ferreira, "The rank reversal problem in multi-criteria decision making: A literature review," *Pesqui. Oper.*, vol. 38, no. 2, pp. 331–362, Aug. 2018, doi: 10.1590/0101-7438.2018.038.02.0331.

BIOGRAPHIES OF AUTHORS



Itishom Al Khoiry is an assistant professor in the Department of Digital Business, Universitas Persatuan Guru Republik Indonesia Semarang, Indonesia. He received a master's degree in Information Systems from Universitas Diponegoro, Indonesia. He also serves as the head of laboratory and the head of the Study Program Quality Assurance Sub-Unit in the Digital Business Study Program. His research interests include artificial intelligence, machine learning, data-driven decision analysis, information systems, and digital business analytics. He can be contacted at email: itishom@upgris.ac.id.



Dhea Rizky Amelia is an assistant professor in the Department of Development Economics, Universitas Negeri Semarang, Indonesia. She received the bachelor's degree in Economics and the master's degree in Economic Planning and Development Policy from Universitas Indonesia. Her research interests include labor economics, public policy, data-driven decision analysis, and applied economic analysis. She can be contacted at email: dhearizkyamelia@mail.unnes.ac.id.