



Methodological Advances in Sensitivity Analysis and Validation for Multi-Criteria Decision-Making (MCDM) Methods: A Comprehensive Review with Advancements, Applications, and Future Directions

Sushil Kumar Sahoo^{1,2*}, Dwarikanath Choudhury³, Prasant Ranjan Dhal^{1,2}, Supriya Sahu^{1,2}, Sudhakar Majhi^{1,2}, Ipsita Dhar^{1,2}

¹ Department of Mechanical Engineering, Indira Gandhi Institute of Technology, Sarang 759146, Odisha, India

² Biju Patnaik University of Technology Rourkela 769015 Odisha, India

³ Great Lakes Institute of Management, Gurgaon Campus, Delhi NCR, 122413, Haryana, India

ARTICLE INFO

Article history:

Received 31 May 2026

Received in revised form 3 August 2026

Accepted 9 September 2026

Available online 25 September 2026

Keywords:

Multi-Criteria Decision-Making; MCDM; MCDA; Sensitivity Analysis; Robustness Analysis; Ranking Stability; Validation; Bibliometric Analysis

ABSTRACT

This study provides a comprehensive review of methodological advances in sensitivity analysis and validation of Multi-Criteria Decision-Making (MCDM) methods. A Systematic Literature Review (SLR) following the PRISMA 2020 guidelines was conducted using the Scopus database, covering the period from 2000 to 2026. A total of 743 records were identified using the defined Scopus search string. Publication trends, citation patterns, leading authors, journals, institutions, countries, and keyword structures were examined through bibliometric analysis, while sensitivity analysis and validation approaches were systematically classified and examined using thematic analysis. The findings reveal a broad range of sensitivity analysis approaches, including deterministic, probabilistic, simulation-based, fuzzy, interval-based, and hybrid techniques. Validation approaches were classified into internal, comparative, statistical, expert-based, and external validation techniques. The bibliometric results show that India was the most productive country, with 167 published papers, while Pamučar recorded the highest number of citations among the analysed authors, with 1,004 citations. The review also highlights the increasing methodological diversity of sensitivity and validation procedures used to evaluate the robustness and reliability of MCDM results. Based on the identified methodological patterns, the study proposes an integrated framework linking uncertainty, sensitivity analysis, robustness assessment, ranking stability, and validation. The proposed framework provides a structured basis for evaluating the reliability of MCDM results and contributes to improving the credibility, transparency, reproducibility, and methodological rigor of MCDM-based decision-making.

1. Introduction

Multi-Criteria Decision-Making (MCDM) or Multi-Criteria Decision Analysis (MCDA) is one of the most popular methods for tackling decision-making problems with multiple and often competing criteria. It offers a structured approach to the assessment of options, considers the preferences of

* Corresponding author.

E-mail address: author.mail@gmail.com

<https://doi.org/10.67334/cds202645>

© The Author(s) 2026 | [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

those involved in decision making and assigns weights to the criteria used, and identifies appropriate solutions. There are many decision approaches, such as AHP, TOPSIS, VIKOR, ELECTRE and PROMETHEE, developed in various contexts of decision making [1, 2]. Further extensions of the field have been made by objective and subjective weighting, fuzzy sets, interval numbers, stochastic models, and hybrid methods. As a result, MCDM is widely used in engineering, manufacturing, energy, healthcare, transportation, supply chain management, environmental assessment, agriculture, and decision-making problems in the area of sustainability.

1.1 Reliability Challenges in MCDM

Although they are widely used, the outcomes from MCDM methods are influenced by the assumptions made, the data used and the procedures followed throughout the decision-making process. An MCDM model includes a set of alternatives, criteria, values of the decision matrix, weights of criteria, normalization methods, aggregation methods, and parameters for the method itself [3, 4]. Modifications to any of these elements can affect rankings and scores. Criteria weights could have subjective and objective uncertainties and decision data could have measurement errors and lack of information. Likewise, normalization, aggregation rules, parameter values, and alternative selection can impact results. So, the most likely option is not necessarily the most reliable one, unless the impact of uncertainty and methodological assumption is examined.

1.2 Importance of Sensitivity Analysis

Sensitivity analysis is one of the valuable methods to investigate the impacts of variation in the input, assumption and structure of an MCDM model. It can explore the impact of changing the relative weights of the criteria, the values on the decision matrix, the normalization method, the aggregation method, method-specific parameter values, and different sets. Classically, MCDM sensitivity analysis mostly concerned with the variation of the weights of the criteria and the corresponding change in the rankings [5]. It has, however, been extended to data perturbation, scenario analysis, rank reversal, alternative addition or removal, probabilistic analysis and approaches based on simulation like Monte Carlo simulation, bootstrap analysis, global sensitivity analysis and stochastic multicriteria acceptability analysis [6, 7]. These approaches allow for a more comprehensive evaluation and ranking of uncertainty and variability. Highly sensitive means that a small change can make a big difference to the result; stable means that the result has high resistance. But the sensitivity analysis is not sufficient to prove the validity of a decision

1.3 Importance of Validation

Validation is crucial to imparting confidence in the decision results obtained using MCDM. The purpose of validation is to judge the credibility and appropriateness of the model and its results, as opposed to sensitivity analysis, where the effect of changes in inputs and assumptions are assessed. Internal consistency checks, comparison with other sources, statistical analysis, simulation, expert judgment, benchmarking, and comparison to experimental or actual results are the typical methods for validation [8–10]. Spearman's correlation coefficient and Kendall's correlation coefficient are also commonly used for ranking comparison between methods of MCDM [11, 12]. Validation practices, however, are still quite varied and disorganized. Do not assume that if two (or more) methods agree, or that the two rank the same, the decision is correct in practice. Thus, a comprehensive validation should include suitable types of evidence based on the nature of the decision problem.

1.4 Sensitivity, Robustness, Stability, and Validation

The concepts of sensitivity, robustness and stability are closely related but distinct components of the reliability of an MCDM decision. Sensitivity is the study of the effects on the results due to input, assumptions, and/or model components. The robustness of a preferred alternative means how it fares when the conditions or scenarios change slightly or there is uncertainty, while stability refers to whether the rankings stay the same among different scenarios or conditions [13]. Validation involves checking the evidence for the validity and usefulness of the model and results of the model. The two should not be used as synonyms since a stable ranking might be based on a wrong type of data or on wrong assumptions. Hence the process of a comprehensive assessment should be carried out starting from the identification of uncertainty sources, sensitivity analysis, robustness and stability analysis, and appropriate validation steps to build confidence on the final decision made.

1.5 Existing Research Problems

Some challenges with respect to sensitivity analysis and validation in MCDM have been highlighted, but this topic has not been the focus of much research. Often the sensitivity analysis and the validation are analysed separately in existing studies, which does not provide a complete picture of how sensitive validation is to sensitivity analysis and how reliable decisions can be. Furthermore, sensitivity analysis is often limited to variations in criteria weights, with less attention paid to uncertainties arising from the criteria values in the decision matrix, the way the criteria are normalized, the method of aggregation, the parameters used by the method, the different sets of criteria, and the method-specific aggregation procedures. The terms sensitivity, robustness, stability and validation are also used loosely, with conceptual overlap. There are also no standardized approaches to the selection of ranges of perturbation, simulation parameters, or stability parameters. Typically, cross-method validation is used; this involves comparing the methods to establish consistency, but it does not guarantee actual validity. Cross-method validation is also carried out by rank-correlations.

1.6 Contributions of the Present Study

To address these problems, this study has reviewed and summarized the advances in methodologies for sensitivity analysis and validation of MCDM methods in detail. It analyses the research trends by using bibliometric analysis and establishes structured sensitivity and validation taxonomies, such as deterministic, probabilistic, simulation-based and uncertainty-aware. The paper also elucidates the connections between sensitivity, robustness, stability and validation, discusses the important application fields, outlines gaps in the field and suggests a combined framework for enhancing the reliability, transparency and credibility of the decision-making process in the MCDM framework and provides research perspectives.

This paper is organized as follows, after this introduction. The existing literature is reviewed (Section 2) and gaps in knowledge identified. The methodology of the systematic review is detailed in Section 3. The bibliometric analysis and the research trends are presented in Section 4. Sections 5 and 6 classify sensitivity analysis and validation methods, respectively. Recent methodological advances are described in Section 7 and major application domains in Section 8. A critical discussion is included in Section 9. Last but not the least, Sections 10 and 11 feature future research directions and conclusion, respectively.

2. Literature Review

The classic sensitivity analysis in MCDM has been extended from the basic procedures of determinant criteria-weight variation to a more comprehensive set of analyses that address the

uncertainty in the decision matrix, various changes, scenarios, probabilistic inputs, simulation, fuzzy representations, and methodological assumptions. These methods have been systematically categorized in existing reviews, and this weight-based sensitivity analysis continues to be the primary way to rank risks.

2.1 Evolution of Sensitivity Analysis in MCDM

The sensitivity analysis in MCDM has been developed from the first few basic deterministic approaches (mostly changes in the weights of the criteria) to the more comprehensive approaches that deal with several uncertainty sources as presented in Table 1. Initial work was on ranking changes due to weight changes, later work on perturbations of the decision matrix, alternative changes, normalization, parameters and scenarios.

Table 1

Representative studies on the evolution of sensitivity analysis in MCDM

Author(s)	Description	Key Findings	Limitation
Saaty [14]	Introduced consistency and sensitivity considerations within the AHP framework.	Demonstrated the importance of examining changes in priorities and judgments.	Primarily associated with hierarchical and pairwise-comparison structures.
Triantaphyllou and Sánchez [15]	Examined sensitivity and ranking behaviour in MCDM methods.	Showed that changes in inputs and methodological procedures can affect decision outcomes.	Focused mainly on selected MCDM approaches.
Ishizaka and Nemery [16]	Discussed sensitivity and robustness as important components of MCDA practice.	Emphasized the role of uncertainty and preference variation in decision analysis.	Broad MCDA coverage rather than a dedicated taxonomy of sensitivity methods.
Więckowski and Sałabun [17]	Conducted a systematic review of sensitivity-analysis approaches in MCDA.	Identified different categories of sensitivity techniques and highlighted the dominance of weight-based analysis.	Validation approaches were not the central focus of the review.
Demir <i>et al.</i> , [18]	Examined sensitivity analysis in MCDM from a state-of-the-art bibliometric perspective.	Revealed the growth, intellectual structure, and research trends of sensitivity-analysis studies.	Focused primarily on bibliometric patterns rather than integrated methodological validation.

2.2 Existing Reviews on Sensitivity Analysis

The available literature on SA is summarized in this section. This section summarizes the existing literature on SA. This means that there are several important reviews available that have brought together different sensitivity analysis methods and highlighted important trends in MCDM research. General reviews on MCDM methods, systematic reviews on sensitivity-analysis techniques and bibliometric studies on the intellectual growth of the field are all found in the literature.

The studies discussed in Table 2 illustrate growing interest in the topics of ranking robustness, uncertainty analysis, and methodological comparison. However, most of the available reviews are only concerned with the sensitivity analysis and the systematic relationship between sensitivity analysis and the validation is not examined as thoroughly. Moreover, recent methodological advancements like the probabilistic ranking, the resampling, the hybrid uncertainty analysis and integrated validation workflows also need to be synthesized in a single framework.

Table 2

Existing studies related to sensitivity analysis in MCDM

Author(s)	Description	Key Findings	Limitation
Sahoo and Goswami [19]	Comprehensive review of MCDM methods, developments, and applications.	Provided a broad overview of major MCDM methods and application areas.	Sensitivity and validation were not the primary focus.
Markatos <i>et al.</i> , [20]	Assessed sustainability rankings using sensitivity analysis aviation case.	Model produced stable, robust rankings under varying conditions.	Limited validation across industries, materials, and applications globally.
Wolters and Mareschal [21]	Introduced three approaches examining rankings, scores, and weights.	Methods support dynamic decisions and analyse weight spaces.	Demonstrated primarily through simulation, limiting practical validation.
Yazdani <i>et al.</i> , [22]	Applied SWARA, WASPAS, MOORA with normalization sensitivity analysis.	New normalization validated robust material ranking across cases.	Only two case studies limited broader generalizability validation.
Boukrouh <i>et al.</i> , [23]	Compared subjective, objective, and hybrid weighting methods using TOPSIS in supplier selection framework.	Hybrid methods showed greatest stability across nineteen sensitivity analysis scenarios.	Framework requires broader validation across industries and real-world applications.

2.3 Existing Studies on MCDM Validation

In the field of MCDM, different methods have been suggested to validate the results, such as by testing for internal consistency, robustness test, cross method comparison, statistical correlation test, expert opinion test, benchmarking, simulation, and comparison with external or empirical evidence. Methods based on pairwise judgements are especially relying on internal validation, whereas comparative and statistical methods are frequently employed to investigate agreement among rankings as presented in Table 3.

Table 3

Existing approaches and studies related to MCDM validation

Author(s)	Description	Key Findings	Limitation
Ayyildiz [24]	Validation used consistency ratios and consolidated expert consensus judgments.	Consistency ratio confirmed reliable and internally consistent MCDM results. RANCOM generally	Lacked sensitivity analysis and comparative validation using alternative methods.
Więckowski <i>et al.</i> , [25]	Systematically compared RANCOM and AHP under expert judgment inaccuracies.	demonstrated greater robustness, maintaining rankings despite increasing expert judgment inaccuracies.	Validation relied on simulations and one housing location decision case, limiting generalizability.
Brunelli <i>et al.</i> , [26]	Consistency validation in pairwise comparisons. Compared ten inconsistency indices to assess agreement among validation measures numerically.	Results revealed consistency indices exhibited highly variable degrees of mutual agreement.	Analysis focused numerical agreement, without examining practical decision-making performance extensively.
Ramírez-Olivares <i>et al.</i> , [27]	Validation combined expert review, reliability testing, consistency analysis, and sensitivity scenarios.	Model demonstrated reliable judgments and robust rankings across changing decision conditions.	Expert-based regional case limits generalizability and retains subjective judgment uncertainty considerably.
Memis & Yildiz [28]	Validation employed Delphi consensus, AHP consistency, and uncertainty-aware PEM evaluation procedures.	Integrated framework systematically prioritized strategies despite uncertainty and subjective expert judgments.	Lacked explicit sensitivity, comparative, and external validation of resulting rankings.

2.4 Summary of Research Gaps

Although more studies have been conducted, there are some limitations and gaps in sensitivity analysis and validation of MCDM methods. Current studies tend to focus on sensitivity, robustness, and validation in isolation, and few studies have paid attention to the uncertainty sources of criteria-weight variation.

2.5 Novelty of the Present Study

The novelty of the study is in combining the methods of systematic literature review, bibliometric analysis, methodological content analysis and conceptual synthesis to study sensitivity analysis and validation in MCDM. It establishes comprehensive classifications, clarifies the relationships between the concepts of sensitivity, robustness, stability and validation, investigates advanced approaches, identifies methodological limitations and suggests an integrated decision-credibility framework.

2.6 Research Questions

The following research questions are proposed based on the identified research gaps.

RQ1: What are the most important sources of uncertainty and sensitivity analysis methods used in MCDM?

RQ2: How to determine the credibility and reliability of the results obtained from MCDM?

RQ3: How to differentiate conceptually and systematically sensitivity from robustness, ranking stability and validation?

RQ4: What are the key methodological pitfalls and research gaps for the use of sensitivity-analysis and validation in the current practice?

2.7 Objectives of the Study

The purpose of the study is to identify and assess the objectives.

- a) To conduct a systematic review of the development of sensitivity analysis and validation practices in MCDM studies.
- b) To perform a bibliometric analysis of the literature selected, to look for trends in publication, influential authors, countries, institutions, sources, collaboration networks and thematic developments.
- c) The taxonomy of sensitivity-analysis methods in MCDM should be developed, which will be useful for identifying and understanding them.
- d) To review the key areas where the use of sensitivity analysis and validation can be found.
- e) To highlight the key research gaps and methodological challenges of existing practices.

3. Methodology

This study aims to review the methodologies development, application, and future directions in conducting sensitivity analysis and validation in MCDM. The study uses Systematic Literature Review (SLR) methodology that includes both bibliometric and thematic analysis. The study follows the Systematic Literature Review (SLR) methodology along with bibliometric and thematic analysis to gain a comprehensive understanding of the evolution and current status of this research field. The methodology for the review is in line with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure transparency, repeatability and methodological rigor in the process of literature selection [29, 30].

Only literature data from Scopus database were collected because it covers a wide range of peer-reviewed scientific journals and multi-disciplinary research papers in the field of MCDM, decision sciences, engineering, management, or related areas. The search spanned the twenty-seven-year timeframe from 2000 to 2026, allowing the study to examine the evolution of sensitivity analysis and

validation from the classical deterministic methods to the recent probabilistic, simulation-based and uncertainty-aware methods.

The literature search was conducted in the Scopus database, and the data were collected on 21 August 2026. The following search string was applied: ("MCDM" OR "MCDA" OR "multi-criteria decision making" OR "multi criteria decision making" OR "multi-criteria decision analysis" OR "multi criteria decision analysis" OR "MADM" OR "multi-attribute decision making" OR "multiple criteria decision making") AND ("sensitivity analysis" OR "sensitivity assessment" OR "sensitivity testing" OR "robustness analysis" OR "robustness assessment" OR "stability analysis" OR "ranking stability" OR "rank stability" OR "rank reversal" OR "uncertainty analysis" OR "perturbation analysis" OR "scenario analysis" OR "Monte Carlo simulation" OR "probabilistic sensitivity analysis" OR "global sensitivity analysis" OR "local sensitivity analysis" OR "SMAA") AND ("validation" OR "model validation" OR "method validation" OR "ranking validation" OR "result validation" OR "comparative validation" OR "statistical validation" OR "external validation" OR "empirical validation" OR "expert validation" OR "benchmarking" OR "cross validation" OR "rank correlation" OR "Spearman correlation" OR "Kendall correlation")

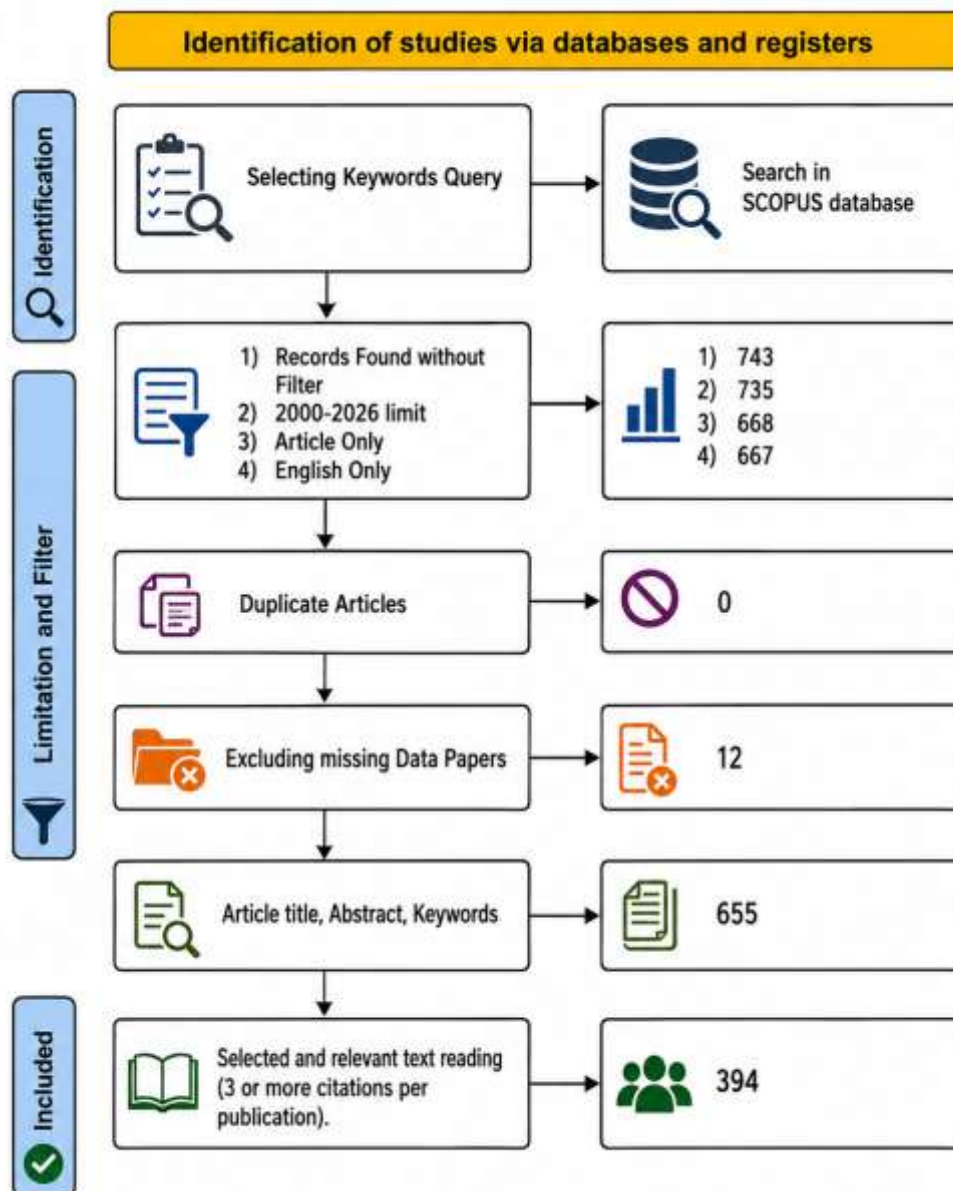


Fig. 1. PRISMA 2020 flow diagram

The initial Scopus search string identified 743 publications on 21 August 2026. The records were not used directly in the final thematic review, but rather a series of "filters" and screening criteria were used to define the literature in scope of the review. The first step was to limit the publication range to 2000-2026, which yielded 735 publications. The dataset was then restricted to journal articles and no other document types which further decreased the number of records to 668. Later, the language was limited to English and 667 publications were published. There were no records found during screening that were duplicated. In total, 12 publications were excluded from further analysis due to missing or incomplete bibliographic data during data preparation, leaving 655 publications for data analysis. The titles, abstracts and key words of these publications were then screened for relevance with respect to the goals and scope of the current project leading to 655 relevant publications. The accessed dataset was then chosen for a bibliometric analysis of its content. To limit the possible age bias of a high citation threshold, but still select publications that have a high degree of establishment and relevance for the subsequent thematic analysis, a threshold of 3 or more citations per publication was used. Based on this criterion, out of 394 publications, selected and relevant publication is used for detailed thematic analysis. The process of selection of the entire study is described in the PRISMA 2020 flow diagram (Figure 1).

4. Bibliometric Analysis

Bibliometric analysis is a quantitative method to analyse intellectual structure and development of a research field [31, 32]. It can examine the pattern of publications, citations, influential authors, journals, countries, collaboration networks, and key words to highlight major research themes, emergent trends, and future directions in a particular scientific field.

4.1 Annual Publication, Citation Trends

Annual publication trend is the yearly pattern and its growth, which indicates the scholarly interest and the research activity in a given field over time [33, 34]. As can be seen from the trend of annual publications, the amount of research on sensitivity analysis and validation in MCDM has shown a significant growth. There was little publication activity for 2000–2015, with less than five publications being published each year. A gradual increase began after 2016, rising from 9 publications in 2016 to 32 in 2020 and 33 in 2021 as shown in Figure 2. The growth in research publications grew significantly after 2022, with 51 publications in 2022, 72 in 2023 and 73 in 2024. The largest growth rates were for 2025 (133 publications) and 2026 (185 publications). The findings suggest that the academic interest in enhancing the reliability, robustness, sensitivity analysis and validation of MCDM-based decision-making methods is on the rise.

Citation trend refers to the annual citation history of a publication, which reflects the scientific impact, influence, visibility and changing importance of a publication in a research area over time [35, 36]. The citation trend demonstrates the increasing scientific impact of studies about sensitivity analysis and validation on MCDM. The level of citations was generally low and variable from 2000 to 2016, as there were few publications in the early phase. A substantial increase began after 2017, with citations rising from 654 in 2017 to 1,532 in 2018 and 1,688 in 2020 as shown in Figure 3. There were fluctuations but the impact on citations was not lost, with the 2023 citation of 2,120 and 1,306 in 2024. The reduction in citations in 2025 and 2026 may be attributable to the brief period of time in which recent publications have been accumulating citations. It illustrates the academic trend and the increasing importance of reliability, robustness, sensitivity and validation in MCDM studies in general.

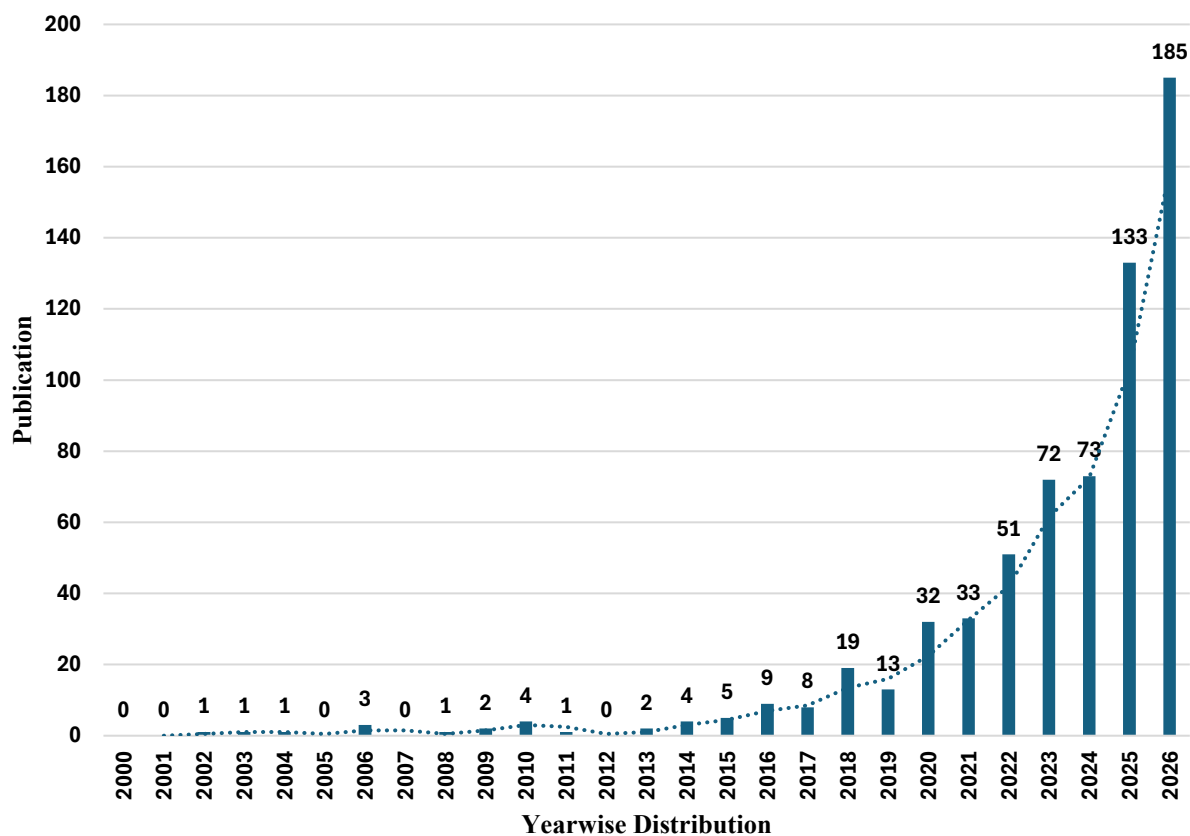


Fig. 2. Publication Count

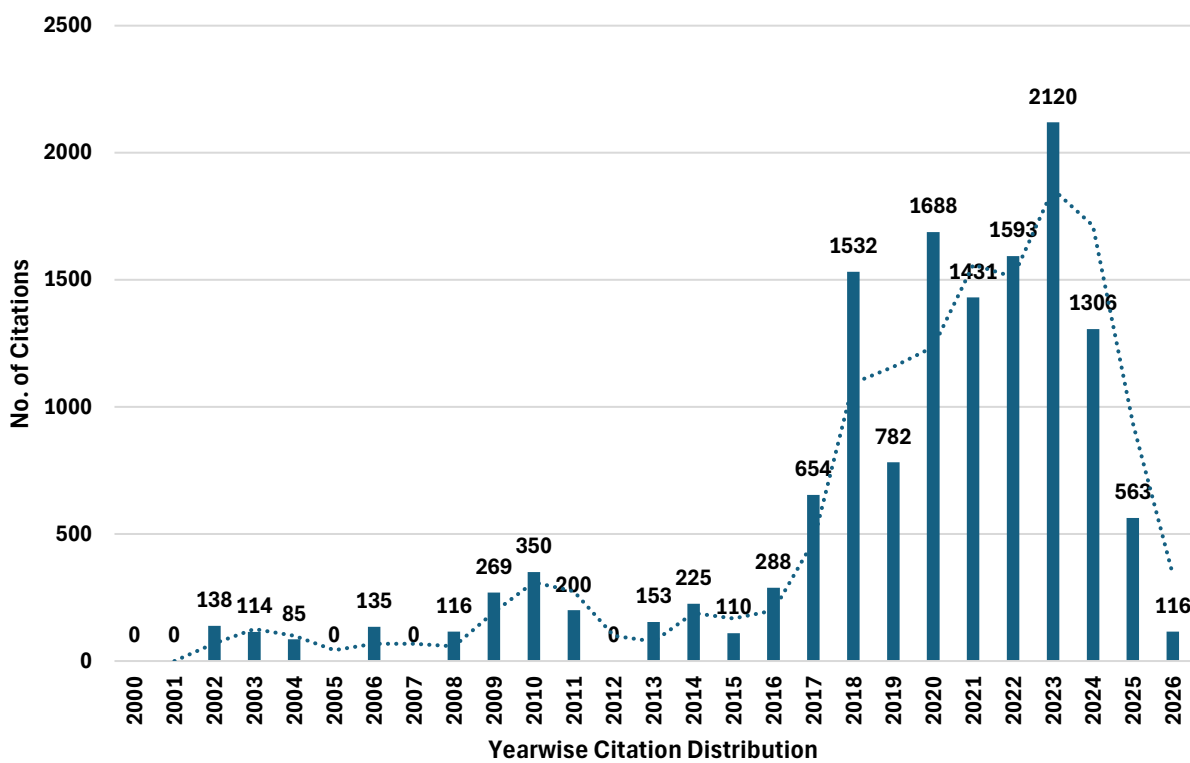


Fig. 3. Citation Count

4.2 Most Productive Sources

As revealed by the source analysis the study of sensitivity and validation in MCDM has been published in multidisciplinary journals, especially in decision sciences, AI, sustainability and environmental research. The most publications were in the category of Sustainability (19) and Scientific Reports (19) (as in Figure 4). The number of publications, however, does not necessarily reflect the impact of citations. The total number of citations received by Expert Systems with Applications and Applied Soft Computing are 506 and 370 from 13 and 15 documents respectively. In particular, Symmetry was cited 577 times by 9 documents and Science of the Total Environment 547 times by 8 documents, which shows high research influence. The overall results prove that the decision making in MCDM is interdisciplinary in nature and is gaining more and more importance in the academic field.

4.3 Leading Authors

Research on sensitivity analysis and validation in MCDM is dominated by a small group of highly productive and influential researchers as depicted in Figure 5. Albahri, A.S. has the greatest number of publications (21), followed by Albahri, O.S. and Pamucar, Dragan (20). As far as citations are concerned, the most cited documents are of Dragan Pamučar with 1,004 citations and of Željko Stević with 817 citations, followed by the documents of Dragan Pamučar with 779 citations. Some other researchers (Zaidan, A.A., Qahtan, Sarah and Deveci, Muhammet) also show high citation influence. Generally, the findings indicate that the research on advanced MCDM, uncertainty analysis, sensitivity analysis, and decision supporting methodologies have contributed significantly.

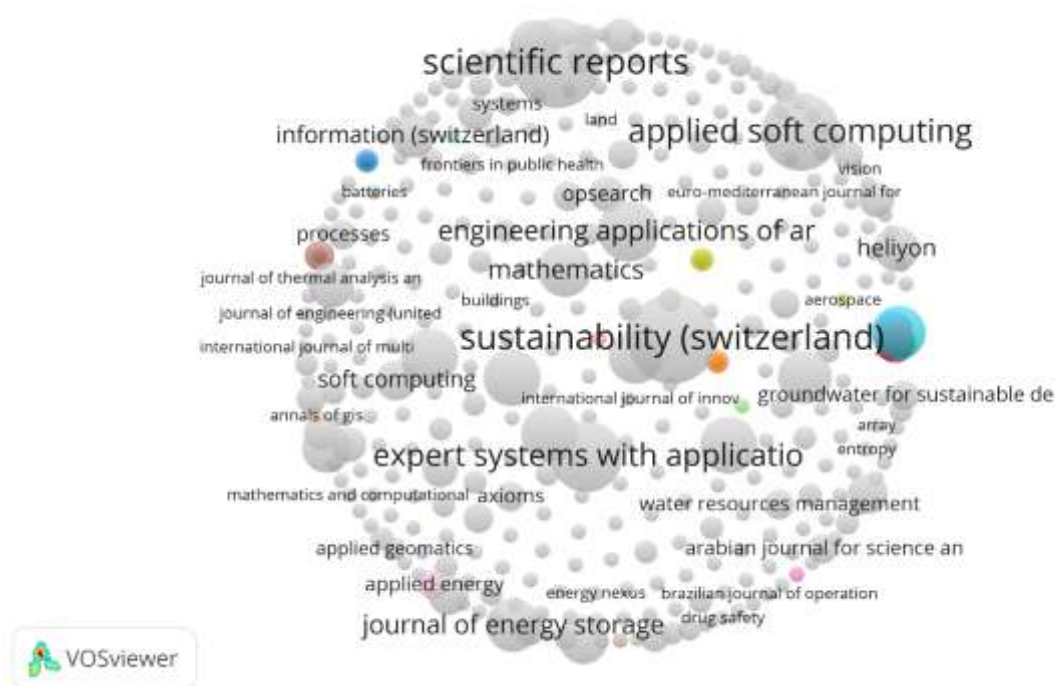


Fig. 4. Most Productive Sources

There is also a strong association between the central cluster and the terms “validation” and “uncertainty”, which reflects the increasing interest in determining the reliability and robustness of the MCDM results. The other crucial keywords are “multi-attribute decision-making”, “fuzzy MCDM”, “analytic hierarchy process (AHP),” “renewable energy”, “sustainability assessment”, “uncertainty quantification”, and “simulation-optimization”. The network also uncovers themes relevant to the application, like the Supply Chain, Electric Vehicles, Optimization in the Construction Industry and Automotive Industries. Overall, the structure of the keywords shows that the current research in MCDM is increasingly incorporating the concepts of sensitivity analysis, validation, modeling uncertainty, fuzzy techniques and application-specific decision frameworks in order to enhance the reliability and robustness of the decision.

5. Sensitivity Analysis Methods in MCDM

Sensitivity analysis plays a crucial role in Multi-Criteria Decision-Making (MCDM) by examining how changes in input parameters, assumptions, and methodological choices affect the scores and rankings of alternatives. The literature has progressively expanded toward other sources of uncertainty and methodological variation, including performance-value perturbation, alternative-set modification, scenario-based analysis, probabilistic and simulation-based approaches, fuzzy and interval analysis, and model- or algorithm-based sensitivity analysis. Accordingly, in this study, sensitivity-analysis methods are systematically classified based on what is varied, how the variation is introduced, which indicators are used to quantify the resulting changes, and the underlying assumptions and limitations. This methodological classification provides a structured basis for comparing different sensitivity-analysis approaches and selecting an appropriate method according to the nature of uncertainty in an MCDM problem.

5.1 Weight-Based Sensitivity Analysis

Among the widely used techniques in the MCDM, the weight-based sensitivity analysis [37] is one of the most common. It discusses the effect of varying the weights of the criteria on the ranking. A common approach is the one-at-a-time or incremental adjustment of the criterion weights, e.g., increasing or decreasing a weight by a certain percentage. The simplicity, transparency and ease in which this approach can be put into practice is really appealing. But it may not be a good illustration of multiple complex and uncertain criteria, though. Other ways to explore weight sensitivity are simultaneous variation, generating weights at random, weight redistribution, "equal-weight" scenarios, "extreme-weight" scenarios, and rank-order-preserving generation. Several weight vectors can be created using a probabilistic method and the frequencies of the weight ranks can be discussed. The literature indicates that using "weight variation" is still a commonly used method, and that the other weight vectors used in the model MCDM can be easily calculated.

5.1.1 One-at-a-Time Weight Variation

The one at a time method involves adjusting one weight at a time, while maintaining the other weights. The change in the ranking is indicated. This technique can be used as a guide to finding the "important" criteria and to determining if a desired alternative is "low" or "high" on a criterion. One-at-a-time weight changes were applied to measure the sensitivity of rankings of sustainable manufacturing process selection by Abdel-Basset *et al.*, and it was found that key factors like energy efficiency, cost, and environmental impact can have a significant impact on the results [38]. Therefore, this approach enhances reliability of the decision-making process in manufacturing. Likewise, Kitila *et al.* used the same for sustainable manufacturing process selection of 316 Stainless Steel to achieve alignment with sustainability targets and considered a wide variety of economic,

environmental and social criteria [39]. Ranjgar *et al.* discussed solar photovoltaic (PV) site selection by systematically altering the weight of each criterion while keeping the others the same to determine the effect of each criterion. In OAT studies, the percentage of variation can vary from $\pm 20\%$ to $\pm 10\%$, $\pm 5\%$, $\pm 2\%$, or $\pm 1\%$. In their study they chose a balanced range of 10% around the mean because it was a compromise between computational feasibility and interpretability [40].

5.1.2 Simultaneous Weight Variation

In sensitivity testing, that several or all of the weights associated with the criteria are changed at once. This is a more typical case where more than one criterion are uncertain, simultaneously. Simultaneous Weight Variation (SWV) sensitivity analysis in MCDM is a sensitivity analysis that ranks simultaneously by varying multiple or all of the weights of the various criteria rather than varying only one criterion at a time (One-At-a-Time/OAT) [41]. Foundational work by Mukhametzanov [42] and O'Shea *et al.*, [43] note that OATs procedures often overestimate decision stability because in reality, trade-offs tend to occur in more than one dimension and affect the relative importance of the other dimensions. Mukhametzanov points out that objective weights, especially the Entropy Weighting Method (EWM), could be useful for obtaining useful data driven weights from the decision matrix. The study, however, shows that EWM is very sensitive to the choice of the assumed probability distribution of the alternatives. The weights and rankings can vary greatly depending on the slight change in the values for the attributes. To preserve the foundational normalization constraint $\sum_{j=1}^n w_j = 1$, modern sensitivity models simultaneously adjust the weights across all n criteria. When each criterion j undergoes a concurrent perturbation factor $\delta_j \in [-\alpha_j, +\alpha_j]$ (or when a selected target weight w_k is perturbed to w_k^*), the normalized simultaneous weight adjustment is modeled as:

$$w_j^* = \frac{w_j(1 + \delta_j)}{\sum_{r=1}^n w_r(1 + \delta_r)} \quad (1)$$

Although alternative rankings can be tested as much as univariate rankings, there are some important shortcomings in the simultaneous weight variation approach. First, it is difficult to explore the multi-dimensional weight hyperspace because of the high computational complexity and combinatorial explosion, and it can be very time consuming to trace the stability thresholds with the use of either extensive Monte Carlo simulations, closed form polyhedral projections or non-linear solvers [44]. Second, it is hard for decision-makers to visualize and intuitively interpret simultaneous multi-parameter movements that are covered in the 3D plot, as opposed to 2D sensitivity plots [45]. Lastly, unconstrained simultaneous perturbations can generate extreme or correlated weight vectors that exceed the cognitive preference hierarchy and knowledge of the decision-maker, which means that decision scenarios mathematically possible but not practically feasible are tested.

5.1.3 Scenario-Based Weight Variation

Various weights may be assigned to different sets of weights, representing the preferences of decision makers, priorities of stakeholders or strategic situations. This is followed by a re-ranking of each scenario. Abbasi [46] investigated on scenario-based weight variation for sustainable recycling of NCM Lithium-ion battery cathodes, which allows to explore the influence of environmental regulation, economic priorities and energy limitations on selection of recycling strategies. Environmental and policy analyses were conducted using scenarios to see if certain extreme market and/or energy constrained conditions would make alternative methods more competitive, but the hydrometallurgical process was always competitive. To improve the low-carbon supplier selection

process, Maden proposed a scenario-based, integrated solution, which is based on the Fuzzy Cognitive Map (FCM) and Evaluation based on Distance from Average Solution (EDAS) [47]. In the same way, Stevic *et al.*, [48] applied the scenario based on changing the weights of various criterion and identified “most significant criterion” for sustainable supplier selection in health care industries by utilizing following Equations.

$$w_c = (1 - w_s) \times \left(\frac{w_c^\circ}{W_c^\circ} \right) = w_c^\circ - \Delta x \alpha_c \quad (2)$$

$$\alpha_c = w_c^\circ / W_c^\circ \quad (3)$$

Where w_c represents the change in the weights of criteria in the sensitivity analysis, represents the w_s weight of the most significant criterion, w_c° represents the original values of the weights of criteria and W_c° represents the sum of the original values of the weights of criteria which are changed. The parameter is α_c defined as the weight coefficient of elasticity, which expresses the relative compensation of other values of the weight coefficients in relation to the changes in the weight of the most significant criterion.

$$w_s = w_s^\circ + \alpha_s \Delta x \quad (4)$$

$$w_c = w_c^\circ + \alpha_c \Delta x \quad (5)$$

Where w_s° is the original weight of the criterion subjected to the sensitivity analysis, w_c° is the original value of variable weights. This new set of criteria always satisfies the universal state of proportionality of weight coefficients that $\sum w_s + \sum w_c = 1$. The new values of the criteria are obtained and the ranks of alternative for the scenario observed are calculated. The method is systematic, mathematically bounded weight changes which always maintain the normalization of the non-focal criteria while maintaining the proportional relative weight ratios without any negative weights being produced [49]. The problem with this approach is that it has a rigid set of proportionality weights for the non-focal criteria, and only allows for one criterion to change at a time rather than the weights changing in multiple ways, independent of each other, across the non-focal criteria.

5.1.4 Probabilistic Weight Analysis

Weights instead of being concrete values, can be represented by probability distributions. Then a series of simulations can be performed to approximate the probabilities of other alternatives receiving a given ranking. Probabilistic weight sensitivity analysis can represent uncertain weights of the criteria and analyze the impact of the sensitivity of the weight on the ranking of the MCDM [50]. In Kala [51], the sensitivity coefficients are regarded as weight factors and the uncertain inputs are modelled probabilistically. The good news is that its positive is that global sensitivity methods include input variability and interactions and are useful for decision analysis. Furthermore, the study indicates that the sensitivity analysis (ROSA) results were not consistent and that interaction effects are strong when failure probabilities are low and difficult to compute when they are low.

Likewise, Sun *et al.*, [52] provides the weight sensitivity analysis of uncertain MCDM based on six fuzzy importance-weight sets generated. Its robustness is that A4 is the best choice for every set, and that its value of belief does not vary. The positive is its robustness: A4 is the best choice for every set, and the value of its belief does not change. It is also a method that shows the effect of the subjective weight changes and corrects the uncertainty of experts. One drawback is that Set 4 alters the ranking of A1 and A3, indicating subjective weights and a potential human-error influence. Furthermore, Feizizadeh and Blaschke [53] use probabilistic weight sensitivity analysis, which is a combination of AHP, PERT probability distributions and Monte Carlo simulation (MCS). Its key advantage is that it allows for the use of probability distributions to replace fixed, subjective weights, and it accounts for

epistemic and aleatory uncertainty, along with quantifying variations with standard deviation, confidence intervals, and percentiles. The method has advantages of robustness and accuracy in landslide susceptibility mapping. Its drawback, though, is that it is limited to expert-defined weight ranges, and assumes distributions, and MCS is still subjective AHP judgments-dependent and is computationally challenging.

5.2 Decision-Matrix Sensitivity Analysis

In MCDM, decision-matrix sensitivity analysis has progressed from a weight perturbation approach to a more general approach which considers the stability of the rankings under various uncertainty structures [54]. According to recent studies, the strength of the results of MCDM is significantly influenced by the dispersion and correlation among the scores of the criteria in the decision matrix, and the compensatory nature of aggregation rules used. For example, Gaona *et al.*, [55] proposed an integrated TOPSIS framework with full range-weight sensitivity analysis, which showed that the local sensitivity analysis methods used in the traditional TOPSIS often miss the global instability regions where rank reversals occur. Likewise, Lian *et al.*, [56] systematically studied fifteen MCDM methods to be used in automotive structural optimization and found that the methods with the highest compensatory characteristics are more likely to have rank reversals when the weights of the criteria vary within a narrow range.

However, existing sensitivity analysis approaches still have many shortcomings, especially with respect to computational complexity and the modelling of deep uncertainty. The strong computational demands of global sensitivity analyses (e.g., Monte Carlo simulations or Stochastic Multicriteria Acceptability Analysis [SMAA]) constrain their use in high dimensional decision problems. Pala [57] points out that there is no systematic way to select an MCDM method depending on their sensitivity profiles and practitioners often select methods which may result in different solutions for the same problem without knowing the stability mechanism. Moreover, Nazarov [58] attributes weaknesses to traditional approaches such as MOORA and WASPAS in using them effectively when uncertainty exists because they are not able to account for qualitative information or changes in the environment over time.

5.3 Alternative-Level Sensitivity Analysis

MCDM alternative-level sensitivity analysis is a significant methodological improvement that allows to assess the robustness of the decision result by examining how the change of the input parameters influences the ranking and selection of particular alternatives [59]. Traditional sensitivity analysis looks at whether the weights for the criteria change as a result of the uncertainty about the data and measurement error, while alternative-level sensitivity analysis examines the values of the performance measures for each individual alternative to see if they are sensitive to data uncertainty and measurement error [60]. Alternative-level sensitivity analysis is particularly common where stakes are high and uncertainty is high, like planning renewable energy, managing water resources, and logistics in supply chains. Sensitivity analysis is used to determine which geographic and/or technical criteria are most important to potential sites and directs data collection on the variables that will have the greatest impact on site suitability in renewable energy site selection [61]. Likewise, water resources planning involves sensitivity analysis of performance values which helps to mitigate the uncertainty that exists in hydrological data to support infrastructure investments that are based on strong and reliable assessments [62]. The ability to pinpoint the most critical performance measures allows for more efficient resource allocation in data gathering and model refinement, ultimately enhancing the credibility of the decision-making process [63].

Furthermore, the sensitivity of changes in alternative ranking because of adding or removing an alternative or slight changes in the weights of the criteria is a central issue in alternative-level sensitivity [64]. It is important to understand these circumstances to guarantee the consistency and reliability of the outcomes of MCDM [65] is achieved. It has been found that various MCDM methods have different sensitivity to rank reversal as some are more stable than others [66]. The use of detailed sensitivity analysis can enable the practitioner to choose the most suitable method for MCDM for a particular context to ensure that the method is discriminative enough in one sense while, at the same time, robust to data uncertainty.

5.4 Scenario-Based Sensitivity Analysis

Sensitivity analysis based on scenarios is one approach to analysing results of the MCDM using different decision situations [67, 68]. In MCDM, scenario-based sensitivity analysis assesses the robustness of the different alternatives' rankings by creating a set of states of the parameters in advance, in accordance with a policy or stakeholder consideration, instead of arbitrarily making continuous perturbations in the weight hyperspaces. In more recent work, Azzini and Munda [69] report that one-at-a-time weight perturbations do not reflect the complexity of real world policies, and when they define scenarios explicitly as weights, they can integrate competing social perspectives into a prioritised set of policy packages that explicitly show where consensus is found and where it is lacking. Likewise, in the realm of industrial warehouse automation, Purwanto *et al.*, [70] examine how the prioritization changes with changing operational pressures, and they stress the importance of evaluating the situation-oriented weighting structures for different combinations of competing operational, ergonomic and space criteria to ensure that the ranking of the automation technologies remains consistent in the event of a sudden change of goals, e.g., from maximum throughput to maximum space saving.

With this regard, from a sustainable infrastructure point of view, Özkurt [71] states that the real-world prioritization process is significantly influenced by the divergent viewpoints of the various stakeholders involved, and that it is useful to assess the various scenarios individually, such as those with an economic dominant view, the environment-focused view, and the technology view, in order to provide decision makers the ability to check if the selected compromise solution remains the top priority even in the event of a significant change in policy. As applied results, benchmark studies for the method highlight the structural behaviour of the scenario test, under uncertainty. Gaona *et al.*, [55] investigate predefined scenarios in addition to full range weight variation and find that scenarios yield an operational baseline in the form of "an anchor of practical sense" that avoids mathematically valid but practically meaningless weight combinations, which help identify realistic rank-reversal thresholds. Hashemkhani and Nemati [72] highlight through multi-scenario simulations the importance of systematic scenario changes to identify the rank changes for non-dominant alternatives when threshold parameter changes and the interactions between the thresholds change.

5.5 Model-Based Sensitivity Analysis

The use of model-based sensitivity analysis (MBSA) in MCDM considers the stability and consistency of alternative rankings across structurally different mathematical models (e.g., distance-to-ideal, outranking, compromise-solution, ratio-based aggregations), rather than only parameter changes [73]. Więckowski and Sałabun [74] make this observation in a thorough comparative study over a wide range of decision contexts: The rank volatility of distance-based models is considerably greater than that of compensatory models while the latter show greater structural stability. Kassara *et al.*, [75] note that the underlying assumptions on which distances are measured – either to an average solution or to minimum individual regret to maximize group utility – often change the

alternative that is ranked first in model selection in machine learning, where VIKOR showed higher stability of the ranking result under the same set of criterion weights. Likewise, in high stakes failure mode prioritization, Şengül and Karahan [76] compare and validate across the three methods by calculating Spearman rank correlation coefficients between the three follow-up results: the core alternative dominance does not change in the secondary and intermediate positions, and high correlation coefficients are obtained, even though ranks change in these positions across the three methods. The most salient problem is the lack of a meta-criterion or mathematical ground truth: if the different algorithms that compete in an MCDM problem yield conflicting rankings for a given set of data, decision analysts have no non-arbitrary basis for deciding which of the rankings is the "best". In addition, cross-method benchmarking requires a significant cognitive burden and computational burden because each method of MCDM has its own structural parameters tuning, thresholds specifications, and mathematical normalization processing.

5.6 Probabilistic and Simulation-Based Sensitivity Analysis

Probabilistic or Simulation-based sensitivity analysis in MCDM investigates the robustness and consistency of the multiple possible rankings that may result from alternative weights of the criteria, performance values, and the decision parameters, which are generally assumed to be continuous random variables or probability distributions [77, 78]. Dunke and Nickel [79] discussed simulation-based sensitivity analysis as an iterative approach for evaluating parameter uncertainty and system dynamics in multi-criteria decision-making. Using global sensitivity analysis, particularly Sobol' sensitivity indices, their methodology identifies influential parameters and strengthens the reliability of decision alternatives under uncertain epidemic conditions. Baležentis and Streimikiene [80] consider offshore engineering systems and address the problem of decommissioning based on dynamic Monte Carlo simulation combined with outranking techniques, noting that sampling of uncertain cost, environmental and technical parameters provide an entire distribution of net flows and probabilities of ranking decommissioning options, which highlights compromise solutions that are robust for thousands of simulated operational systems. Likewise, in an industrial materials optimization setting, Rahman *et al.*, [81] use a combination of objective criteria weighting and sensitivity analysis based on Monte Carlo simulations, noting that the Monte Carlo simulations validate the predictive consistency of the structure of compromise alternatives by monitoring the ranking variations in a large number of simulations.

It extends these findings to more complex operational and technological planning, and Valenzuela [82] develop a model based on distance-from-average decision models and Monte Carlo simulation that enables decision makers to examine the expected performance score as well as the variance spreads of the plans, and they show that these more reliable models lead to significantly better selections than deterministic or isolated multi-criteria models. Moreover, the resulting ranking probabilities are significantly affected by the distribution forms that are specified (e.g., uniform, normal, or beta distributions), as well as by the set of variance bounds that are assigned to them, which introduces a second level of subjectivity by the modeler that can lead to invalid output if the underlying statistical properties are miss specified.

5.7 Fuzzy, Interval, and Uncertainty-Based Sensitivity Analysis

In fuzzy, interval and uncertainty-based sensitivity analysis in MCDM, the parameters are modeled as a fuzzy set, membership function, grey interval or neutrosophic number instead of a deterministic scalar, and the alternative rankings are evaluated with respect to the epistemic imprecision, linguistic vagueness and incomplete cognitive information [83]. In a complex industrial system, Kaspar and Kaliyaperumal [84] study the technological roadmaps by a hybrid neutrosophic

approach; their sensitivity tests show that, by progressively changing the value of truth, indeterminacy and falsity parameters, a systematic stress test can be performed to identify ranking vulnerabilities that are not revealed by crisp or purely additive evaluation. In the case of sustainable manufacturing process selection, Lebepe and Mathaba [85] employ an integrated fuzzy multi-criteria approach, and, using a systematic fuzzy weight adjustment procedure, they show that the inclusion of triangular fuzzy boundaries in the sensitivity analysis yields robust compromise rankings that are able to resist real-world uncertainties in the subjective preferences of stakeholders.

In addition to these applied research, methodological researchers highlight the crucial diagnose properties of interval-based testing and hesitation-aware testing. Ibrahim *et al.*, [86] investigate the rank stability of a modified fuzzy AHP–TOPSIS model and conclude that monitoring the ranking behavior in the vicinity of the fuzzy boundaries and changing the fuzzy set boundaries make the model less rank-reversal sensitive, which is a common problem in crisp distance-based models. Moreover, Ayyildiz *et al.*, [87] shows in a facility location model, that the sensitivity analysis of an intuitionistic fuzzy interval can reveal subtle crossover points at which a small change in degree-of-hesitation can cause the relative closeness of the competing alternatives to flip. The major positive result of these investigations is that fuzzy and interval-based sensitivity analyses allow combining the mathematical rigor with the cognitive ambiguity of human decision making, allowing decision analysts to stress test the boundaries of stability within realistic operating intervals that are not crisp. The biggest problem is the uncertainty propagation when propagating intervals and fuzzy numbers through non-linear multi-criteria aggregation mechanisms, the intervals can get too large and the output intervals too wide, creating large overlapping ranges of rank where alternatives compete with each other and are statistically indistinguishable.

5.8 Comparative Summary of Sensitivity Analysis Methods

As seen in Table 4, all these sensitivity analysis methods have various features and characteristics in terms of parameters analyzed, representation of uncertainty, computational complexity, and robustness of the results when it comes to ranking. These approaches can be compared and contrasted in their merits and disadvantages and select suitable approaches according to the type of uncertainty and the characteristics of the MCDM problem.

Table 4

Comparative Summary of Sensitivity Analysis Methods

Sensitivity Analysis Method	What is Perturbed?	How is it Perturbed?	Main Output Indicator	Key Assumptions	Advantages	Limitations
Weight-Based	Criterion weights	One-at-a-time (OAT), simultaneous, incremental, or weight redistribution	Rank change, rank stability, rank correlation, critical weight	Weights can be meaningfully varied while satisfying weight constraints	Simple, transparent, easy to implement	Mainly considers uncertainty in criterion importance
Probabilistic Weight	Criterion weights	Probability distributions, random sampling, PERT distribution, Monte Carlo simulation	Rank probability, confidence interval, variance	Probability distributions adequately represent weight uncertainty	Captures uncertainty and variability in weights	Depends on distribution assumptions and may require high computation

Table 4
Continued

Sensitivity Analysis Method	What is Perturbed?	How is it Perturbed?	Main Output Indicator	Key Assumptions	Advantages	Limitations
Decision-Matrix	Alternative performance values	Percentage change, random noise, error range, interval variation	Score change, rank change, rank reversal, rank correlation	Perturbation represents estimation uncertainty	Captures uncertainty in performance data	Results depend on selected perturbation magnitude
Alternative-Level	Alternatives and alternative set	Adding, removing, replacing, or modifying alternatives	Rank reversal, rank stability, rank correlation	Changes in the alternative set are realistic	Examines structural stability of the decision model	Can be computationally demanding
Scenario-Based	Multiple decision inputs	Baseline, optimistic, pessimistic, stakeholder	Scenario ranking	Plausible decision conditions	Represents multidimensional uncertainty	Scenario construction may be subjective
Simulation-Based	Multiple uncertain inputs	Repeated random sampling and simulation, e.g., Monte Carlo	Variance, confidence interval, ranking	Input distributions are appropriately specified	Captures simultaneous uncertainty and interactions	Computationally intensive
Fuzzy	Fuzzy weights, performance values, or expert judgments	Fuzzy numbers, membership functions, fuzzy parameter variation	Fuzzy score/rank, defuzzied rank, rank stability	Fuzzy numbers adequately represent vagueness	Suitable for linguistic and subjective information	Membership-function selection can be subjective
Interval-Based	Weights, parameters	Lower–upper bounds and interval variation	Score range, rank stability	uncertainty	Does not require probability distributions	Does not provide probability of occurrence
Model-Based	MCDM method, normalization, aggregation	Alternative methods or methodological configurations	Rank correlation, rank change	compared under comparable conditions	Identifies methodological dependency	Different MCDM have different assumptions

6. Validation Methods in MCDM

Along with sensitivity analysis, validation will bring evidence to support the credibility, consistency, suitability and usefulness of an MCDM model and its results [88, 89]. The major approaches found in the literature that are followed in the present study are as follows: Internal validation, External validation, Robustness checks, Monte Carlo simulation, Expert validation, Statistical validation, Cross-method triangulation.

6.1 Internal Validation

The Analytic Hierarchy Process (AHP) is a structured MCDM method that uses pairwise comparison to express preferences, according to Saaty [90]. The internal consistency of these judgments needs to be checked prior to the interpretation of the resulting weights and rankings as they are the basis for deriving criterion priorities. Internal validation, thus, examines the logical consistency of the decision model and the consistency of the underlying judgments and computations

involved in the model. The validation is especially applicable to the pairwise comparison methods like AHP and ANP.

6.1.1 Consistency Ratio

Saaty suggested a measure of consistency of pairwise-comparison judgments, called the Consistency Ratio (CR). It compares the inconsistency in a comparison matrix to the inconsistency that would be expected from random judgments. The CR is derived from the following:

$$CR = \frac{CI}{RI} \quad (6)$$

where CI is the Consistency Index and RI is the Random Index. A commonly used criterion is $CR \leq 0.10$, indicating an acceptable level of inconsistency, whereas higher values suggest that the pairwise judgments should be reconsidered.

6.1.2 Consistency Index

To measure the deviation from perfect consistency of a matrix of pairwise comparisons, Saaty developed a new metric, called the Consistency Index (CI). It is computed from the following equation:

$$CI = \frac{\lambda_{max} - n}{n - 1} \quad (7)$$

where λ_{max} represents the maximum eigenvalue of the comparison matrix and n is its order. For a perfectly consistent matrix, $\lambda_{max}=n$, and therefore $CI = 0$. An increasing CI indicates increasing inconsistency in the pairwise judgments.

6.1.3 Logical Consistency

The pairwise comparisons must be consistent with an admissible consistency level, since decision makers can make judgments that are contradictory or hard to reconcile, Saaty stated. Logical consistency, then, is the study of if the preference relationships between the criteria are coherent. For instance, if C1 is preferred over C2, and C2 is preferred over C3, then the resulting preference structure should not be in contrast to the relationship between C1 and C3. These quality checks are useful for detecting faulty judgments at an early stage before they get reflected in criterion weights and final rankings.

6.1.4 Redundancy and Convergence Checks

Saaty pointed out that it is important to build a sound hierarchical structure with relevant criteria and alternatives. In the same sense, internal model checking can also be used to check if the criteria itself give different information for decision making, and if iterative procedures (where applicable) can lead to a stable solution. When there are multiple criteria, they can provide redundant information and have the unwanted effect of giving more weight to one decision dimension. The convergence checks are applicable principally to iterative or computational procedure, and the stopping criterion should be explicitly defined for that purpose. These checks should not be used as evidence of the structure's external validity or as indicators of robustness, but rather are used to support structural validity.

6.2 Robustness and Stability Validation

Cinelli *et al.*, [91] identify difference between robustness and sensitivity analysis in MCDA. Robustness refers to whether a recommendation will stand when most or all potential combination of uncertain performance values and preference parameters are considered, while sensitivity

analysis focuses on how a current recommendation will change if certain assumptions and inputs are modified. Accordingly, robustness and stability should be evaluated by systematically varying the model's inputs that are likely to be relevant and checking if the scores, rankings or preferred alternatives are consistent. The steps of a practical procedure may be written as in Figure 9.

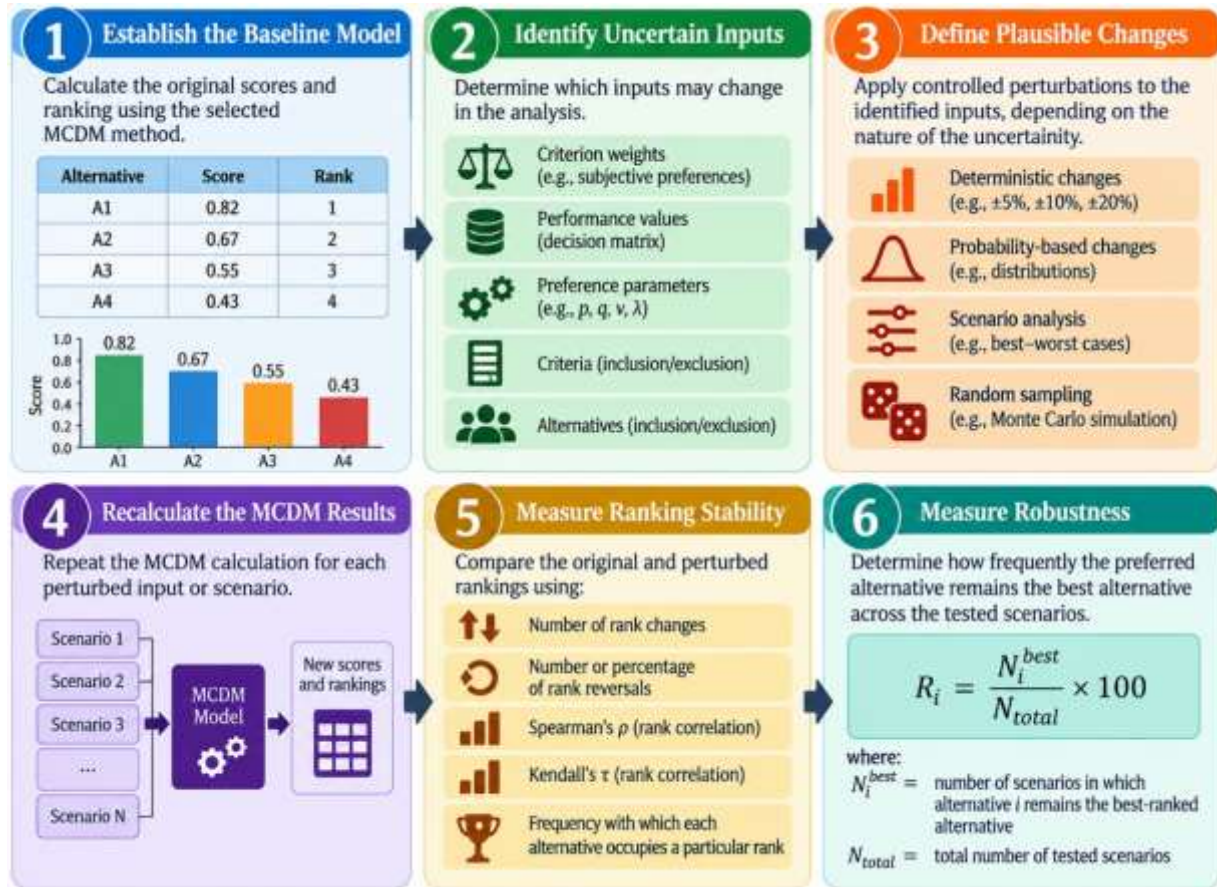


Fig. 9. Robustness and stability validation in MCDM

Original and perturbed rankings are compared based on the value of measures like rank changes, rank correlation, rank reversal frequency and ranking stability. Ganji *et al.*, [92] showed an approach that was more formal based on a modified version of the Sobol sensitivity-analysis framework. Their approach assesses the robustness or stability of the relative ranking of alternatives with respect to shifts in the criteria for assessing the alternatives and changes in the preferences of the stakeholders, and offers sensitivity measures of the ranking.

A more recent proposal by Więckowski *et al.*, [93] proposes an integrated benchmarking framework that takes into account a change in criterion weights, as well as a change in criterion exclusion and a change in decision-matrix values. They run repeated perturbed cases with the MCDA methods to discuss their differences in stability and robustness in their simulation approach. This shows how robustness testing can be expanded beyond one source of uncertainty to many sources of uncertainty interacting. Thus, in the context of Stability, the question is, "What happens if the inputs to the model are modified?"; in the context of robustness, the question is, "What happens if there is uncertainty in the inputs for the model, but the preferred decision remains valid? External validation, however, is whether the decision corresponds to independent evidence or actual results. This is significant because even if the result of the MCDM is very stable, the data processed, the assumptions made and the criteria used, and the structure of the model, may all be inappropriate.

Thus, alongside external validation, stability and robustness should be considered as measures of reliability.

6.3 Comparative Validation

Duško and Bozanic [94] proposed a framework for the validation of different MCDM methods in which they can be compared by the same decision conditions under which they were used to rank the alternatives, to check the consistency and effectiveness of the results obtained. Their approach employs the standard measures of difference between rankings, such as mean absolute rank deviation and Kendall's tau. This is consistent with the suggestion that cross-method comparison can be used in the evaluation of ranking agreement. Comparative validation is conducted by using two or more MCDM methods to analyze a set of alternatives, a set of criteria, performance data, and preferably a uniform weighting scheme, and then comparing the resultant rankings. For instance, TOPSIS, VIKOR, EDAS and COPRAS can be used to the same decision problem [95]. Spearman's rank correlation coefficient, Kendall's tau, rank differences, rank reversal frequency, and the agreement of the top-ranked alternatives are all measures of the degree of agreement. If there is a high rank correlation and consistent identification of leading alternatives, then there is greater stability across methods.

More recently, Spearman's rank correlation is used to measure the agreement between different MCDM approaches to produce rankings. Baydaş *et al.*, [96] noted, however, that MCDA is not a single method with a single philosophy, and that the method chosen should be a function of the structure of the decision problem, not a universal best method. Thus, consistency of the rankings obtained by the various MCDM methods should be interpreted as a convergence, rather than as evidence that the chosen alternative is truly optimal. Therefore, comparative validation helps to answer the question: "Are the different decisions yielded by the different MCDM methods the same for the same problem?" It can give evidence about methodological stability and ranking agreement, if no independent ground truth is available. But no agreement does not mean correctness, optimality or external validity since there may be several methods using the same data, assumptions, criteria or weighting judgments. Comparative validation is not an alternative to sensitivity analysis, robustness assessment, expert validation or external validation, but rather a complement to these techniques.

6.4 Statistical Validation

The statistical validation of MCDM should be treated as a group of quantitative processes used to test the validity of certain properties of the results of the decision rather than as a test of the overall validity of the model. statistical validation used mean absolute rank deviation and Kendall's τ to quantify the consistency of MCDM rankings with reference rankings. Their study shows that it is possible to get quantitative evidence of agreement or difference in ranking by using statistical measures, but it is important that each of these measures relates to a particular property being evaluated.

Spearman [97] developed a rank correlation measure to evaluate the association between two ordered variables. Spearman's rank correlation coefficient (ρ_s) is a coefficient that can be used in MCDM to assess the similarity between two alternative rankings. When there are no tied ranks, it is computed as:

$$\rho_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (8)$$

where d_i is the difference between the ranks assigned to alternative i by the two methods or conditions. A value near +1 means a strong agreement in terms of rank ordering while a value near

-1 means an inverse ordering of the ranks. Therefore, the main purpose of Spearman's ρ_s test is to determine if there is an association between the rankings, rather than if the ranking is truly correct.

Kendall [98] created Kendall's rank correlation as a non-parametric measure of ordinal association. It considers the relative number of concordant and discordant pairs for two rankings:

$$\tau = \frac{C - D}{\frac{1}{2}n(n-1)} \quad (9)$$

where C is the number of concordant pairs and D is the number of discordant pairs. The higher the τ , the greater the ordinal agreement of the ranks. Kendall's τ is used as a statistic to assess the consistency of the rankings obtained by using the MCDM approach. Hence, Spearman's ρ_s and Kendall's τ are best thought of as ranking agreement rather than measures of external validity or correctness.

If the objective is to determine the numerical association of two sets of MCDM scores rather than ordinal rankings, Pearson's product-moment correlation coefficient (r) can be used to quantify the linear association [99]:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (10)$$

where X_i and Y_i are the scores produced by two methods or conditions, and $\bar{X}$ and $\bar{Y}$ are their respective means. A high r indicates that the two sets of scores have a similar linear pattern. However, Pearson's correlation does not establish equality, correctness, or causal validity. It should therefore be interpreted as score association, rather than ranking validation.

Osmond *et al.*, developed the analysis of variance approach to deciding whether there are significant differences between group means [100]. The ANOVA can be used in the MCDM research if the numerical results are generated using various MCDM methods, scenarios, parameter settings, or experimental conditions. The basic one-way ANOVA statistic is:

$$F = \frac{MS_{between}}{MS_{within}} \quad (11)$$

where $MS_{between}$ represents variation between group means and MS_{within} represents variation within groups. A statistically significant F-test indicates that at least one group mean differs from another. Therefore, ANOVA evaluates statistical differences between numerical groups, rather than ranking agreement or ranking robustness.

Neyman provided the framework of the "confidence-interval" for estimating an unknown population parameter using the information in the sample [101]. In MCDM, confidence intervals may be created when the scores, the probabilities of ranking, or other statistics are estimated based on repeated observation, simulation, or resampling. In particular, a confidence interval for the mean score is:

$$\bar{X} \pm t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}} \quad (12)$$

where $\bar{X}$ is the sample mean, s is the sample standard deviation, n is the sample size, and $t_{\frac{\alpha}{2}, n-1}$ is the appropriate critical value. A narrower interval generally indicates greater precision of the estimated quantity. Confidence intervals therefore assess estimation uncertainty, rather than agreement between rankings or external validity. The summary of statistical validation is shown in Table 5.

Table 5
Statistical Measures and What this method Validate

Statistical procedure	What is evaluated?	Primary validation question
Spearman's ρ_s	Association between rankings	Do two rankings have a similar order?
Kendall's τ	Ordinal agreement	How strongly do two rankings agree?
Pearson's r	Linear association between scores	Do two sets of numerical scores vary similarly?
ANOVA	Differences between group means	Are observed numerical differences statistically significant?
Confidence interval	Precision/uncertainty of an estimate	How uncertain is the estimated quantity?

6.5 Expert-Based Validation

The expert-based validation is especially important for MCDM problems where there is no objective ground truth, benchmark or experimentally obtained optimum. MCDA requires preference and value dependent judgements, so the quality and appropriateness of the criteria, preferences, and the model assumptions are relevant issues within the decision analysis process. These criteria can then be validated by experts to see if they are relevant and complete; if the assigned weights are reasonable; if the assumptions are appropriate; and if the resulting ranking is consistent with domain knowledge. Lawshe [102] suggested a quantitative method for determining whether an item is essential, the Content Validity Ratio (CVR). It can also be applied in an MCDM context when experts decide whether or not a criterion has to be included or not in the decision problem [103]. The CVR is calculated as:

$$CVR = \frac{n_e - N/2}{N/2} \quad (12)$$

where n_e is the number of experts identifying an item as essential and N is the total number of experts. Higher CVR values indicate stronger expert agreement regarding the relevance of an item. Thus, expert validation can be used to examine criterion relevance, rather than simply accepting criteria selected by the researcher. The specific acceptance threshold, however, depends on the number of experts and the adopted validation procedure.

Elangavan and Sundaravel [104] differentiate between content validity and face validity. Content validity investigates if the intended concept is covered with the proper range of items and face validity investigates whether the model or items seem relevant and sensible to knowledgeable reviewers. These ideas can subsequently be applied in MCDM to evaluate if the chosen criteria and model architecture are suitable representations of the decision problem. Dalkey and Helmer [105] developed the Delphi technique as a systematic method for asking and combining experts' opinions in a series of meetings. In MCDM, Delphi procedures can be employed to elicit expert consensus on the selection of the criteria, the definition of the criteria, the weights, the performance measurement or the plausibility of the final ranking. One common approach is to gather individual judgments, give feedback in a controlled manner, and re-evaluate the same judgment again and again until a certain consensus or stopping criterion is met. Delphi procedures remain popular in contemporary MCDA applications for the construction of experts' consensus in the absence of sufficient and consistent empirical data [106].

The expert validation can also be used after the alternative ranks and criterion weights have been obtained from MCDM model. Vinogradova *et al.*, [107] proposed a re-calculation and combination of the weights of two sets of experts, which would increase the consistency and accuracy of the judgments, but this would be effective only if the expert judgments are reliable and the methods of the weights' calculation are appropriate. Experts might be invited to comment on the reasonableness of the relative importance of criteria and the resulting ranking, as well as the consistency with the known domain knowledge or professional experience. The agreement of the experts in the ranking, however, should be understood as being expert or face valid, not as an objective demonstration that

the alternative selected is optimal. There are more than a few influences that can affect expert judgment, such as experience, knowledge, preference, group dynamics, and even confirmation bias. The choice and makeup of the expert panel will be subject to judgments and therefore should be reported. The validation procedure should contain a description of the background of the experts, the number of experts, the selection criteria, the scale on which the experts' judgments were made, the number of Delphi rounds (if any), consensus level, how the disagreement is treated, and how the experts' judgments are aggregated.

For instance, Pevcin and Debelak [108] validated the decision-support structure of the AHP-based framework by involving experts in the validation process, highlighting the significance of including domain knowledge explicitly in the evaluation of the criteria and the decision-support structure. Thus, there are two main questions addressed in the process of expert-based validation: "Are the criteria, the assumptions, the weights, and the resulting decision consistent with informed domain knowledge?" It does not demonstrate statistical significance, ranking stability, probabilistic robustness, or external validity; it only demonstrates evidence of content relevance, practical plausibility, and expert agreement. As such, expert validation is recommended in addition to statistical, sensitivity, robustness, and external validation especially if an objective ground truth is not available, as stated in Table 6.

Table 6
Expert Validation Checks

Expert-validation component	What is evaluated?	Primary question
Criterion validation	Relevance and completeness of criteria	Are the selected criteria appropriate?
Content validity	Coverage of the decision concept	Do the criteria adequately represent the decision problem?
Face validity	Apparent appropriateness of the model	Does the model make practical sense to experts?
Weight validation	Reasonableness of criterion importance	Are the assigned weights defensible?
Delphi consensus	Agreement among experts	Can experts reach an acceptable level of consensus?
Ranking validation	Practical plausibility of final ranking	Does the ranking agree with domain knowledge?

6.6 Benchmark and External Validation

External validation is a comparison of the outcomes of the use of the MCDM with other evidence such as the optimal solution, ranking, experimental observation, historical performance, industry standards, benchmark models, and real-world outcomes [109, 110]. The benchmarking and ranking of multiclass classification models using multiple and conflicting evaluation criteria were discussed by Alsalem *et al.*, [111] using a combination of BWM and VIKOR. Results of the study showed that almost identical rankings were obtained in the two decision-making methods, internal and external, and that significant differences were found among the group scores as validated by objective measures. Their study was limited to a small dataset and a particular group of classification models. Similarly, Zaidan *et al.*, [112] proposed an external evaluator based MCDM benchmarking method that incorporates external aggregation with AHP based preference weights and TOPSIS, to demonstrate that the external aggregation can facilitate the reliable comparison of alternatives under different evaluator preferences and only applied to a small example of digital-watermarking, thereby not generalizing to a broader context. To enhance the stability of the ranking, Tomar *et al.*, [113] developed a benchmark-driven MCDM model with fixed benchmark values and 0–100 scores, showing the effectiveness of other addition/removal tests and sensitivity analysis and studied rank-

reversal resistance and benchmark stability, but there is still a need for broader independent validation across a variety of MCDM applications. The intelligent supplier-benchmarking decision-support system, under interval uncertainty, was developed by Goswami *et al.*, [114] and the results were compared with the existing methods which gave comparative evidence for the validity of supplier efficiency rankings. Overall, the results of these studies show that external validation and benchmarking can improve the credibility and robustness of MCDM rankings, but independent datasets, different types of decision contexts, other weighting approaches and cross-method validation are still significant research challenges.

6.7 Simulation and Resampling Validation

The use of simulation and resampling validation techniques to verify the robustness, stability, predictive capability and uncertainty of the MCDM results is growing. Researchers can re-sample data, test different scenarios, or simulate changes to the criteria and weights to see if rankings change and if conclusions hold true if the decision environment changes [115].

Eloudi *et al.*, [116] tested the accuracy of MCDM, statistical and machine-learning ensemble models to gully-sensitivity mapping, and they concluded that MCDM-WoE-AHP and WoE-RF models performed well in predicting gully-sensitivity, with AUC values of 0.886 and 0.888, respectively. The study, however, was geographically limited to a semi-arid region, and was not fully validated because it was primarily prediction-oriented rather than in its entirety on the stability of the MCDM ranking under repeated resampling or under uncertainty. Raheja *et al.*, [117] proposed a hybrid approach of AHP-CODAS and applied the simulation and sensitivity analysis to analyze the ranking of polluted cities in India. The model was validated with AHP and TOPSIS with the help of Spearman correlation and scenario-based sensitivity analysis was conducted to investigate the impact of pollutant weights. However, the validation was still conducted for a small number of cities, pollutants and alternative MCDM methods and so the results were limited to those. Likewise, Kalantar *et al.*, [118] showed the benefit of resampling-based validation by employing cross-validation, bootstrap and optimism-bootstrap methods for forest-fire susceptibility models. The results indicated that resampling led to better predictive performance; BRT performed optimistically with AUC score of 0.91. It is worth noting that the study mainly validates the prediction models using machine learning methods, but not the stability of the weights of the criteria and the alternative rankings obtained by MCDM, which leaves a methodological gap for resampling validation of MCDM.

7. Methodological Advances in Sensitivity Analysis and Validation

The literature suggests there is a definite methodological shift, from being strictly a sensitivity test method, to a more general uncertainty-aware approach. The first set of methods involved changing the importance of the criteria; the latter set of methods also take into account the values of the decision matrix, alternatives, scenarios, methodological assumptions, and multiple uncertain inputs and interactions. This is the next step after fixed inputs and deterministic perturbations to probabilistic inputs. One-at-a-time procedures have been extended to so-called global sensitivity analysis that takes into account the variations of multiple input variables and their potential interaction. Probabilistic approaches also enable uncertain parameters to be represented by a probability distribution and Monte Carlo simulation enables the evaluation of multiple combinations of weights and performance values to be repeated.

So instead of reporting just one ranking, researchers can construct estimates for the probabilities of the other rank being the first. The effect of sampling variability on the outcome of MCDM is also discussed in the Bootstrap and resampling methods. The ranking stability and validation assessment has also been improved as a result of the last years developments. The focus has shifted over to other

elements of the decision uncertainty, such as rank probability, rank acceptability, rank volatility, stability indices and stochastic dominance, instead of only the “stable” vs. “unstable” dichotomy. Likewise, the more elementary and basic consistency checks have been expanded to a more complex multi-level validation of internal, comparative, statistical, expert and external validation. There are other ways of representing uncertain information, such as fuzzy or interval numbers, which don't need to be represented numerically.

8. Applications of Sensitivity Analysis and Validation in MCDM

Sensitivity and validation processes are more and more widely used in many areas of decision making. How they contribute to the uncertainty depends on the type of uncertainty and the external evidence available. The areas of engineering and industry in which MCDM applications are used are very broad. One of the various types of analysis available to an engineer or manufacturer in material selection, process selection, equipment evaluation, machining, supplier selection and technology assessment is the sensitivity analysis. This uncertainty in energy demand, environmental conditions, technological performance and even policy conditions is particularly useful in energy and renewable-energy applications for probabilistic sensitivity analysis and comparisons of different approaches. Their techniques are used in the selection of renewables, assessment of energy technologies, energy-management strategies and grid planning.

In the healthcare field, the application of the MCDM methods was observed for hospitals assessment, medical equipment selection, prioritization of treatment and allocation of resources and in the case of clinical and experimental uncertainty, statistical validation can be used as an alternative tool. In supply chain operations and logistics, the sensitivity analysis can be used to test the changes in supplier rank under changes in supplier cost, lead time, reliability, and sustainability rating, particularly in a dynamic environment or operations, using scenarios and probabilistic approaches. In the construction and infrastructure sector, MCDM can be used to select contractors, for material selection, site selection, infrastructure prioritization and sustainability assessment, and sensitivity analysis can be used to assess the impact of costs, safety and project duration, while validation can be done by engineering standards and regulatory requirements. In smart-cities and emerging applications uncertainty and ranking stability are analyzed with hybrid (deterministic, fuzzy and simulation) approaches, which are suitable for dynamic and heterogeneous data.

9. Comparative Discussion and Critical Analysis

This reviewed literature indicates that the sensitivity analysis and validation in MCDM has shifted from the relatively simple criterion weight perturbation to more complex techniques, such as using uncertain performance values, alternative sets, method selection, simulation, statistical analysis and external evidence. The comparative evidence leads to the introduction of a seven-dimensional taxonomy of sensitivity sources, which are expressed in the following equation: equation 13 and are discussed in Table 7 and 8.

$$S = \{W, X, A, N, G, P, M\} \quad (13)$$

where W denotes criterion weights, X the decision-matrix or performance values, A the set of alternatives, N the normalization procedure, G the aggregation procedure, P model-specific parameters and M the selected MCDM method. The selection of the appropriate analysis is done as indicated in Figure 10.

Table 7
Taxonomy of Sensitivity Sources in MCDM

Symbol	Sensitivity source	What is varied?	Typical perturbation	Main question
W	Weights	Criterion importance	$\pm\%$, OAT, simultaneous variation, probability distributions	Does changing criterion importance alter the ranking?
X	Decision matrix	Alternative performance values	$\pm\%$, random error, interval/noise	Does data uncertainty alter the ranking?
A	Alternatives	Alternative set	Addition, removal, replacement	Is the ranking stable to changes in alternatives?
N	Normalization	Normalization technique	Change normalization method	Is the result dependent on data transformation?
G	Aggregation	Aggregation/decision rule	Change aggregation operator	Is the ranking dependent on the aggregation mechanism?
P	Parameters	Method-specific parameters	Parameter variation	Is the result sensitive to model settings?
M	MCDM method	Decision-making method	TOPSIS, VIKOR, EDAS, COPRAS, etc.	Is the result method-dependent?

Table 8
Operational mapping of sensitivity sources to perturbation and validation measures

Source	Perturbation	Output	Recommended indicator	What it establishes
W	Weight variation	Scores/ranks	Rank change, Spearman ρ , Kendall τ , critical weight	Weight sensitivity/stability
X	Performance-value variation	Scores/ranks	Score variation, rank reversal, ρ/τ	Data sensitivity
A	Add/remove alternatives	Ranking	Rank reversal, rank-position frequency	Structural stability
N	Normalization methods	Scores/ranks	Rank correlation, ranking difference	Normalization dependence
G	Aggregation operators	Final scores/ranks	Rank correlation, rank reversal	Aggregation dependence
P	Parameter variation	Scores/ranks	Sensitivity index, rank change, variance	Parameter sensitivity
M	MCDM-method variation	Rankings	Spearman ρ , Kendall τ , rank deviation	Cross-method agreement

The complete framework can thus be illustrated as in Figure 11. This sequence helps to establish an operational connection between what is not known, how it is disturbed, how the decision changes, how stable and robust the resulting ranking is, and what evidence underpinning the final recommendation exists. Importantly, it does not presume a stable ranking is always correct, nor does it presume that if one method of MCDM agrees with another, they are correct. Instead, various types of evidence address various methodological questions: stability questions whether sensitive to changes in specified ways; robustness asks whether the results are robust to plausible uncertainty; comparative and statistical analyses ask whether the results are in agreement or disagreement; and external validation asks whether they are consistent with other evidence. This distinction offers a scientifically defensible way to evaluate the trustworthiness of the recommendations of MCDM.

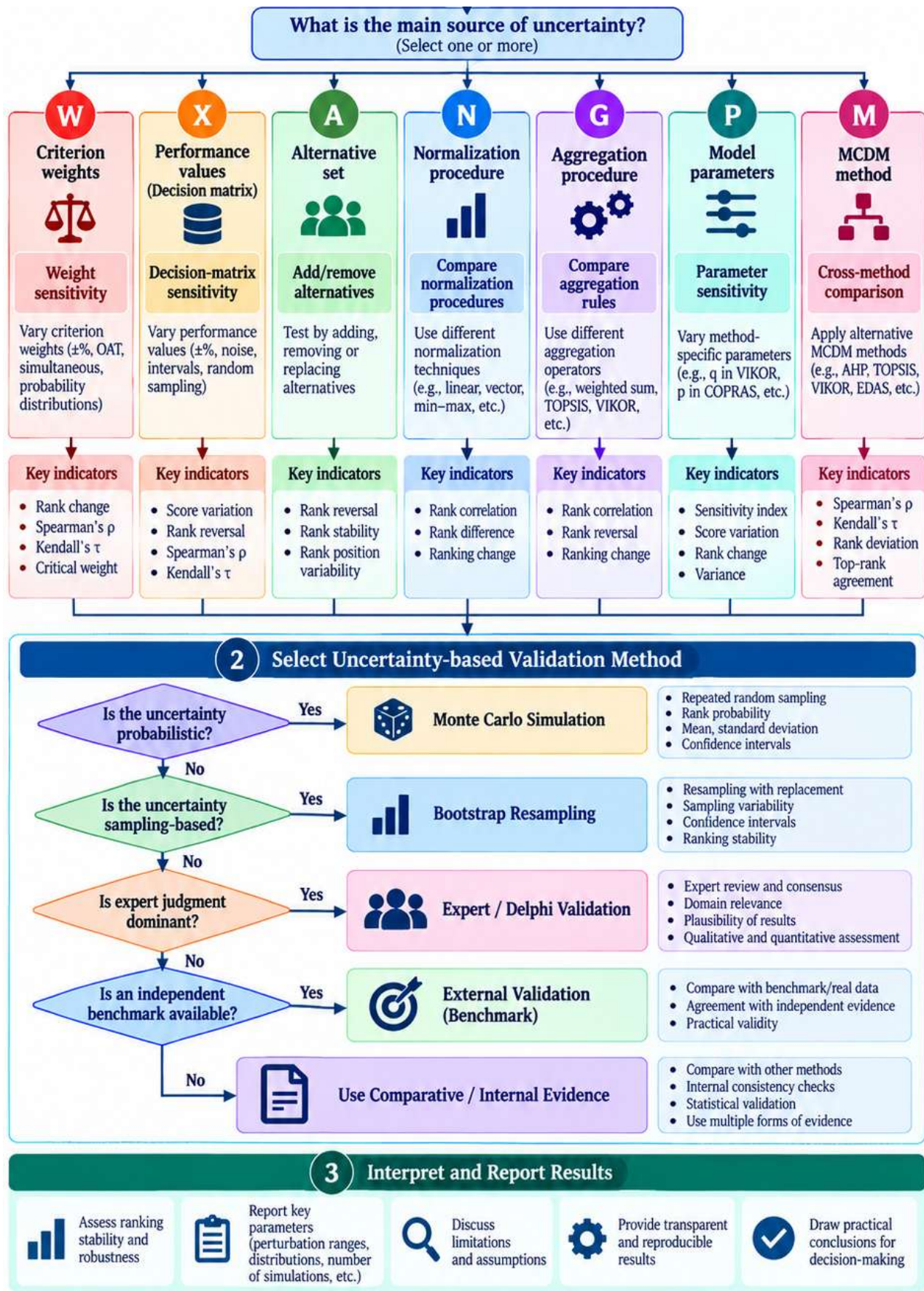


Fig. 10. Decision tree for selecting sensitivity and validation procedures in MCDM

Stability, robustness and external validity can conceptually be stated in the following equations:

$$\text{Stability} = f(\text{rank changes}, \rho, \tau, \text{rank frequency}) \quad (14)$$

$$\text{Robustness} = f(\text{rank probability}, \text{preferred alternative frequency}) \quad (15)$$

$$\text{External Validity} = f(\text{agreement with independent evidence}) \quad (16)$$



Fig. 11. Integrated Sensitivity–Validation Framework for Selecting Sensitivity, Stability, Robustness, and Validation Analyses in MCDM

10. Research Gaps and Future Research Directions

Based on the systematic and thematic analysis, it is found that there are several gaps in the sensitivity analysis and validation of MCDM as discussed in Table 9. The literature searched indicated that criterion-weight sensitivity is the most extensively discussed source of uncertainty while the uncertainty related to values in the decision matrix, alternative sets, normalisation process, aggregation rules, model-specific parameters and choice of MCDM methods is comparatively less systematically discussed. Furthermore, there is also significant variation in the range of perturbations, the assumptions made for the probability, the simulation parameters, the stability measure, and the validation method, which reduce the comparability and reproducibility of the studies. The literature also provides evidence that stability of ranking, methodological agreement, external validity, and robustness are often measured independently, and little evidence of their being synthesized in a common analytical framework. The results suggest the need for standardized and more multi-dimensional sensitivity-validation procedures instead of just one sensitivity indicator and validation method.

Table 9
Identified Literature Gaps and Future Research Directions

Identified gap	observation from the reviewed studies	Future research direction
Lack of standardized sensitivity procedures	Studies use different perturbation ranges, scenario numbers, probability distributions, simulation settings, indicators, and reporting practices.	Standardized sensitivity and validation protocols
Dominance of criterion-weight sensitivity	Weight variation is widely investigated, while decision-matrix values, alternatives, normalization, aggregation, parameters, and MCDM-method dependence receive comparatively less systematic attention.	Multidimensional sensitivity analysis
Heterogeneous representation of uncertain decision information	Existing studies employ deterministic ranges, intervals, probability distributions, fuzzy information, and stochastic approaches with different assumptions.	Probabilistic and Bayesian uncertainty modelling
Limited integration of simulation and resampling approaches	Monte Carlo, SMAA, bootstrap, and global/variance-based sensitivity methods are generally applied for different purposes and are not consistently integrated.	Advanced simulation and resampling frameworks
Predominantly static decision structures	Many sensitivity analyses evaluate a fixed set of weights, performance values, and alternatives rather than their evolution over time.	Dynamic MCDM sensitivity analysis
Increasing complexity of multidimensional sensitivity analysis	Simultaneous perturbation can generate a large number of scenarios and computationally demanding analyses.	AI-assisted sensitivity and validation
Limited use of independent evidence for validation	Cross-method agreement and statistical association are more readily applied than comparison with independent empirical or real-world evidence.	External validation and benchmark comparison

Table 9
Continued

Identified gap	observation from the reviewed studies	Future research direction
Inconsistent reporting and limited reproducibility	Perturbation ranges, distributions, simulation runs, preprocessing procedures, algorithms, and stability indicators are not consistently reported.	Reproducibility and transparent reporting standards
Separate treatment of stability, robustness, and validation evidence	Different forms of evidence are often reported independently, making it difficult to assess the overall credibility of an MCDM recommendation.	Integrated decision-credibility assessment

11. Conclusion

This study conducted a comprehensive review of methodological advances in sensitivity analysis and validation of MCDM methods by conducting a systematic literature review, bibliometric analysis

and thematic analysis. The results indicate a significant increase in research interest and the development of the sensitivity analysis from the classical approach of varying the criteria to the current approach of perturbing the decision matrix, scenario analysis, alternative-level analysis, probabilistic approach, Monte Carlo simulation, bootstrap method, fuzzy approach, interval approach, and hybrid approach.

The study also revealed that validation includes internal consistency assessment, robustness and stability analysis, cross-method comparison, statistical validation, expert evaluation, benchmarking, and external and empirical validation. But other uncertainties and external validation are relatively underdiscussed compared to criteria-weight sensitivity. Lack of standardization of sensitivity procedures, indicators, terminology and reporting practices also restrict comparability and reproducibility. From these results, the study suggests an integrated view that takes into account a combination of uncertainty, sensitivity, robustness, ranking stability and validation to make up the decision credibility. The future research directions include standardized procedures, probabilistic and simulation-based approaches, dynamic and AI based MCDM, better external validation and reproducible research.

Acknowledgement

We would like to offer our warmest thanks to our mentors, our colleagues and sources of research, whose support and help were of value to this research.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Sařabun, W., Wątróbski, J., & Shekhovtsov, A. (2020). Are MCDA methods benchmarkable? a comparative study of topos, vikor, copras, and promethee ii methods. *Symmetry*, 12(9), 1549. <https://doi.org/10.3390/sym12091549>
- [2] Ullah, K., Rehman, N., & Ali, A. (2026). Business-oriented stock market decision analysis using circular complex picture fuzzy sets and advanced MCDM based on the CRITIC–WASPAS method. *Journal of Contemporary Decision Science*, 2(1), 1-54. <https://doi.org/10.67334/cds2120268>
- [3] Boz, E., Çalik, A., Alshalfan, A., & Tirkolaee, E. B. (2026). Advancements of Multi-Criteria Decision-Making Approaches Integrated with Artificial Intelligence and Machine Learning in Sustainable Supply Chain Management. *Decision Making Advances*, 4(1), 257-289. <https://doi.org/10.31181/dma412026189>
- [4] Katrancı, A., Kundakcı, N., & Arman, K. (2026). Fuzzy SIWEC and Fuzzy RAWEC methods for sustainable waste disposal technology selection. *Spectrum of Operational Research*, 3(1), 87-102. <https://doi.org/10.31181/sor31202633>
- [5] Sarkar, A., & Goswami, S. S. (2027). A comprehensive review: The novel weighting methods for multi-criteria decision-making (MCDM). *Spectrum of Operational Research*, 1-26. <https://doi.org/10.31181/sor202781>
- [6] Guliyev, J., & Deveci, M. (2026). A Stochastic Approach to Criterion Weighting in Multi-Criteria Decision Making: Entropy-Based Monte Carlo Modeling. *Measurement: Digitalization*, 7, 100043. <https://doi.org/10.1016/j.measdig.2026.100043>
- [7] Yalçın, G. C., Özyürek, H., & Dişçi, C. (2026). Benchmarking the Financial Performance of the Corporate Governance Index under Borsa Istanbul Companies Using SITDE and MABAC Framework. *Journal of Contemporary Decision Science*, 2(1), 169-193. <https://doi.org/10.67334/cds21202614>
- [8] Sarkar, A., Goswami, S. S., & Gupta, K. K. (2027). Sensitivity Analysis and Validation in MCDM Methods: A Comprehensive Review with Advancements, Applications, and Future Directions. *Spectrum of Decision Making and Applications*, 1-14. <https://doi.org/10.31181/sdmap41202769>
- [9] Tahir, M., & Shahid, M. I. (2026). An intelligent industry 5.0 logistics decision system: circular supply chain management with fermatean neutrosophic hypersoft sets and machine learning. *Journal of Contemporary Decision Science*, 2(1), 260-286. <https://doi.org/10.67334/cds21202618>
- [10] Khan, H. U., & Ali, Y. (2026). Assessment framework for authentication algorithms in IoT-driven transportation systems and vehicle road corporation systems using hybrid MCDM methodology. *Case Studies on Transport Policy*, 101955. <https://doi.org/10.1016/j.cstp.2026.101955>

- [11] Lin, M., & Mi, G. (2026). A Multi-Criteria Evaluation Method for Interactive Interface Layouts in Smart Digital Museums Based on RANCOM–LOPCOW–TOPSIS. *Applied Sciences*, 16(15), 7747. <https://doi.org/10.3390/app16157747>
- [12] Sahoo, S. K., Choudhury, B. B., Dhal, P. R., & Hanspal, M. S. (2025). A comprehensive review of multi-criteria decision-making (MCDM) toward sustainable renewable energy development. *Spectrum of Operational Research*, 2(1), 309-325. <https://doi.org/10.31181/sor21202527>
- [13] Jameel, T., Yasin, Y., & Riaz, M. (2026). An integrated hybrid MCDM framework for renewable energy prioritization in sustainable development. *Spectrum of decision making and applications*, 3(1), 124-150. <https://doi.org/10.31181/sdmap31202640>
- [14] Saaty, T. L. (1994). *Fundamentals of decision making and priority theory with the analytic hierarchy process*. RWS publications.
- [15] Triantaphyllou, E., & Sánchez, A. (1997). A sensitivity analysis approach for some deterministic multi-criteria decision-making methods. *Decision sciences*, 28(1), 151-194. <https://doi.org/10.1111/j.1540-5915.1997.tb01306.x>
- [16] Ishizaka, A., & Nemery, P. (2013). *Multi-criteria decision analysis: methods and software*. John Wiley & Sons.
- [17] Więckowski, J., & Sałabun, W. (2023). Sensitivity analysis approaches in multi-criteria decision analysis: A systematic review. *Applied Soft Computing*, 148, 110915. <https://doi.org/10.1016/j.asoc.2023.110915>
- [18] Demir, G., Chatterjee, P., & Pamucar, D. (2024). Sensitivity analysis in multi-criteria decision making: A state-of-the-art research perspective using bibliometric analysis. *Expert systems with applications*, 237, 121660. <https://doi.org/10.1016/j.eswa.2023.121660>
- [19] Sahoo, S. K., & Goswami, S. S. (2023). A comprehensive review of multiple criteria decision-making (MCDM) methods: advancements, applications, and future directions. *Decision making advances*, 1(1), 25-48. <https://doi.org/10.31181/dma1120237>
- [20] Markatos, D. N., Malefaki, S., & Pantelakis, S. G. (2023). Sensitivity analysis of a hybrid MCDM model for sustainability assessment—an example from the aviation industry. *Aerospace*, 10(4), 385. <https://doi.org/10.3390/aerospace10040385>
- [21] Wolters, W. T. M., & Mareschal, B. (1995). Novel types of sensitivity analysis for additive MCDM methods. *European Journal of Operational Research*, 81(2), 281-290. [https://doi.org/10.1016/0377-2217\(93\)E0343-V](https://doi.org/10.1016/0377-2217(93)E0343-V)
- [22] Yazdani, M., Zavadskas, E. K., Ignatius, J., & Abad, M. D. (2016). Sensitivity analysis in MADM methods: Application of material selection. *Engineering Economics*, 27(4), 382-391. <https://doi.org/10.5755/j01.ee.27.4.14005>
- [23] Boukrouh, I., Tayalati, F., & Azmani, A. (2024). A comprehensive framework for supplier selection: using subjective, objective, and hybrid multi-criteria decision-making techniques with sensitivity analysis. *IEEE Access*, 12, 145550-145569. <https://doi.org/10.1109/ACCESS.2024.3462348>
- [24] Ayyildiz, E. (2023). Interval valued intuitionistic fuzzy analytic hierarchy process-based green supply chain resilience evaluation methodology in post COVID-19 era. *Environmental Science and Pollution Research*, 30(15), 42476-42494. <https://doi.org/10.1007/s11356-021-16972-y>
- [25] Więckowski, J., Wątróbski, J., & Sałabun, W. (2024). Inaccuracies in expert judgment: Comparative analysis of RANCOM and AHP methods in housing location selection problem. *IEEE Access*, 12, 142083-142100. <https://doi.org/10.1109/ACCESS.2024.3443451>. <https://doi.org/10.1109/ACCESS.2024.3443451>
- [26] Brunelli, M., Canal, L., & Fedrizzi, M. (2013). Inconsistency indices for pairwise comparison matrices: a numerical study. *Annals of Operations Research*, 211(1), 493-509. <https://doi.org/10.1007/s10479-013-1329-0>
- [27] Ramírez-Olivares, E., Robles, J. A., & Miranda, T. B. (2026). Prioritization of Critical Sustainability Factors in Medium-Scale Mining: The Coquimbo Region Case. *Sustainability*, 18(14), 7314. <https://doi.org/10.3390/su18147314>
- [28] Memis, S., & Yildiz, R. O. (2026). Green human resource management: a hybrid soft decision-making approach. *Management Decision*, 1-58. <https://doi.org/10.1108/MD-06-2025-1700>
- [29] Sahoo, S. K., Choudhury, B. B., Dhal, P. R., Sahu, S., Majhi, S., & Dhar, I. (2026). Artificial Intelligence in Open Innovation and Project Management: A Systematic Literature Review of Technologies, Applications, and Future Research Directions. *Applied Research Advances*, 2(1), 199-224. <https://doi.org/10.65069/ara21202629>
- [30] Kim, A. H., Feingold, C. L., Kotlier, J. L., Lin, E. H., Fathi, A., Agarwalla, A., & Liu, J. N. (2026). Appraisal of the Presence of Spin in Abstracts of Systematic Reviews and Meta-analyses Regarding Free Bone Block Procedures for Glenohumeral Instability. *The American Journal of Sports Medicine*, 54(1), 199-207. <https://doi.org/10.1177/036354652513380>
- [31] Lyu, P., Liu, X., & Yao, T. (2026). A bibliometric analysis of literature on bibliometrics in recent half-century. *Journal of Information Science*, 52(2), 324-344. <https://doi.org/10.1177/01655515231191233>

- [32] Özdemir, Ö. (2026). The evolution, intellectual structure, and developmental trends of global mental health nursing research: A bibliometric analysis (2015–2025). *Archives of Psychiatric Nursing*, 152055. <https://doi.org/10.1016/j.apnu.2026.152055>
- [33] Özekenci, E. K., Onat, K. T., & Pamucar, D. (2026). A comprehensive bibliometric analysis of objective weighting methods in multi-criterion decision-making. *Management Science Advances*, 3(1), 246-265. <https://doi.org/10.31181/msa31202645>
- [34] Bhattacharjee, S. (2026). A Novel Multi Criteria Decision Framework for Modeling Trust and Loyalty in Social Media Marketing Ecosystems. *Applied Decision Analytics*, 2(1), 352-370. <https://doi.org/10.66972/ada21202629>
- [35] Wang, T., He, J., Yan, W., Chen, K., Zhang, X., Zhang, N., & Liang, W. (2026). Artificial intelligence in primary health care: A bibliometric analysis of publications from 2015 to 2024. *Digital Health*, 12, 20552076251411631. <https://doi.org/10.1177/20552076251411631>
- [36] Ivezić, I., Hrustek, N. Ž., & Su, Y. (2026). Analytical DEA-MACBETH-OPA Approach for Selecting the Location of a Student Apartment. *Applied Decision Analytics*, 2(1), 371-389. <https://doi.org/10.66972/ada21202630>
- [37] Ravichandran, K. S. (2025). Variance Measure in Decision-Making Process: An Interpretation. *Management Science Advances*, 2(1), 282-287. <https://doi.org/10.31181/msa21202519>
- [38] Abdel-Basset, M., Mohamed, R., Sallam, K., & Elhoseny, M. (2020). A novel decision-making model for sustainable supply chain finance under uncertainty environment. *Journal of Cleaner Production*, 269, 122324. <https://doi.org/10.1016/j.jclepro.2020.122324>
- [39] Kitila, L. G., Duc, E., Durieux, S., & Taddese, G. A. (2026). An integrated fuzzy AHP-fuzzy TOPSIS approach for multi-criteria decision-making framework in sustainable manufacturing process selection. *The International Journal of Advanced Manufacturing Technology*, 142(1-2), 559-583. <https://doi.org/10.1007/s00170-025-16881-1>
- [40] Ranjgar, B., Niccolai, A., & Leva, S. (2026). Solar PV site selection using GIS-MCDM: A comprehensive approach with spatially explicit collinearity and sensitivity analyses. *Renewable Energy*, 270, 125873. <https://doi.org/10.1016/j.renene.2026.125873>
- [41] Anu, P. S., & Arulselvam, A. (2026). An Enhanced Multi-Criteria Decision-Making Model Using Prospect Theory Under Interval-Valued Intuitionistic Quadri-Partitioned Neutrosophic Soft Set Environment. *Journal of the Indonesian Mathematical Society*, 32(2), 2226-2226. <https://doi.org/10.22342/jims.v32i2.2226>
- [42] Mukhametzhanov, I. (2021). Specific character of objective methods for determining weights of criteria in MCDM problems: Entropy, CRITIC and SD. *Decision Making: Applications in Management and Engineering*, 4(2), 76-105. <https://doi.org/10.31181/dmame210402076i>
- [43] O'Shea, R., Deeney, P., Triantaphyllou, E., Díaz-Balteiro, L., & Tarim, S. A. (2026). Weight stability intervals for multi-criteria decision analysis using the weighted sum model. *Expert Systems with Applications*, 296, 128460. <https://doi.org/10.1016/j.eswa.2025.128460>
- [44] Mazurek, J., & Strzałka, D. (2022). On the Monte Carlo weights in multiple criteria decision analysis. *PLoS One*, 17(10), e0268950. <https://doi.org/10.1371/journal.pone.0268950>
- [45] Cajot, S., Schüler, N., Peter, M., Koch, A., & Maréchal, F. (2019). Interactive optimization with parallel coordinates: exploring multidimensional spaces for decision support. *Frontiers in ICT*, 5, 32. <https://doi.org/10.3389/fict.2018.00032>
- [46] Abbasi, S. (2026). A robust multi-criteria decision framework for sustainable recycling of NCM lithium-ion battery cathodes under uncertainty and policy scenarios. *Results in Engineering*, 30, 110015. <https://doi.org/10.1016/j.rineng.2026.110015>
- [47] Maden, A. (2025). Low-Carbon Supplier Selection Using a Scenario-Based Integrated FCM-EDAS Approach: A Textile Company Case Study: A. Maden: Low-Carbon Supplier Selection Using a Scenario-Based Integrated FCM-EDAS Approach. *International Journal of Fuzzy Systems*, 28(5), 1816-1835. <https://doi.org/10.1007/s40815-025-02049-2>
- [48] Stević, Ž., Pamučar, D., Puška, A., & Chatterjee, P. (2020). Sustainable supplier selection in healthcare industries using a new MCDM method: Measurement of alternatives and ranking according to COMpromise solution (MARCOS). *Computers & industrial engineering*, 140, 106231. <https://doi.org/10.1016/j.cie.2019.106231>
- [49] Bhattacharya, R. K., Chatterjee, N. D., & Das, K. (2020). Sub-basin prioritization for assessment of soil erosion susceptibility in Kangsabati, a plateau basin: a comparison between MCDM and SWAT models. *Science of the Total Environment*, 734, 139474. <https://doi.org/10.1016/j.scitotenv.2020.139474>
- [50] Mohammadi, M., & Rezaei, J. (2020). Bayesian best-worst method: A probabilistic group decision making model. *Omega*, 96, 102075. <https://doi.org/10.1016/j.omega.2019.06.001>
- [51] Kala, Z. (2020). Sensitivity analysis in probabilistic structural design: A comparison of selected techniques. *Sustainability*, 12(11), 4788. <https://doi.org/10.3390/su12114788>

- [52] Sun, C., Li, S., & Deng, Y. (2020). Determining weights in multi-criteria decision making based on negation of probability distribution under uncertain environment. *Mathematics*, 8(2), 191. <https://doi.org/10.3390/math8020191>
- [53] Feizizadeh, B., & Blaschke, T. (2014). An uncertainty and sensitivity analysis approach for GIS-based multicriteria landslide susceptibility mapping. *International Journal of Geographical Information Science*, 28(3), 610-638. <https://doi.org/10.1080/13658816.2013.869821>
- [54] Mukhametzyanov, I., & Pamucar, D. (2018). A sensitivity analysis in MCDM problems: A statistical approach. *Decision making: applications in management and engineering*, 1(2), 51-80. <https://doi.org/10.31181/dmame1802050m>
- [55] Gaona, À., Guisasola, A., & Baeza, J. A. (2025). An integrated TOPSIS framework with full-range weight sensitivity analysis for robust decision analysis. *Decision Analytics Journal*, 17, 100642. <https://doi.org/10.1016/j.dajour.2025.100642>
- [56] Fengmin, L., Dengfeng, W., Ying, X., Zihao, M., & Jing, C. (2025). Classification, selection of MCDM methods and robust decision-making in multidisciplinary design optimization of automotive structures. *Structural and Multidisciplinary Optimization*, 68(11), 232. <https://doi.org/10.1007/s00158-025-04150-4>
- [57] Pala, O. (2026). A Framework to Select Multi Criteria Decision Making (MCDM) Methods for Either Ranking or Choosing Problem. *International Journal of Information Technology & Decision Making*, 1-30. <https://doi.org/10.1142/S0219622026500471>
- [58] Nazarov, A. (2025). Development of a decision support methodology for optimizing ROI in project management. *Technology audit and production reserves*, 2(2(82)), 58-65. <https://doi.org/10.15587/2706-5448.2025.326385>
- [59] Paradowski, B. (2025). Towards robust objective decision-making: sensitivity analysis and comparative evaluation of the EORS approach. *Procedia Computer Science*, 270, 6223-6232. <https://doi.org/10.1016/j.procs.2025.10.092>
- [60] Paradowski, B., Więckowski, J., & Sałabun, W. (2024). PySensMCDA: A novel tool for sensitivity analysis in multi-criteria problems. *SoftwareX*, 27, 101746. <https://doi.org/10.1016/j.softx.2024.101746>
- [61] Hassan, I., Alhamrouni, I., & Azhan, N. H. (2023). A CRITIC–TOPSIS multi-criteria decision-making approach for optimum site selection for solar PV farm. *Energies*, 16(10), 4245. <https://doi.org/10.3390/en16104245>
- [62] Ali, J., Popa, I. L., & Al-Kenani, A. N. (2026). Entropy-Based Multi-attribute Decision Analysis with Hamacher Operators for Water Resource Management Challenges. *International Journal of Fuzzy Systems*, 1-22. <https://doi.org/10.1007/s40815-026-02265-4>
- [63] Mosadeghi, R., Warnken, J., Tomlinson, R., & Mirfenderesk, H. (2013). Uncertainty analysis in the application of multi-criteria decision-making methods in Australian strategic environmental decisions. *Journal of Environmental Planning and Management*, 56(8), 1097-1124. <https://doi.org/10.1080/09640568.2012.717886>
- [64] Liu, X., & Liu, Y. (2024). Sensitivity analysis of the parameters for preference functions and rank reversal analysis in the PROMETHEE II method. *Omega*, 128, 103116. <https://doi.org/10.1016/j.omega.2024.103116>
- [65] Nabavi, S. R., Wang, Z., & Rangaiah, G. P. (2023). Sensitivity analysis of multi-criteria decision-making methods for engineering applications. *Industrial & Engineering Chemistry Research*, 62(17), 6707-6722. <https://doi.org/10.1021/acs.iecr.2c04270>
- [66] Li, Y., He, X., Martinez, L., Zhang, J., Wang, D., & Liu, X. A. (2024). Comparative analysis of three categories of multi-criteria decision-making methods. *Expert Systems with Applications*, 238, 121824. <https://doi.org/10.1016/j.eswa.2023.121824>
- [67] Więckowski, J., Białoń, P., & Sałabun, W. (2024). Comparative sensitivity analysis of single and multiple modifications in multi-criteria decision analysis. *Procedia Computer Science*, 246, 91-102. <https://doi.org/10.1016/j.procs.2024.09.231>
- [68] Bonakdari, H., Khazaei, M., Gholian-Jouybari, F., & Hajiaghahi-Keshteli, M. (2025). Sustainable wastewater-to-hydrogen systems in Canada: A scenario-based decision support analysis. *Renewable Energy*, 256, 124308. <https://doi.org/10.1016/j.renene.2025.124308>
- [69] Azzini, I., & Munda, G. (2025). Sensitivity and Robustness Analyses in Social Multi-Criteria Evaluation of Public Policies. *Journal of Multi-Criteria Decision Analysis*, 32(1), e70006. <https://doi.org/10.1002/mcda.70006>
- [70] Purwanto, C., Ats-Tsauri, M. I., Syahliantina, A., Azizah, G. F., & Pratiwi, C. D. (2026). A Multi-Criteria Decision Making for Automated Storage System Selection in the Warehousing Industry Using AHP-TOPSIS. *Golden Ratio of Data in Summary*, 6(1), 48-65. <https://doi.org/10.52970/grdis.v6i1.1929>
- [71] Özkurt, P. (2026). A scenario-robust intuitionistic fuzzy AHP–TOPSIS model for sustainable healthcare waste treatment selection: Evidence from Türkiye. *Sustainability*, 18(3), 1167. <https://doi.org/10.3390/su18031167>
- [72] Hashemkhani Zolfani, S., & Nemati, A. (2026). An Assessment of Indifference Threshold Values to Achieve Full Objective Indifference Threshold-Based Attribute Ratio Analysis. *Mathematics*, 14(2), 235. <https://doi.org/10.3390/math14020235>

- [73] Maliene, V., Dixon-Gough, R., & Malys, N. (2018). Dispersion of relative importance values contributes to the ranking uncertainty: Sensitivity analysis of Multiple Criteria Decision-Making methods. *Applied Soft Computing*, 67, 286-298. <https://doi.org/10.1016/j.asoc.2018.03.003>
- [74] Więckowski, J., & Sałabun, W. (2025). Comparative sensitivity analysis in composite material selection: Evaluating OAT and COMSAM methods in multi-criteria decision-making. *Spectrum of mechanical engineering and operational research*, 2(1), 1-12. <https://doi.org/10.31181/smeor21202524>
- [75] Kassara, W. A., Hasan, B. F., & Zahi, B. R. (2026). A hybrid MCDM perspective on model selection for X-ray image classification: comparative insights from TOPSIS, VIKOR, and EDAS. *OPSEARCH*, 1-32. <https://doi.org/10.1007/s12597-026-01189-9>
- [76] Şengül, D., & Karahan, F. K. (2026). A Proportional Intuitionistic Fuzzy AHP–EDAS Framework for Symmetric Risk Assessment in Automotive Assembly Lines: A 12-Failure-Mode PFMEA Study. *Symmetry*, 18(7), 1115. <https://doi.org/10.3390/sym18071115>
- [77] Guan, L., Abbasi, A., & Ryan, M. J. (2021). A simulation-based risk interdependency network model for project risk assessment. *Decision Support Systems*, 148, 113602. <https://doi.org/10.1016/j.dss.2021.113602>
- [78] Da Silva, L. B. L., Alencar, M. H., & de Almeida, A. T. (2022). Exploring global sensitivity analysis on a risk-based MCDM/A model to support urban adaptation policies against floods. *International Journal of Disaster Risk Reduction*, 73, 102898. <https://doi.org/10.1016/j.ijdrr.2022.102898>
- [79] Dunke, F., & Nickel, S. (2026). Simulation-based multi-criteria decision making: an interactive method with a case study on infectious disease epidemics. *Annals of Operations Research*, 360(1), 67-96. <https://doi.org/10.1007/s10479-021-04321-8>
- [80] Baležentis, T., & Streimikiene, D. (2017). Multi-criteria ranking of energy generation scenarios with Monte Carlo simulation. *Applied energy*, 185, 862-871. <https://doi.org/10.1016/j.apenergy.2016.10.085>
- [81] Ab Rahman, M. H., Mahmud, J., Jumaidin, R., & Rahman, S. M. A. (2025). Integrated CRITIC-TOPSIS and Monte Carlo sensitivity analysis for optimal various natural fibre selection in sustainable building insulation composites to support the Sustainable Development Goals (SDGs). *ASEAN Journal of Science and Engineering*, 5(2), 231-260. <https://doi.org/10.17509/ajse.v5i2.85614>
- [82] Valenzuela, U. M. (2026). An Evaluation of Production Plans Under Uncertainty in the Automotive Industry Using EDAS and Monte Carlo Simulation. *Engineering, Technology & Applied Science Research*, 16(2), 34517-34524. <https://doi.org/10.48084/etasr.17868>
- [83] Alvarez, D. A. (2009). Reduction of uncertainty using sensitivity analysis methods for infinite random sets of indexable type. *International journal of approximate reasoning*, 50(5), 750-762. <https://doi.org/10.1016/j.ijar.2009.02.002>
- [84] Kaspar, K., & Kaliyaperumal, P. (2026). Single-valued neutrosophic sets for enhanced evaluation of IT roadmaps: a hybrid neutrosophic multi-criteria decision-making (MCDM) model for the IT industry. *International Journal of System Assurance Engineering and Management*, 17(1), 273-285. <https://doi.org/10.1007/s13198-025-02969-0>
- [85] Lebepe, P., & Mathaba, T. N. (2025). Enhancing energy resilience in enterprises: a multi-criteria approach. *Sustainable Energy Research*, 12(1), 7. <https://doi.org/10.1186/s40807-025-00148-0>
- [86] Ibrahim, S., Boukari, S., & Alaramma, K. S. (2026). An Improved Fuzzy AHP–TOPSIS Framework for Functional Requirements Selection and Rank-Reversal Mitigation. *FUDMA Journal of Sciences*, 10(15), 61-74. <https://doi.org/10.33003/fjs-2026-1015-5617>
- [87] Ayyildiz, E., ERDOĞAN, M., & Gul, M. (2025). An integrated intuitionistic fuzzy framework for emergency response center location planning in industrial areas. *International Journal of Critical Infrastructure Protection*, 52, 100813. <https://doi.org/10.1016/j.ijcip.2025.100813>
- [88] Wu, J. Z., & Tiao, P. J. (2018). A validation scheme for intelligent and effective multiple criteria decision-making. *Applied Soft Computing*, 68, 866-872. <https://doi.org/10.1016/j.asoc.2017.04.054>
- [89] Baydaş, M., & Pamučar, D. (2022). Determining objective characteristics of MCDM methods under uncertainty: an exploration study with financial data. *Mathematics*, 10(7), 1115. <https://doi.org/10.3390/math10071115>
- [90] Saaty, T. L. (1990). How to make a decision: the analytic hierarchy process. *European journal of operational research*, 48(1), 9-26. [https://doi.org/10.1016/0377-2217\(90\)90057-I](https://doi.org/10.1016/0377-2217(90)90057-I)
- [91] Cinelli, M., Kadziński, M., Gonzalez, M., & Słowiński, R. (2020). How to support the application of multiple criteria decision analysis? Let us start with a comprehensive taxonomy. *Omega*, 96, 102261. <https://doi.org/10.1016/j.omega.2020.102261>
- [92] Ganji, A., Maier, H. R., & Dandy, G. C. (2016). A modified Sobol' sensitivity analysis method for decision-making in environmental problems. *Environmental Modelling & Software*, 75, 15-27. <https://doi.org/10.1016/j.envsoft.2015.10.001>

- [93] Więckowski, J., Kizielewicz, B., Paradowski, B., & Sałabun, W. (2025). A robust framework for benchmarking MCDA methods: Sensitivity, stability, and method selection decision support. *Applied Soft Computing*, 187, 114335. <https://doi.org/10.1016/j.asoc.2025.114335>
- [94] Duško, T., & Bozanic, D. (2023). Optimizing military decision-making: Application of the FUCOM–EWAA–COPRAS–G MCDM model. *Acadlore Trans. Appl Math. Stat*, 1(3), 148-160. <https://doi.org/10.56578/atams010303>.
- [95] Belton, V., & Stewart, T. (2012). Multiple criteria decision analysis: an integrated approach. Springer Science & Business Media. <https://doi.org/10.1007/978-1-4615-1495-4>
- [96] Baydaş, M., Eren, T., Stević, Ž., Starčević, V., & Parlakkaya, R. (2023). Proposal for an objective binary benchmarking framework that validates each other for comparing MCDM methods through data analytics. *PeerJ Computer Science*, 9, e1350. <https://doi.org/10.7717/peerj-cs.1350>
- [97] Spearman, C. (1987). The proof and measurement of association between two things. *The American journal of psychology*, 100(3/4), 441-471. <https://doi.org/10.2307/1422689>
- [98] Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, 30(1-2), 81-93. <https://doi.org/10.1093/biomet/30.1-2.81>
- [99] Emerson, R. W. (2015). Causation and Pearson's correlation coefficient. *Journal of visual impairment & blindness*, 109(3), 242-244. <https://doi.org/10.1177/0145482X151090031>
- [100] Osmond, R., Mollahoseini, Z., Singh, J., Gautam, A., Seethaler, R., Golovin, K., & Milani, A. S. (2021). A group multicriteria decision making with ANOVA to select optimum parameters of drilling flax fibre composites: a case study. *Composites Part C: Open Access*, 5, 100156. <https://doi.org/10.1016/j.jcomc.2021.100156>
- [101] Neyman, J. (1937). Outline of a theory of statistical estimation based on the classical theory of probability. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 236(767), 333-380. <https://www.jstor.org/stable/91337>
- [102] Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel psychology*, 28(4), 563. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- [103] Mombeni, S., Darestani, S. A., & Shemami, N. H. (2023). Evaluation of the effective factors of unloading and loading goods in ports and the impact on the environment using the MCDM method. *Environmental Science and Pollution Research*, 30(6), 14873-14883. <https://doi.org/10.1007/s11356-022-23219-x>
- [104] Elangovan, N., & Sundaravel, E. (2021). Method of preparing a document for survey instrument validation by experts. *MethodsX*, 8, 101326. <https://doi.org/10.1016/j.mex.2021.101326>
- [105] Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management science*, 9(3), 458-467. <https://doi.org/10.1287/mnsc.9.3.458>
- [106] Toufighi, S. P., Tafazoli, R., Habibi, S., & Vang, J. (2026). A Delphi-informed fuzzy multi-criteria decision framework for prioritizing sustainable development goals in industrial strategy: An application to the paint and coatings sector. *Journal of Industrial Information Integration*, 51, 101091. <https://doi.org/101091.10.1016/j.jii.2026.101091>
- [107] Vinogradova, I., Podvezko, V., & Zavadskas, E. K. (2018). The recalculation of the weights of criteria in MCDM methods using the Bayes approach. *Symmetry*, 10(6), 205. <https://doi.org/10.3390/sym10060205>
- [108] Pevcin, P., & Debelak, K. (2026). Validating an AHP-based decision-support framework for AI adoption in public administration. *Business, Management and Economics Engineering*, 24(2), 387-406. <https://doi.org/10.3846/bmee.2026.27100>
- [109] Héberger, K. (2023). Selection of optimal validation methods for quantitative structure–activity relationships and applicability domain. *SAR and QSAR in Environmental Research*, 34(5), 415-434. <https://doi.org/10.1080/1062936X.2023.2214871>
- [110] Singh, B., Grover, S., & Singh, V. (2017). An empirical study of benchmarking evaluation using MCDM in service industries. *Managerial Auditing Journal*, 32(2), 111-147. <https://doi.org/10.1108/MAJ-11-2015-1274>
- [111] Alsalem, M. A., Zaidan, A. A., Zaidan, B. B., Albahri, O. S., Alamoodi, A. H., Albahri, A. S., Mohsin A. S., & Mohammed, K. I. (2019). Multiclass benchmarking framework for automated acute Leukaemia detection and classification based on BWM and group-VIKOR. *Journal of medical systems*, 43(7), 212. <https://doi.org/10.1007/s10916-019-1338-x>
- [112] Zaidan, B. B., Zaidan, A. A., Karim, H. A., & Ahmad, N. N. (2017). A new digital watermarking evaluation and benchmarking methodology using an external group of evaluators and multi-criteria analysis based on ‘large-scale data’. *Software: Practice and Experience*, 47(10), 1365-1392. <https://doi.org/10.1002/spe.2465>
- [113] Tomar, S., Jaiswal, A., Jaiswal, A. K., & Singh, L. K. (2026). Design and implementation of a benchmark-driven MCDM framework for eliminating rank reversal in decision-making. *International Journal of System Assurance Engineering and Management*, 1-15. <https://doi.org/10.1007/s13198-026-03411-9>
- [114] Goswami, M., Daultani, Y., Chan, F. T., & Pratap, S. (2022). Assessing the impact of supplier benchmarking in manufacturing value chains: an Intelligent decision support system for original equipment manufacturers.

- International Journal of Production Research, 60(24), 7411-7435.
<https://doi.org/10.1080/00207543.2022.2075811>
- [115] Kumar, P., Choudhary, M. P., & Mathur, A. K. (2026). Systematic review of conceptual and methodological frameworks for sustainable landfill site selection using multi-criteria decision-making techniques. *Waste Management & Research*, 0734242X261435305. <https://doi.org/10.1177/0734242X261435305>
- [116] Eloudi, H., Reddad, H., Hssaisoune, M., Estrany, J., Krimissa, S., Elaloui, A., Namous, M., Ouatiki, H., Aboutaib, F., Ouayah, M., Jadoud, M., Edahbi, M., & Bouchaou, L. (2022). Assessing the performance of MCDM, statistical and machine learning ensemble models for gully sensitivity mapping in a semi-arid context. *Geocarto International*, 37(27), 17435-17464. <https://doi.org/10.1080/10106049.2022.2129818>
- [117] Raheja, S., Obaidat, M. S., Kumar, M., Sadoun, B., & Bhushan, S. (2022). A hybrid MCDM framework and simulation analysis for the assessment of worst polluted cities. *Simulation Modelling Practice and Theory*, 118, 102540. <https://doi.org/10.1016/j.simpat.2022.102540>
- [118] Kalantar, B., Ueda, N., Idrees, M. O., Janizadeh, S., Ahmadi, K., & Shabani, F. (2020). Forest fire susceptibility prediction based on machine learning models with resampling algorithms on remote sensing data. *Remote Sensing*, 12(22), 3682. <https://doi.org/10.3390/rs12223682>