Assessing the impact of criteria removal on Multi-Criteria Decision-Making stability: A simulation-based sensitivity analysis

Jakub Więckowski
National Institute of Telecommunications
Warsaw, Poland *j.wieckowski@il-pib.pl*

Joanna Kołodziejczyk
West Pomeranian University of Technology
Szczecin, Poland *joanna.kolodziejczyk@zut.edu.pl*

Wojciech Sałabun
National Institute of Telecommunications
Warsaw, Poland *w.salabun@il-pib.pl*

Abstract

Selecting a Multi-Criteria Decision-Making (MCDM) method is critical for developing robust Decision Support Systems (DSS), yet limited attention has been given to assessing their stability under structural changes in decision problems. This study proposes a simulation-based framework for evaluating the robustness of MCDM methods when the least important criteria are iteratively removed. Four selected methods, namely Additive Ratio ASsessment (ARAS), COmplex PRoportional ASsessment (COPRAS), Measurement Alternatives and Ranking according to COmpromise Solution (MARCOS), and MultiAttributive Ideal-Real Comparative Analysis (MAIRCA) were tested across thousands of randomized scenarios, with performance assessed through mean ranking correlation, frequency of ranking alterations, and distribution of similarity values. The findings reveal consistent stability trends across methods while identifying differences in sensitivity to criteria reduction. Notably, MAIRCA and COPRAS exhibited more concise performance distributions, suggesting stronger resilience to problem changes. This work addresses a critical gap in understanding method robustness, supporting more informed selection of MCDM techniques for uncertain decision environments and enhancing the reliability of decision-making processes.
Keywords: decision support systems, decision-making, sensitivity analysis, rankings stability

1. Introduction

Reliable and flexible Decision Support Systems (DSS) are essential in a wide range of domains, from environmental planning and engineering to supply chain management [1], [13], [15], where they enable effective decision-making in complex, uncertain, and dynamic environments [5]. Many of these systems are grounded in Multi-Criteria Decision-Making (MCDM) methods, which offer structured frameworks for evaluating multiple, often conflicting criteria, thereby supporting more conscious, consistent, and transparent decision-making processes. MCDM methods provide systematic approaches for selecting the most rational decision variants by aggregating diverse decision factors, allowing decision-makers to assess trade-offs between competing objectives [7].

One of the main strengths of MCDM methods is their ability to incorporate decision-makers' preferences, allowing for tailored decisions [6]. By assigning weights to criteria, decision-makers can emphasize aspects most aligned with their strategic goals or priorities, enhancing the adaptability of MCDM methods across diverse applications [17]. However, the growing variety of MCDM techniques poses a challenge in selecting the most suitable one for a given problem [12]. Each method is based on different assumptions and computational principles, leading to varied outcomes depending on the problem structure and inputs [2]. Some methods are more sensitive to changes in weights, matrix values, or the number of criteria, which may reduce their reliability in practice [16]. Due to the dynamic nature of decision-making environments, it is essential to evaluate how MCDM methods respond to such changes and whether they maintain consistent performance.

In this context, sensitivity analysis plays a key role in evaluating the robustness and stability of MCDM approaches [8]. It helps reveal how much the outcomes of a method are affected by changes in inputs or problem structure, providing insights into their reliability and suitability for integration into resilient DSS [3]. In particular, regarding the dimensionality of the problem, the number of criteria within the problem can significantly influence the decision's outcome. Real-world situations often require adding or removing criteria due to shifts in decision-maker preferences, availability of data, or changing objectives, making it highly valuable to understand how such structural changes affect method performance.

To support the development of robust decision-making strategies, this study conducts a simulation-based sensitivity analysis to evaluate the stability of selected MCDM methods under changes in the number of criteria. Specifically, it examines how the sequential removal of the least important criteria affects final rankings. Four widely used compensatory, utility-based MCDA methods, namely Additive Ratio ASsessment (ARAS) [18], COmplex PRoportional ASsessment (COPRAS) [11], Measurement Alternatives and Ranking according to COmpromise Solution (MARCOS) [9], and MultiAttributive Ideal-Real Comparative Analysis (MAIRCA) [10]. Simulations covered decision problems with 4 to 12 criteria and a fixed set of 10 alternatives, with 100,000 iterations per configuration, resulting in a total of 900,000 scenarios. The intentional variation in the criteria set is not proposed as a normative MCDA practice, but rather as an exploratory framework to assess method behavior under dynamically constrained conditions. The goal is to identify which methods maintain stable rankings despite structural changes, aiding in the selection of MCDA tools for evolving decision environments. The main contributions of the study are: 1) a simulation-based analysis of criteria removal effects; 2) a comparative evaluation of four selected MCDA methods; and 3) a focus on stability as a key dimension in designing reliable DSS.

The rest of the paper is organized as follows. Section 2 shows the simulation-based experiments for the sensitivity analysis of selected MCDM methods. Section 3 discusses the results obtained from the simulation runs, focusing on the differences in the stability of the examined techniques. Finally, Section 4 presents conclusions drawn from the study with further research directions.

2. Simulation study

To evaluate the stability and robustness of MCDM methods, a simulation-based sensitivity analysis was conducted. The study focused on analyzing how the removal of criteria influences the ranking stability of selected MCDM methods. The simulation approach involved systematically eliminating criteria based on their importance and observing the impact on final rankings. The study aimed to identify which methods are more resilient to variations in input data, ensuring reliable decision support in dynamic environments and selecting those MCDM methods that are characterized by greater stability of operation.

The simulation setup considered decision problems with a varying number of criteria (from 4 to 12) and a fixed number of 10 alternatives. Four widely used MCDM methods, namely ARAS, COPRAS, MARCOS, and MAIRCA, were selected for evaluation based on their popularity and validated effectiveness in diverse practical decision-making contexts [4], [14]. These methods belong to the same class of compensatory, utility-based MCDA techniques, which use normalized scores and additive or proportional aggregation for ranking alternatives. This structural similarity allows for a controlled comparison of their behavior under criteria exclusion, where differences in ranking stability are more likely due to methodological nuances rather than fundamental model differences. Each scenario with a specific number of criteria was simulated over 100,000 iterations, totaling 900,000 decision-making simulations. The aim was to assess how ranking stability is influenced by the systematic removal of criteria, providing insight into the robustness of these MCDA methods under dynamic decision conditions.

The proposed experiment is designed to evaluate the robustness of MCDA methods under the exclusion of criteria based on their importance. The procedure begins by defining key experimental parameters, including the number of alternatives, criteria, MCDA methods, and the specified order for criterion removal. These parameters enable the generation of decision problems with varying dimensions and complexities, facilitating a thorough exploration of MCDA behavior under systematic reductions in criteria.

For each experimental scenario, a random decision matrix is generated with random criteria weights. Criteria are assigned alternating types (profit or cost) to simulate diverse decision-making contexts. Each MCDA method is initially applied to the initial decision matrix, establishing baseline rankings and preference scores. Subsequently, the criteria are sorted according to the specified removal order, based on their importance in either descending or ascending order. In this study, we ordered the criteria weights in ascending order. The iterative removal process begins with the sorted criteria list, systematically excluding one criterion at a time. The pseudocode of the performed experiments is outlined in Algorithm 1,

Algorithm 1: Pseudocode of the experiment used for examining the robustness of MCDA methods under excluding criteria based on their relative importance

```
Input:
    List with number of criteria C
    List with MCDA methods M
    Number of iterations I
Output: List with results from evaluations
Procedure:
    for each criteria number n in C do
        for each iteration in range(I) do
            Generate random decision matrix D of size 10 × n
            Generate random criteria weights w where Σ_{i=1}^{n} w_i = 1
            Generate criteria types ct (alternating between profit and cost)
            for each MCDA method method in M do
                Evaluate initial decision problem using method
                Save initial preference scores and rankings
            Sort criteria based on weights in ascending order
            for each criterion c_k in sorted C do
                Remove criterion c_k from C and ct
                Redistribute removed weight to remaining criteria
                Normalize modified criteria weights
                for each MCDA method method in M do
                    Evaluate the decision problem with modified criteria using method
                    Save new preference scores and rankings
                    Compute the correlation of rankings using WS
                    Save results for analysis
                if rankings are altered OR only two criteria remain then
                    Stop the removal process
```

3. Results

The obtained results were analyzed from three perspectives: 1) the mean WS correlation value for a given number of criteria simulated in the problem, along with the number of removed criteria that led to ranking alterations; 2) the count of criteria removals for a given number of criteria simulated in the problem, and the number of removed criteria causing ranking alterations; and 3) the distribution of WS correlation values for a given number of criteria simulated in the problem, and the number of removed criteria causing ranking alterations.

Figure 1 displays heatmaps of the mean WS correlation values calculated from the simulation results. For each scenario, where a particular number of criteria was modeled, the heatmap illustrates cases in which a specific number of criteria, ordered by least relevance, were removed. The mean WS values for these cases were computed and visualized. Scenarios where removal exceeded available criteria or caused no ranking change were grouped and assigned a WS value of 0.000, as the process stopped when only two criteria remained.

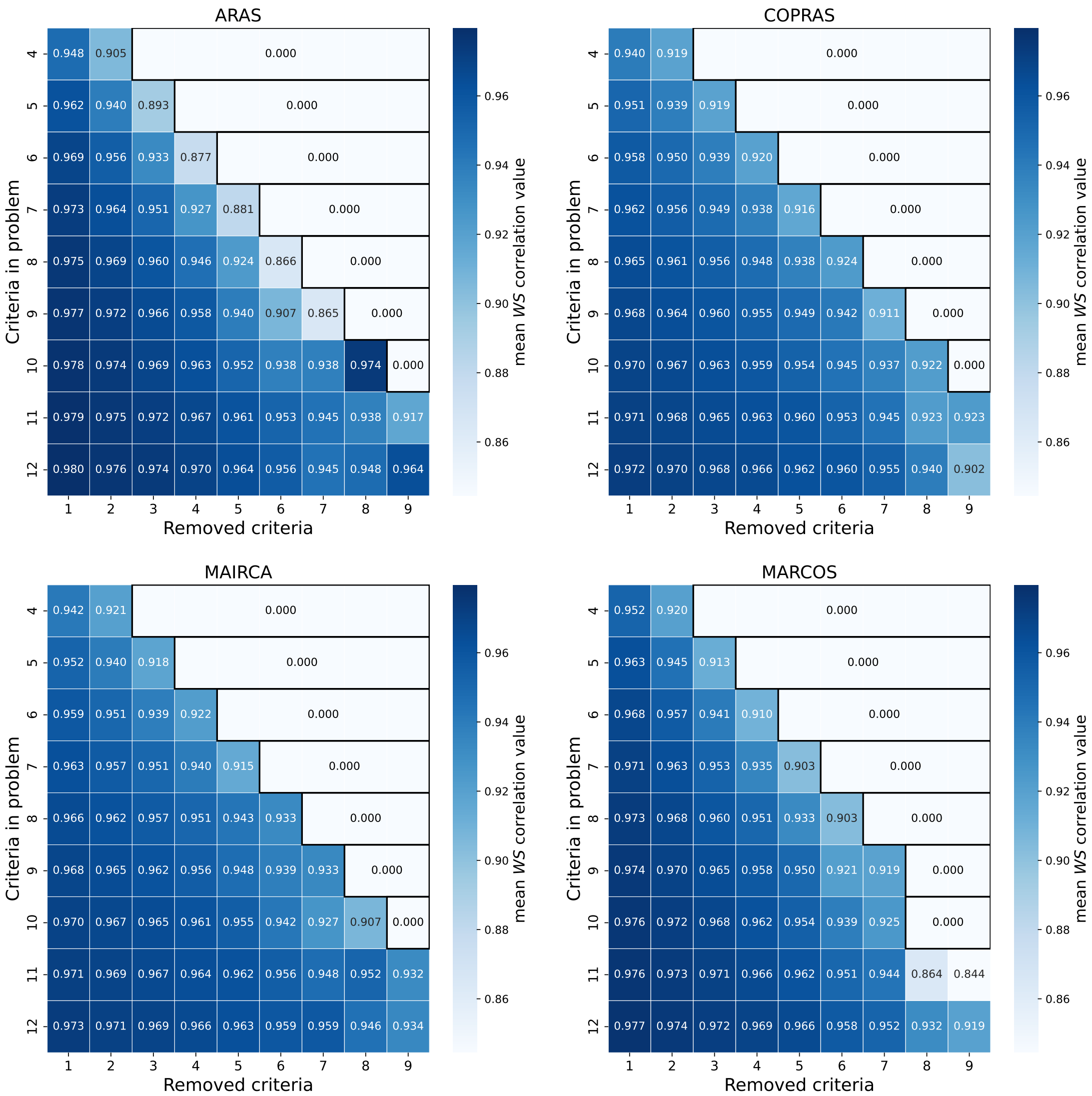

Fig. 1. Heatmaps presenting the mean WS correlation values between reference rankings and rankings calculated with removed criteria

The ideal outcome would be to maintain high mean WS correlation values, with minimal changes to the ranking caused by the removal of a greater number of criteria. This would indicate that a given MCDA method is more stable and does not rely excessively on particular criteria in its calculations. Based on the presented visualizations, a general trend emerges: as the number of criteria in the problem increases and the number of removed criteria decreases, the mean WS correlation values for the examined MCDA methods tend to be higher. The variation in correlation values across methods was minimal, ranging from 0.972 to 0.980 for 12 criteria and 1 removed criterion. A similar trend was observed for scenarios with fewer criteria in the problem and more removed criteria that led to ranking alteration.

Furthermore, border cases, where the maximum possible number of criteria was removed, leaving only two criteria, are noteworthy. In these cases, the ARAS method exhibited the lowest stability, with mean WS values approximately 0.02 to 0.07 lower than those of the other methods. This suggests that the ARAS method may be less stable when most of the less important criteria are removed from the problem.

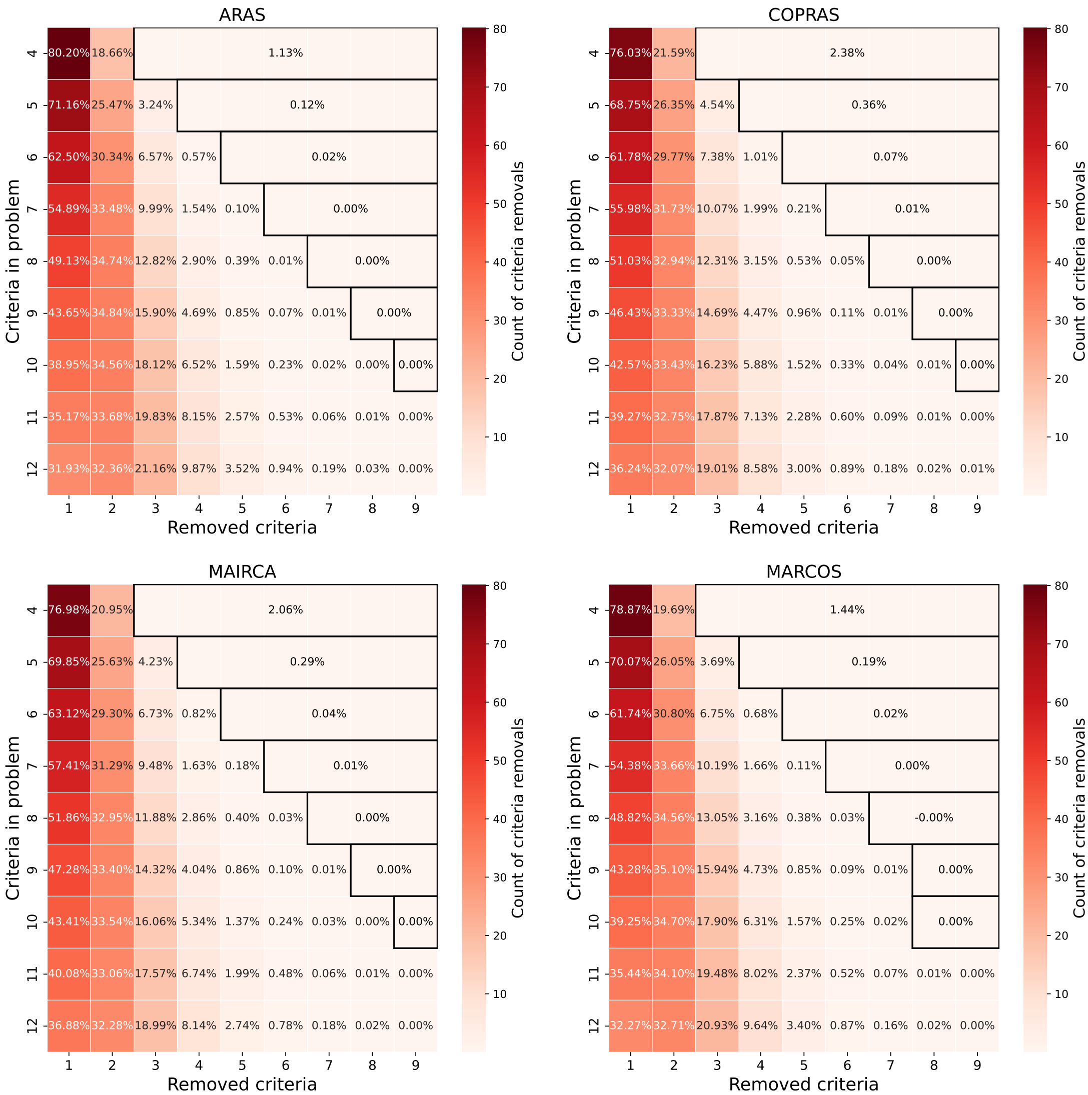

Fig. 2. Heatmaps presenting the count of values of how frequently a given number of removed criteria cause ranking alteration

Figure 2 presents heatmaps showing the count of scenarios in which a particular number of criteria had to be removed to alter the ranking order, for a given number of criteria in the problem. The count is represented as a percentage of the 100,000 iterations performed for each scenario. As in the previous visualization, scenarios that were not feasible, such as when the number of removed criteria exceeds the available criteria (leaving only two criteria) or when no scenario resulted in ranking alteration, were grouped together and assigned a remaining percentage value corresponding to the scenario where only two criteria remained in the problem without affecting the ranking.

The results reveal that for most of the scenarios causing ranking alterations, only one least important criterion needed to be removed (particularly for problems with 7 or fewer criteria). However, for problems with 8 to 12 criteria, the majority of scenarios that caused ranking alterations required the removal of up to three least important criteria. Interestingly, the methods performed differently across various aspects of the removal process. The COPRAS method achieved the lowest percentage for scenarios in which removing just one criterion caused ranking alterations when there were 4 criteria in the problem. In contrast, the ARAS and MARCOS methods had higher percentages of scenarios where removing two criteria led to ranking alterations compared to the removal of only one criterion from the decision matrix, particularly for problems with 12 criteria.

The COPRAS and MAIRCA methods achieved a higher number of scenarios where the ranking remained unchanged after reaching two criteria in the problem, compared to the ARAS and MARCOS methods. On the other hand, the ARAS and MARCOS methods consistently reached higher percentage values for scenarios where more than two criteria needed to be removed to cause ranking alterations, especially for problems with 8 or more criteria. This suggests that the ARAS and MARCOS methods may exhibit greater stability when dealing with the removal of criteria in more complex decision problems. Figure 3 presents the distribution of WS correlation values across the examined methods and varying numbers of criteria in the problem.

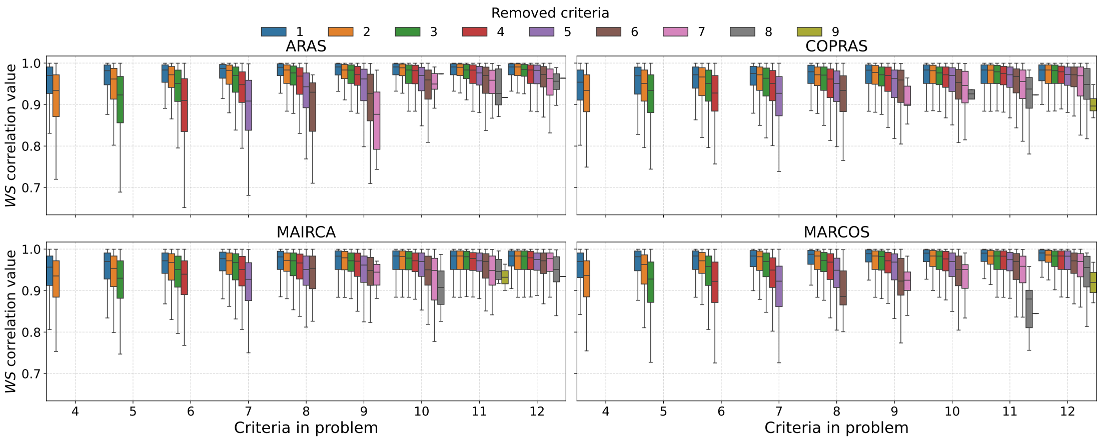

Fig. 3. Distribution of WS correlation values for the examined number of criteria in the problem

The simulation results show that ARAS and MARCOS generally achieved slightly higher WS correlation values, especially when only a few criteria needed to be removed to change the rankings. However, ARAS also showed the widest spread in correlation values, particularly in cases with nine or fewer criteria, where the values ranged from 0.65 to 1.00. It points to less consistent stability. In contrast, the COPRAS and MAIRCA methods maintained higher and more concised WS values, usually between 0.75 and 1.00, across all scenarios. The ARAS and MARCOS methods had similar patterns, especially in edge cases near the stop condition, where ranking stability varied more. The COPRAS and MAIRCA methods showed more stable behavior, with narrower WS distributions and stronger resistance to changes, making them more robust across different setups and input variations.

Although the study offers valuable insights, it also has some limitations. Equal weight redistribution was used for simplicity and consistency, but it may not fully reflect how decision-makers adjust weights in real-world scenarios. This approach allowed for isolating the effects of criterion removal across a large number of randomized cases. It is also important to distinguish between two types of ranking changes: 1) beneficial sensitivity, where methods respond meaningfully to relevant input changes, and 2) problematic instability, where methods react inconsistently or excessively. In this study, criteria were removed in descending order of importance, and ranking stability was considered desirable, as the least important criteria were excluded. However, in a reversed scenario, where the most important criteria are removed, unchanged rankings would be questionable, as such changes should naturally impact results. It would be meaningful to explore the boundary between beneficial sensitivity and instability to better support decision-makers working in dynamic and evolving decision environments.

4. Conclusions and future directions

The simulation results provide valuable insights into the stability of four widely used MCDM methods under criteria removal scenarios. While all methods showed generally high robustness, COPRAS and MAIRCA demonstrated consistently strong stability across all settings, maintaining tighter distributions of WS rank similarity values and stable rankings. This suggests reduced sensitivity to structural changes in the decision problem. In contrast, ARAS and MARCOS showed slightly higher mean WS values when only a few criteria were removed, particularly in larger problems, but exhibited greater variability as more criteria were excluded. These differences indicate that COPRAS and MAIRCA may be more suitable for dynamic or incomplete decision environments, underscoring the importance of selecting MCDA methods based on the specific characteristics and stability requirements of the application context. Despite the indicated limitations of the study, the obtained results provide valuable knowledge about the performance of the examined methods in decision problems vulnerable to changes in the set of criteria.

For future research, expanding the experimental design to include varying numbers of alternatives, more diverse benefit/cost modeling strategies, and alternative weight distributions would improve the depth of findings drawn from the results. Moreover, incorporating a broader range of MCDA methods and validating the findings using real-world decision problems would significantly improve the practical relevance and applicability of the proposed approach.

References

[1] Alvandi, E., Sadoddin, A., Sheikh, V.: A web-based integrated Bayesian decision support system for rangeland management in semi-arid watersheds. Environmental Science & Policy 141, pp. 80–96 (2023)

[2] Bandyopadhyay, S.: Comparison among multi-criteria decision analysis techniques: A novel method. Progress in Artificial Intelligence 10(2), pp. 195–216 (2021)

[3] Demir, G., Chatterjee, P., Pamucar, D.: Sensitivity analysis in multi-criteria decision making: A state-of-the-art research perspective using bibliometric analysis. Expert Systems with Applications 237, pp. 121660 (2024)

[4] Gopisetty, Y.B., Sama, H.R.: Skewness impact through distributional evaluation (SITDE) method: a new method in multi-criteria decision making. Journal of the Operational Research Society pp. 1–21 (2024)

[5] Gupta, S., Modgil, S., Bhattacharyya, S., Bose, I.: Artificial intelligence for decision support systems in the field of operations research: review and future scope of research. Annals of Operations Research 308(1), pp. 215–274 (2022)

[6] Kizielewicz, B., Tomczyk, T., Gandor, M., Sałabun, W.: Subjective weight determination methods in multi-criteria decision-making: a systematic review. Procedia Computer Science 246, pp. 5396–5407 (2024)

[7] Kumar, R.: A Comprehensive Review of MCDM Methods, Applications, and Emerging Trends. Decision Making Advances 3(1), pp. 185–199 (2025)

[8] Nabavi, S.R., Wang, Z., Rangaiah, G.P.: Sensitivity analysis of multi-criteria decision-making methods for engineering applications. Industrial & Engineering Chemistry Research 62(17), pp. 6707–6722 (2023)

[9] Nguyen, H.Q., Nguyen, V.T., Phan, D.P., Tran, Q.H., Vu, N.P.: Multi-criteria decision making in the PMEDM process by using MARCOS, TOPSIS, and MAIRCA methods. Applied sciences 12(8), pp. 3720 (2022)

[10] Ozdemir, S., Demirel, N., Zaralı, F., Çelik, T.: Multi-criteria assessment framework for evaluation of Green Deal performance. Environmental Science and Pollution Research 31(3), pp. 4686–4704 (2024)

[11] Rehman, S., Rehman, N., Naz, M., Mumtaz, A., Jianglin, Z.: Application of Grey-Based SWARA and COPRAS Techniques in Disease Mortality Risk Assessment. Journal of healthcare engineering 2021(1), pp. 7302157 (2021)

[12] Sałabun, W., Wątróbski, J., Shekhovtsov, A.: Are mcda methods benchmarkable? a comparative study of topsis, vikor, copras, and promethee ii methods. Symmetry 12(9), pp. 1549 (2020)

[13] Shided Keniwe, M.S., Ali, A.H., Abdelaal, M.A., Yassin, A.M., Kineber, A.F., Nosier, I.A.R., El Monayeri, O.D., Elsayad, M.A.: Introducing a novel decision support system to enhance performance in infrastructure sanitation projects within the construction industry. Engineering, Construction and Architectural Management (2024)

[14] Tamer, Z., Demir, G., Darıcı, S., Pamučar, D.: Understanding twitter in crisis: a roadmap for public sector decision makers with multi-criteria decision making. Environment, Development and Sustainability pp. 1–37 (2025)

[15] Unhelkar, B., Joshi, S., Sharma, M., Prakash, S., Mani, A.K., Prasad, M.: Enhancing supply chain performance using RFID technology and decision support systems in the industry 4.0–A systematic literature review. International Journal of Information Management Data Insights 2(2), pp. 100084 (2022)

[16] Więckowski, J., Sałabun, W.: Sensitivity analysis approaches in multi-criteria decision analysis: A systematic review. Applied Soft Computing 148, pp. 110915 (2023)

[17] Więckowski, J., Shekhovtsov, A., Franczyk, B., Wątróbski, J., Sałabun, W.: Personalized decision support enhanced by multiple Expected Solution Points in the Characteristic Objects Method. IEEE Access (2024)

[18] Zavadskas, E.K., Turskis, Z.: A new additive ratio assessment (ARAS) method in multicriteria decision-making. Technological and economic development of economy 16(2), pp. 159–172 (2010)