Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Policy problem coherence: preference–impact alignment in multicriteria decision-making

Monique Borges

monique@ua.pt

Research Unit GOVCOPP

João Lourenço Marques

Research Unit GOVCOPP

Paulo Batista

Research Unit GOVCOPP

Eduardo Anselmo Castro

Research Unit GOVCOPP

Article

Keywords: Policy problem structuring, Governance complexity, Participatory processes, Multicriteria Decision-making, Preference–impact alignment

Posted Date: May 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9567597/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Policy problem coherence: preference–impact alignment in multicriteria decision-making

Abstract

Multicriteria decision-making (MCDM) methods are increasingly used to support complex planning and policy processes, particularly in participatory contexts where multiple objectives and stakeholder preferences must be reconciled. However, when applied to the same decision problem, these methods often produce divergent rankings of alternatives, raising challenges for interpretation and decision-making.

This paper examines argues that variation in outputs reflects both methodological differences and structural properties of the decision problem. It focuses on the relationship between collective preference structures and the distribution of impacts across policy alternatives, conceptualising this relationship through the notion of preference–impact alignment.

The analysis draws on three participatory strategic planning processes and applies four widely used MCDM methods (AHP, TOPSIS, SAW, and WPM). The results show that higher levels of alignment are associated with greater convergence across methods, whereas lower alignment corresponds to increased divergence and sensitivity to aggregation logic. These patterns are observed consistently across cases, although their intensity varies depending on the configuration of preferences and impacts.

The findings demonstrate that divergence across go beyond technical features and evidence a diagnostic signal of misalignment in policy problem structuring. In this sense, multicriteria tools are repositioned not only as instruments for evaluation, but as devices for appraising the coherence between values, policy design, and expected outcomes within complex governance contexts. By making this relationship explicit, the paper contributes to the policy sciences by operationalising the appraisal function of the decision process and advancing a problem-oriented interpretation of analytical outputs in participatory policy-making.

Keywords: Policy problem structuring; Governance complexity; Participatory processes; Multicriteria Decision-making; Preference–impact alignment

1. Introduction

Decision-making in planning and policy contexts is inherently complex, involving the reconciliation of multiple objectives, heterogeneous information, and diverse stakeholder preferences. In participatory planning processes in particular, decisions are shaped not only by technical assessments but also by collective perceptions, expectations, and value judgments. Decisions involve processes of problem structuring, value articulation, and collective interpretation shaped by institutional settings and social dynamics (Jann & Wegrich, 2007; Majone, 2006). In this context, governance arrangements play a central role in shaping how decisions are made, particularly in participatory and multilevel settings where multiple actors interact across scales (Kooiman, 2003; Rhodes, 1996; Teles, 2016). These processes are further influenced by cognitive and social mechanisms, including bounded rationality and the role of heuristics in judgement and decision-making (Simon, 1960; Tversky & Kahneman, 1974).

Multicriteria decision-making (MCDM) methods have been widely adopted as tools to support such complex decisions, enabling the structured evaluation of alternatives across multiple criteria and facilitating the incorporation of stakeholder preferences (Dodgson et al., 2009; Tzeng & Huang, 2011).

In fact, in public sector planning contexts, decisions often rely on multicriteria evaluation methods to prioritise interventions, allocate resources, and support strategic choices. However, different methods frequently produce inconsistent rankings, creating ambiguity rather than clarity for decision-makers. This raises a critical but underexplored problem: how can decision-makers assess whether a given ranking is robust and reliable when methodological choices lead to divergent outcomes? This paper addresses this problem by introducing a diagnostic perspective that links preference structures and impact distributions, allowing the identification of conditions under which decision outcomes are stable or method-sensitive.

Multicriteria analysis has proven to be a practical operational approach for decision-making processes. It is particularly well-suited for managing the complexity that arises from conflicting objectives, diverse data and information structures, and competing interests and perspectives (Bana e Costa et al., 2014; Garmendia & Gamboa, 2012; Hwang & Yoon, 1981; Keeney, 2004; Kwan, 1995; Madadian et al., 2012; Marques et al., 2021; Nogueira et al., 2017; Thokala & Duenas, 2011; Valiris et al., 2005; Wolf et al., 2021; Yilmaz & Harmancioglu, 2010). By considering all of these factors, the multicriteria analysis provides a comprehensive framework for making informed decisions that can yield optimal outcomes. The multicriteria methods that are used most often are outlined in Table 1.

Table 1: List of commonly used multicriteria methods

Acronym	**Method**	**References**
MAUT	Multi-Attribute UtilityTheory	(Fishburn, 1970; Fishburn & Keeney, 1974)
ELECTRE	Elimation Et Choix Traduisant la REalité	(Roy, 1991)
PROMETHEE	Preference Ranking Organization Method for Enrichment of Evaluations	(Brans & Vincke, 1985)
AHP	Analytic Hierarchy Process	(Saaty, 1977, 1990)
ANP	Analytic Network Process	(Saaty, 2006, 2008; Saaty & Vargas, 2013)
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution	(Hwang et al., 1993; Hwang & Yoon, 1981)
MACBETH	Measuring Attractiveness by a Categorical Based Evaluation Technique	(Bana e Costa et al., 2011)
SMART	Simple Multi-Attribute Rating Technique	(Edwards & Barron, 1994)
SAW	Simple Additive Weighting	(Triantaphyllou, 2000)
WPM	Weighted Product Method	

Despite their widespread application, the use of MCDM methods in real-world planning contexts raises persistent challenges. In practice, the selection of methods often prioritizes theoretical simplicity and ease of application, rather than their adequacy to the specific characteristics of the decision problem (Cinelli et al., 2022; Wang et al., 2009). As a result, similar decision problems are frequently approached using different methods, which may produce divergent rankings of alternatives. While such divergence is often attributed to differences in the mathematical structure and aggregation logic of each method (Triantaphyllou, 2000; Zanakis et al., 2000), it also creates practical difficulties for interpretation and decision-making, particularly when results are used to support collective or policy-relevant choices.

Existing research has largely focused on the internal properties of multicriteria methods characteristics (Belton & Stewart, 2002; Wątróbski et al., 2019), including consistency, robustness, and methodological performance (Athawale & Chakraborty, 2012; Azhar et al., 2021;

Figueira et al., 2005; Greco et al., 2025; Roy, 1996). However, less attention has been given to how the characteristics of the decision problem itself, namely the structure of collective preferences and the distribution of impacts across alternatives, influence the results obtained. In planning practice, where alternatives often reflect complex and interdependent interventions, understanding this relationship is critical to ensure that decision-support tools are used appropriately and interpreted meaningfully (Keeney, 2004; Marttunen et al., 2017).

From a policy sciences perspective, this is a limitation. Policy problems are not given but constructed, shaped by institutional logics, governance arrangements, and social representations (Moscovici, 1988; Schneider & Ingram, 2008; Surel, 2000). Analytical tools, therefore, do not merely process information, they actively participate in structuring policy problems and framing decision outcomes. In addition, planning and policy research has highlighted the importance of interpreting analytical outputs within their institutional and participatory context, rather than treating them as neutral or purely technical results (Fischer et al., 2007; Healey, 1997; Innes & Booher, 2010). This perspective recognises that planning processes are socially constructed and shaped by interaction, power relations, and context-specific problem framings (Flyvbjerg & Richardson, 2002), and that analytical tools embed normative assumptions that influence how decision problems are structured and interpreted (Marttunen et al., 2017; Munda, 2004).

This paper builds on this premise by proposing a shift in perspective: rather than interpreting variation across MCDM methods as a purely technical issue, it is examined as a reflection of the structural properties of the decision problem. Specifically, the analysis focuses on the relationship between collective preference structures and the configuration of impacts across alternatives. To capture this relationship, the paper introduces the concept of *preference-impact alignment*, defined as the degree of correspondence between the hierarchy of values expressed through stakeholder preferences and the distribution of impacts associated with policy alternatives. This concept is proposed as a diagnostic dimension that supports the interpretation of decision outcomes in context.

The empirical analysis draws on data from three participatory strategic planning processes, in which individual preferences were elicited and aggregated to define collective weight structures. Four widely used MCDM methods (AHP, TOPSIS, SAW, and WPM) are applied to a common decision structure, allowing for a comparative analysis of ranking outcomes. By examining both the similarity of rankings and their relationship to the underlying preference and impact structures, the paper provides insights into why different methods may converge or diverge in practice.

By shifting the focus from methodological comparison to the interpretation of results in context, this study contributes to a more nuanced understanding of decision-support tools in planning processes. It highlights the importance of analysing the relationship between preferences and impacts when applying multicriteria methods and discusses the implications of these dynamics for participatory decision-making and policy design.

This paper makes three main contributions. First, it shifts the focus from comparing multicriteria methods as technical artefacts to interpreting their results within the context of real-world planning processes. In doing so, it addresses a persistent gap between the theoretical development of MCDM methods and their practical application in participatory decision-making contexts.

Second, the paper introduces the concept of preference–impact alignment as an analytical dimension that captures the relationship between the structure of collective preferences and the distribution of impacts across alternatives. This perspective provides a structured way to interpret convergence and divergence across methods, moving beyond explanations based solely on methodological differences.

Third, drawing on empirical data from three participatory strategic planning processes, the paper demonstrates how variations in alignment are associated with differences in ranking outcomes. By linking these patterns to the characteristics of the decision problem, the analysis contributes to a more informed and critical use of decision-support tools in planning practice.

Together, these contributions support a more context-sensitive interpretation of multicriteria analysis, emphasizing the need to consider both methodological properties and problem structure when applying MCDM in policy and planning settings.

2. Conceptual framing and analytical approach

Policy sciences conceptualise decision-making as a process that extends beyond the selection of alternatives to include problem structuring, framing, and interpretation (Jann & Wegrich, 2007; Sabatier, 2007). Within this perspective, policy-relevant knowledge is inherently contextual, shaped by governance arrangements, institutional settings, and cognitive processes. At the same time, insights from cognitive and behavioural perspectives highlight that decision-making is bounded and influenced by heuristics, biases, and social representations (Borges et al., 2020; Kahneman et al., 1982; Moscovici, 2001; Simon, 1960). These dimensions are particularly relevant in participatory settings, where preferences are formed, negotiated, and transformed through interaction (Nogueira et al., 2017; Wolf et al., 2021). Within this framework, MCDM methods can be understood as tools that support specific stages of the policy process, particularly appraisal and evaluation. However, their outputs are contingent on how the decision problem is structured, including the selection of criteria, the weighting of preferences, and the representation of impacts.

MCDM models are typically structured around three core elements: a set of evaluation criteria, a vector of weights reflecting the relative importance assigned to those criteria, and a decision matrix describing the performance of each alternative with respect to each criterion. Different methods combine these elements through specific aggregation procedures, which may lead to different rankings of alternatives even when applied to the same data.

Most comparative studies in the MCDM literature focus on differences in aggregation rules, normalization procedures, or consistency conditions across methods (Athawale & Chakraborty, 2012; Wątróbski et al., 2019). Within this perspective, variation in results is primarily interpreted as a consequence of methodological design. However, this approach tends to treat the decision problem as fixed and neutral, overlooking the extent to which the empirical configuration of the data may condition the behaviour of different methods.

Recent work in multicriteria decision analysis has increasingly emphasized that the performance and suitability of MCDM methods depend not only on their internal properties but also on the characteristics of the decision context, including the structure of preferences, the nature of criteria, and the distribution of performances across alternatives (Cinelli et al., 2022; Greco et al., 2025). In particular, concerns related to robustness and sensitivity have highlighted that variations in input data and modelling assumptions may lead to substantially different outcomes, raising questions about how such variation should be interpreted (Więckowski & Sałabun, 2023).

Building on this perspective, this paper adopts a problem-structural view of multicriteria decision-making. Rather than focusing exclusively on differences between methods, it examines how the structure of the decision problem interacts with aggregation procedures to produce convergence or divergence in ranking outcomes. In this context, the structure of the decision problem is understood as the joint configuration of two elements: i) the collective preference structure, represented by the vector of weights assigned to evaluation criteria, and ii) the impact

configuration, represented by the distribution of alternative performances across those criteria in the decision matrix.

This definition emphasizes that MCDM outcomes depend not only on how weights and impacts are aggregated, but also on how they relate to each other. Accordingly, variation in rankings across methods is interpreted not only as a consequence of methodological differences, but also as a function of how alternatives are positioned relative to the underlying preference structure..

2.1 Preference-impact alignment

Within this framework, a key analytical dimension concerns the degree of alignment between the preference structure and the impact configuration. Preference-impact alignment refers to the extent to which the relative performance of an alternative across criteria corresponds to the relative importance assigned to those criteria.

More specifically, alignment is higher when an alternative performs comparatively better on criteria with higher weights, and lower when its strongest performances are associated with less important criteria or when its performance profile does not reflect the hierarchy of preferences. This concept captures the degree to which the structure of impacts is consistent with the structure of preferences.

While this relationship is implicitly embedded in most MCDM formulations, it is rarely operationalized as an explicit analytical dimension. In this paper, preference-impact alignment is introduced as an empirical measure that allows the relationship between weights and impacts to be examined systematically across alternatives and cases.

Rather than assuming a predefined functional form, alignment is assessed empirically by examining the association between the ranking of criteria weights and the ranking of criterion-specific impacts for each alternative. This correspondence is approximated using rank-based correlation measures, which allow for the identification of alternatives that are more or less aligned with the collective preference structure while remaining robust to differences in scale and distribution.

This operationalization does not aim to provide a complete representation of the relationship between weights and impacts, but rather to capture a comparable and interpretable proxy of their relative correspondence across cases. As such, it enables the analysis of alignment as a structural property of the decision problem that can be related to variation in ranking outcomes across methods.

Despite its intuitive relevance, this relationship has rarely been treated as a distinct analytical dimension in the interpretation of multicriteria results. Existing work on robustness and sensitivity analysis has primarily focused on how variations in inputs, parameters, or modelling assumptions affect decision outcomes. These approaches are essential for assessing the stability of results under change. However, they typically treat divergence as a response to perturbations of the model, rather than as a property that may already be embedded in the structure of the decision problem itself.

In contrast, this study does not analyse how rankings change under alternative specifications, but examines how differences between methods emerge from a fixed empirical configuration of weights and impacts. The proposed concept of preference–impact alignment addresses this gap by characterising the internal correspondence between the collective preference structure and the distribution of impacts across alternatives.

This distinction is analytically relevant because existing approaches do not provide a direct way to interpret whether divergence across methods reflects methodological differences or structural tensions within the decision problem. By making this relationship explicit, preference–impact alignment allows identifying when convergence is driven by coherent problem structure and when divergence signals a misalignment between what is valued and how alternatives perform.

In this sense, alignment is not an alternative to robustness or sensitivity analysis, but a complementary structural descriptor that helps explain why different methods may produce similar or conflicting rankings even in the absence of model variation.

2.2 Structural conditions and ranking behaviour

Different MCDM methods combine weights and impacts in distinct ways, which makes them more or less sensitive to particular configurations of the decision problem. From this perspective, convergence and divergence in rankings emerge from the interaction between method-specific aggregation rules and the structural characteristics of the decision problem.

When the impact profiles of alternatives are broadly consistent with the preference structure, aggregation procedures tend to reinforce similar ordering patterns, resulting in higher convergence across methods. In such cases, the relative dominance of alternatives is less sensitive to the specific aggregation logic applied.

In contrast, when impact profiles are unevenly distributed or weakly aligned with the preference structure, different methods may emphasize different aspects of the same performance profile. Under these conditions, aggregation procedures may privilege distinct trade-offs between criteria, leading to greater divergence in rankings.

This interpretation aligns with broader concerns in MCDA regarding the sensitivity of results to modelling assumptions and the need to interpret ranking outcomes in light of both method design and problem structure (Greco et al., 2025; Pinzon Amorocho & Hartmann, 2022; Więckowski & Sałabun, 2023). Rather than treating divergence as an indicator of methodological inconsistency, it is therefore interpreted here as an expression of structural tension within the decision problem.

2.3 Analytical expectations

Based on this framing, two analytical expectations guide the empirical analysis. First, if differences in outcomes are associated with the structure of the decision problem, then variations in ranking similarity across methods should be systematically related to variations in the configuration of weights and impacts.

Second, higher levels of preference-impact alignment are expected to be associated with greater convergence in rankings across methods, while lower levels of alignment are expected to correspond to increased divergence and sensitivity to aggregation logic.

These expectations do not imply deterministic or causal relationships, but provide a basis for assessing whether consistent patterns can be identified across empirical cases. The aim is not to establish the superiority of any particular method, but to examine whether differences in outcomes can be interpreted in light of the structural properties of the decision problem.

This paper contributes to the MCDA literature by providing a more explicit and operational account of how decision outcomes depend on the empirical configuration of the problem. While

it is well established that multicriteria results are context-dependent, this study advances the discussion by introducing the concept of preference-impact alignment. Rather than proposing a new decision rule, alignment is conceptualised as a diagnostic analytical dimension that captures the degree of correspondence between collective preference hierarchies and alternative-specific impact profiles. This approach enables a more systematic interpretation of ranking convergence and divergence across methods, linking cross-method variation to observable properties of the decision problem itself.

Existing work on robustness and sensitivity analysis has primarily focused on how variations in inputs, parameters, or modelling assumptions affect decision outcomes. These approaches are essential for assessing the stability of results under change. However, they typically treat divergence as a response to perturbations of the model, rather than as a property that may already be embedded in the structure of the decision problem itself.

In contrast, this study does not analyse how rankings change under alternative specifications, but examines how differences between methods emerge from a fixed empirical configuration of weights and impacts. The proposed concept of preference–impact alignment addresses this gap by characterising the internal correspondence between the collective preference structure and the distribution of impacts across alternatives.

This distinction is analytically relevant because existing approaches do not provide a direct way to interpret whether divergence across methods reflects methodological differences or structural tensions within the decision problem. By making this relationship explicit, preference–impact alignment allows identifying when convergence is driven by coherent problem structure and when divergence signals a misalignment between what is valued and how alternatives perform. In this sense, alignment is not an alternative to robustness or sensitivity analysis, but a complementary structural descriptor that helps explain why different methods may produce similar or conflicting rankings even in the absence of model variation.

Existing work on robustness and sensitivity analysis has primarily focused on how variations in inputs, parameters, or modelling assumptions affect decision outcomes. These approaches are essential for assessing the stability of results under change. However, they typically treat divergence as a response to perturbations of the model, rather than as a property that may already be embedded in the structure of the decision problem itself.

In contrast, this study does not analyse how rankings change under alternative specifications, but examines how differences between methods emerge from a fixed empirical configuration of weights and impacts. The proposed concept of preference–impact alignment addresses this gap by characterising the internal correspondence between the collective preference structure and the distribution of impacts across alternatives.

This distinction is analytically relevant because existing approaches do not provide a direct way to interpret whether divergence across methods reflects methodological differences or structural tensions within the decision problem. By making this relationship explicit, preference–impact alignment allows identifying when convergence is driven by coherent problem structure and when divergence signals a misalignment between what is valued and how alternatives perform. In this sense, alignment is not an alternative to robustness or sensitivity analysis, but a complementary structural descriptor that helps explain why different methods may produce similar or conflicting rankings even in the absence of model variation.

3. Methods

3.1 Empirical setting and data

The empirical analysis is based on three participatory strategic planning processes in which stakeholder preferences were elicited using the Analytic Hierarchy Process (AHP). Across the three cases, a total of 146 individual responses were collected and used to derive collective preference structures. In each case, the evaluation framework was organised hierarchically, through structured pairwise comparisons, reflecting the decomposition of the decision problem into multiple levels of criteria and sub-criteria (see Figure 1). Individual judgments were subsequently aggregated to derive a single vector of criteria weights for each case, representing the relative importance assigned to each evaluation dimension. These weights reflect a collective preference structure grounded in participatory elicitation. This structure enabled the systematic elicitation of preferences while capturing the multidimensional nature of the planning objectives.

Criteria *level 0* →	A				B				C				D			E		
Criteria *level 1* →	A 1	A 2	A 3	A 4	B 1	B 2	B 3	B 4	C 1	C 2	C 3	C 4	D 1	D 2	D 3	E 1	E 2	E 3
Alternative 1	1,4	0,1	0,9	0,8	0,2	0,7	0,8	1,0	1,9	2,6	3,3	4,5	1,0	1,7	0,5	2,9	1,9	1,2
Alternative 2	1,8	0,5	1,8	0,5		0,5	0,5			4,6	0,5	0,7	1,6	0,7	0,7	2,3	5,0	4,1
Alternative 3	1,8		1,8	0,6		0,6	0,6	0,3	2,5	1,2		0,3	3,1	3,1	1,2	3,1	3,1	1,8
Alternative 4	1,2	0,8	0,6	0,6	0,8		0,2	0,6	1,9	0,8	1,9	1,6	3,5	3,5	1,0	3,5	1,9	1,2
Alternative 5			0,5	0,2	0,2	1,5	2,2	1,2	1,0	3,2	0,2	2,5	3,2	1,5	0,7	4,4	1,0	2,2
Alternative 6	2,3		1,1	0,4	2,3	2,3	2,3	2,3	0,8	4,5		1,1	1,1			3,0	1,1	3,0
Alternative 7	1,0	1,3	2,0	2,0	0,3	3,0	2,3	1,0	1,0	3,0	1,3	2,0	0,7			3,0	3,0	1,0
Alternative 8	5,1	5,1	4,1	3,1	1,0			1,0	0,5	0,5	0,5		2,6		3,3			1,0
Alternative 9	6,7	2,4	6,7	6,3				0,4	1,6	2,4			0,8		1,6			0,4
Alternative 10	5,6	0,7	3,5	2,8	0,7	1,4	0,7	0,7		4,2	1,4	2,8	1,4		1,4	0,7	0,7	

Figure 1: Example of a decision or impact matrix

The empirical dataset therefore consists of two main components: i) a vector of criteria weight, and ii) a decision matrix $X = [x_{ij}]$, where x_{ij} represents the performance of alternative i with respect to criterion j. Together, these elements define the empirical structure of the decision problem.

However, the multicriteria methods considered in this study do not operate on hierarchical criteria structures in a directly comparable manner. To ensure consistency across methods, a transformation of the original decision structure was required. Specifically, the hierarchical criteria were aggregated into a single-level structure by considering only the top-level criteria (level 0), which are consistent across the three cases.

The initial decision matrix was constructed by summing the impacts associated with the lower-level criteria within each top-level criterion for each alternative. This aggregation preserves the overall contribution of each criterion while allowing for comparability across methods that assume a flat decision structure.

Following this step, the resulting values were normalised under the assumption of equal weights across all evaluation criteria. Given that five criteria are considered, each criterion was assigned a weight of 0.2. This assumption allows for isolating the effect of the aggregation logic of each method, independent of differences in weighting schemes derived from the original AHP application.

To ensure comparability across alternatives under this assumption, a redistribution ratio was defined and applied to the impact values of each alternative. This transformation guarantees that, when criteria are equally weighted, all alternatives would have the same total valuation, thereby neutralising scale effects and enabling a consistent comparison of ranking outcomes across methods.

The redistribution ratio is defined as:

$$R_{ij} = \frac{\sum_{i,j=1}^{n} S_{ij}}{\bar{X}(\sum_{j}^{m} \sum_{i}^{n} S_{ij})} \quad (1)$$

Where:

R, is the redistribution ratio for matrix normalization.

S, is the sum of impacts assigned to each criterion.

i, is each criterion.

j, is each alternative.

***n*,** is the total number of criteria.

m, is the total number of alternatives.

The final decision matrices obtained after applying this transformation constitute a common analytical basis for the application and comparison of the multicriteria methods. This procedure ensures that observed differences in ranking outcomes can be attributed to the aggregation logic of each method and to the structure of the decision problem, rather than to inconsistencies in data representation. The use of real-world decision contexts allows the analysis of MCDM behaviour under empirically grounded conditions, although it limits experimental control over the configuration of the data.

3.2 Multicriteria decision methods

Four multicriteria decision-making (MCDM) methods were selected for the analysis: the Analytic Hierarchy Process (AHP), the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), the Simple Additive Weighting method (SAW), and the Weighted Product Method (WPM). These methods are widely used in applied decision-support contexts and represent different aggregation logics, allowing for a structured comparison of ranking outcomes under a common decision framework.

The selection of methods was guided by two criteria. First, all methods are commonly applied in planning and policy-related contexts, ensuring the practical relevance of the analysis. Second, they reflect distinct ways of aggregating criteria and evaluating alternatives. SAW represents a fully additive and compensatory approach, in which poor performance in one criterion can be offset by good performance in another. WPM follows a multiplicative structure, preserving proportional relationships between criteria and reducing full compensability. TOPSIS is based on a distance-to-ideal solution logic, evaluating alternatives according to their relative closeness to a hypothetical best and worst solution. AHP, although originally designed for hierarchical decision structures, is implemented here in a rating-based form to ensure comparability with the other methods under a single-level decision matrix.

All methods were applied to the same transformed decision matrices described in Section 3.1, ensuring that differences in ranking outcomes result from the aggregation logic of each method rather than from variations in data structure or weighting schemes. To further enhance comparability, a consistent set of assumptions was adopted across methods, including the use of identical criteria weights and the treatment of all criteria as benefit criteria after normalisation.

The implementation of each method follows standard formulations as described in the literature (e.g., (Afshari et al., 2010; Belton & Stewart, 2002; Hwang & Yoon, 1981; Rosszkowska, 2011; Saaty, 2008; Triantaphyllou, 2000). For AHP, the rating-based approach was used to derive scores for alternatives under each criterion. For TOPSIS, vector normalisation and Euclidean distance were used to calculate the relative closeness to the ideal solution. SAW was implemented through linear aggregation of weighted normalised scores, and WPM through multiplicative aggregation of weighted criteria values.

By applying these methods under a consistent analytical framework, the study isolates the effect of different aggregation logics on ranking outcomes, enabling a systematic examination of convergence and divergence across methods in relation to the structure of the decision problem.

3.3 Analytical approach

To examine differences in ranking outcomes across multicriteria methods, a comparative analysis was conducted using a consistent decision framework for all cases. The four methods (AHP, TOPSIS, SAW, and WPM) were applied to the transformed decision matrices described in Section 3.1, ensuring that all methods operated on identical data structures, criteria sets, and weighting assumptions. This allows differences in results to be attributed to the aggregation logic of each method and to the structure of the decision problem, rather than to inconsistencies in input data.

The analysis focuses on two complementary dimensions: i) the degree of similarity between rankings produced by different methods, and ii) the relationship between ranking outcomes and the underlying configuration of preferences and impacts across alternatives.

To assess the degree of similarity between ranking outcomes, the Spearman's rank correlation coefficient (ρ) was used. For each case, ρ was computed based on the ordinal rankings of alternatives obtained from each method. This measure captures the strength and direction of the monotonic relationship between rankings, allowing for the identification of convergence and divergence patterns across methods. Spearman's coefficient was selected as it is appropriate for ordinal data and does not assume linear relationships between values, making it suitable for comparing ranking-based outputs.

In addition to ranking similarity, the analysis examines how ranking outcomes relate to the structure of the decision problem, specifically the correspondence between collective preference weights and the distribution of impacts across alternatives. This relationship is operationalised through the concept of *preference-impact alignment*, which captures the extent to which alternatives that perform well in highly weighted criteria also concentrate their impacts in those criteria. Alignment is assessed at the level of individual alternatives and interpreted in relation to the rankings produced by each method.

By combining these two analytical dimensions (ranking similarity and preference–impact alignment) the study provides a structured basis for interpreting why different methods may converge or diverge when applied to the same decision problem. This approach allows the analysis to move beyond method comparison as an end in itself and instead focus on how methodological differences interact with the characteristics of the decision context.

4. Results

4.1 Structural coherence in policy problem formulation

Variation in rankings across multicriteria methods reflects differences in the structural coherence of the policy problem, rather than solely the properties of the methods themselves. In particular, the extent to which policy alternatives embody collectively articulated values, expressed through the relationship between preference structures and impact distributions, conditions the degree of convergence or divergence observed across methods.

The application of AHP, TOPSIS, SAW, and WPM to the three decision contexts results in distinct rankings of alternatives, as presented in Tables 2–4. Across the three municipalities, a consistent pattern can be identified in the behaviour of different MCDM methods. While there is general agreement in the identification of top- and bottom-ranked alternatives, differences emerge in the relative ordering of intermediate alternatives. These variations are not uniform across contexts, suggesting that ranking behaviour is shaped by the underlying configuration of the decision problem.

In Municipality A, greater variation is observed across methods, particularly in the positioning of mid-ranked alternatives. TOPSIS shows the highest level of divergence, producing rankings that differ substantially from those obtained with SAW and WPM. In contrast, SAW and WPM display a high degree of consistency, generating similar ordering patterns across most alternatives. AHP exhibits intermediate behaviour, with its level of agreement varying depending on the configuration of the data.

A similar pattern is observed in the remaining cases, although with different intensity. In Municipalities B and C, the degree of agreement between methods is generally higher, particularly between SAW and WPM, while TOPSIS continues to display greater variability. These differences indicate that variation in rankings is not uniform across contexts and may be associated with the structural characteristics of the decision problem.

From this perspective, convergence across methods can be interpreted as an indication of coherent policy problem structuring, where the distribution of impacts is consistent with the hierarchy of preferences. Conversely, divergence signals potential misalignment between what is collectively valued and how policy alternatives perform, pointing to underlying tensions in the formulation of the policy problem.

Table 2: Scores and rankings obtained by applying AHP, TOPSIS, SAW, and WPM in Municipality A

Alternatives	AHP		TOPSIS		SAW		WPM	
	Scores	Rank	Scores	Rank	Scores	Rank	Scores	Rank
Alternative 1	93,0	10	49,8	6	82	10	3,8	9
Alternative 2	97,5	4	52,8	4	94	4	28,1	4
Alternative 3	98,1	3	48,6	7	100	1	100	1
Alternative 4	96,5	5	46,5	8	97	2	38,9	2
Alternative 5	95,3	9	34,5	10	94	5	11,0	5
Alternative 6	95,4	8	44,2	9	88	6	7,9	7
Alternative 7	95,6	7	52,1	5	86	7	9,6	6
Alternative 8	100,0	1	81,1	2	94	3	29,2	3
Alternative 9	99,3	2	100,0	1	86	8	2,5	10
Alternative 10	95,7	6	72,6	3	83	9	6,3	8

Table 3: Scores and rankings obtained by applying AHP, TOPSIS, SAW, and WPM in Municipality B

Alternatives	AHP		TOPSIS		SAW		WPM	
	Scores	Rank	Scores	Rank	Scores	Rank	Scores	Rank
Alternative 1	98,8	2	100,0	1	100	1	81,1	2
Alternative 2	83,4	6	60,6	5	76	6	4,2	6
Alternative 3	100,0	1	49,3	7	96	2	79,8	3
Alternative 4	95,4	3	68,9	3	91	3	100	1
Alternative 5	83,8	5	63,3	4	75	7	3,9	7
Alternative 6	87,1	4	83,8	2	86	4	16,2	4
Alternative 7	80,6	7	58,8	6	81	5	4,5	5

Table 4: Scores and rankings obtained by applying AHP, TOPSIS, SAW, and WPM in Municipality C

Alternatives	AHP		TOPSIS		SAW		WPM	
	Scores	Rank	Scores	Rank	Scores	Rank	Scores	Rank
Alternative 1	88,3	4	98,9	2	86	5	14,4	5
Alternative 2	90,6	3	70,4	5	88	4	15,0	3
Alternative 3	96,7	2	100,0	1	90	3	14,9	4
Alternative 4	100,0	1	77,6	3	100	1	100	1
Alternative 5	78,0	7	65,8	6	74	7	1,2	7
Alternative 6	87,9	5	62,7	7	93	2	15,5	2
Alternative 7	79,7	6	77,5	4	77	6	1,7	6

To assess convergence and divergence more systematically, pairwise correlations between rankings were computed using Spearman's rank correlation coefficient (Table 5). The results confirm the patterns observed in the ranking comparisons.

Across the three cases, additive and multiplicative models (SAW and WPM) exhibit strong and stable correlations, indicating that these methods tend to produce similar ordering patterns under the conditions analysed. In contrast, TOPSIS displays lower and more variable correlations with the other methods, in some cases approaching zero or negative values, particularly in Municipality A. This suggests that TOPSIS follows a distinct aggregation logic that is more sensitive to the distribution of impacts across criteria.

AHP shows moderate to strong correlations with SAW and WPM in Municipalities B and C, while weaker associations are observed in Municipality A. This variation suggests that the relative behaviour of AHP depends on the empirical configuration of the decision problem.

These results indicate that convergence and divergence across MCDM methods are not constant but vary across cases. While differences in aggregation logic play an important role, the observed variation suggests that method behaviour is also conditioned by the structure of the underlying decision problem.

Table 5: Correlation (Spearman's rho) between the rankings obtained by applying AHP, TOPSIS, SAW, and WPM.

Municipality A

	AHP	TOPSIS	SAW
AHP	1	0,648*	0,503
TOPSIS	,648*	1	-0,285
SAW	0,503	-0,285	1
WPM	0,37	-0,321	,952**

Level of significance: *5%; ** 1%

Municipality B

	AHP	TOPSIS	SAW
AHP	1	0,214	0,821*
TOPSIS	0,214	1	0,286
SAW	,821*	0,286	1
WPM	0,714	0,357	,893**

Level of significance: *5%; ** 1%

Municipality C

	AHP	TOPSIS	SAW	WPM
AHP	1	0,607	,786*	0,75
TOPSIS	0,607	1	0,143	0
SAW	,786*	0,143	1	0,964**
WPM	0,75	0	,964**	1

Level of significance: *5%; ** 1%

4.2 Alignment between values and policy outcomes

To further interpret the variation in ranking outcomes observed in Section 4.1, the structural characteristics of the decision problems were analysed through the relationship between preference structures and impact matrices (Tables 6 and 7). These elements define how alternatives are positioned relative to the evaluation criteria and provide the basis for assessing preference-impact alignment.

The analysis reveals substantial heterogeneity in how alternatives distribute their impacts across criteria. In several cases, alternatives concentrate a significant share of their impacts on a limited number of criteria, while others display more balanced performance profiles. More importantly, the distribution of impacts does not necessarily correspond to the hierarchy of preferences: some alternatives perform strongly on criteria with lower weights, while showing weaker performance on the most valued criteria. Conversely, other alternatives exhibit impact profiles that are more closely aligned with the preference structure.

This heterogeneity constitutes a key structural feature of the decision problem. It implies that alternatives differ not only in their overall level of performance but also in how their impacts are positioned relative to the structure of preferences.

Table 6: Example of a collective preference structure

Criteria	**Weights**
Criteria A	0,211
Criteria B	0,196
Criteria C	0,175
Criteria D	0,207
Criteria E	0,211

Table 7: Example of a normalized decision or impact matrix (characterization of each alternative in relation to evaluation criteria) & Percentage distribution of the impacts of each alternative on the criteria (derived from the normalized decision matrix)

	Alternatives									
Criteria	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Criteria A	0,63	0,97	0,92	0,66	0,15	0,74	1,22	3,38	4,10	2,38
	12%	18%	17%	12%	3%	14%	23%	62%	76%	44%
Criteria B	0,53	0,19	0,33	0,33	1,08	1,78	1,29	0,40	0,07	0,66
	10%	4%	6%	6%	20%	33%	24%	7%	1%	12%
Criteria C	2,43	1,21	0,86	1,32	1,44	1,26	1,42	0,30	0,73	1,58
	45%	22%	16%	24%	27%	23%	26%	6%	14%	29%
Criteria D	0,63	0,63	1,58	1,69	1,13	0,22	0,13	1,14	0,44	0,53
	12%	12%	29%	31%	21%	4%	2%	21%	8%	10%
Criteria E	1,19	2,42	1,72	1,40	1,60	1,41	1,35	0,20	0,07	0,26
	22%	45%	32%	26%	30%	26%	25%	4%	1%	5%

Preference-impact alignment was assessed for each alternative using rank-based correlation measures (Table 8). The results show considerable variation in alignment levels across alternatives and cases, ranging from strongly positive to negative values. This variation indicates that some alternatives are closely aligned with the collective preference structure, while others exhibit significant misalignment. Rather, it is proposed as a diagnostic analytical dimension that

captures the degree of correspondence between the hierarchy of collective preferences and the distribution of alternative performances across criteria. Its value lies in helping interpret ranking behaviour across methods by making explicit one structural feature of the decision problem that is usually implicit in multicriteria formulations.

A consistent relationship emerges between alignment and ranking behaviour. Alternatives with higher levels of preference-impact alignment tend to occupy more stable positions across methods, exhibiting higher agreement in their relative ranking. In contrast, alternatives with lower or negative alignment display greater variability, with their ranking positions shifting more substantially across methods.

This pattern is particularly evident in Municipality A, where lower overall alignment levels correspond to weaker correlations between methods, as observed in Table 5. Conversely, in Municipalities B and C, where impact distributions are more balanced and alignment levels are generally higher, stronger agreement across methods is observed.

Table 8: Correlation (Spearman's rho) between the collective preference structure and the decision matrix

Municipality A

	Alternatives									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Correlation coefficient	-0,132	0,308	0,718	0,205	-0,103	-0,205	-0,410	0,205	0,026	-0,205
Sig.	0,416	0,307	0,086*	0,370	0,435	0,370	0,246	0,370	0,483	0,370
Rank	7	2	1	3	6	8	10	3	5	8

*Level of significance: 10%

Municipality B

	Alternatives						
	1	**2**	**3**	**4**	**5**	**6**	**7**
Correlation coefficient	0,600	-0,564	0,600	-0,100	-0,718	0,100	-0,600
Sig.	0,142	0,161	0,142	0,436	0,086*	0,436	0,142
Rank	1	5	1	4	7	3	6

* Level of significance: 10%

Municipality C

	Alternatives						
	1	**2**	**3**	**4**	**5**	**6**	**7**
Correlation coefficient	0,300	0,300	0,051	0,800	-0,700	0,500	0,300
Sig.	0,312	0,312	0,467	0,052*	0,094*	0,196	0,312
Rank	3	3	5	1	6	2	3

* Level of significance: 10%

Table 9: Correlation (Spearman's rho) between the ranking obtained with the multicriteria methods and the ranking that would result from considering the preference structure

	AHP	**TOPSIS**	**SAW**	**WPM**
Municipality A	0,96**	0,52	0,61	0,50
Municipality B	0,86*	0,32	0,96**	0,86*
Municipality C	0,68	0,21	0,89**	0,86*

Level of significance: * 5% | ** 1%

These results suggest that divergence between methods is not random but reflects structural characteristics of the decision problem. When alternatives are more closely aligned with the preference structure, different aggregation procedures tend to reinforce similar ordering patterns,

resulting in higher convergence. When alignment is weaker or uneven, different methods emphasise distinct aspects of the same performance profile, leading to greater divergence.

Overall, the findings indicate that variation in multicriteria decision outcomes cannot be fully understood as a property of methods alone. Instead, ranking behaviour reflects the interaction between aggregation logic and the structural configuration of the decision problem, with preference-impact alignment emerging as a relevant analytical dimension for interpreting convergence and divergence across methods.

This structural relationship is further supported when examining the aggregated distribution of impacts across criteria in relation to the preference structure (Table 10). The results show that, particularly in Municipality A, higher accumulated impacts are not consistently associated with the most valued criteria, indicating a clear structural misalignment. In contrast, Municipalities B and C exhibit a closer correspondence between the distribution of impacts and the hierarchy of preferences. This aggregated perspective reinforces the interpretation that variation in ranking outcomes is rooted not only in the behaviour of individual alternatives but also in broader structural characteristics of the decision problem..

Table 10: Relationship between the accumulated value of the normalized decision matrix and the collective preferences' structure – Municipality A, B e C

		Criteria				
		A	**B**	**C**	**D**	**E**
Municipality A	**Preference structure** (weight)	**0,211**	0,20	**0,175**	0,207	**0,211**
	Decision matrix (impact)	**15,65**	6,70	**12,53**	7,95	**11,30**
Municipality B	**Preference structure** (weight)	**0,272**	0,158	**0,142**	**0,222**	0,206
	Decision matrix (impact)	**2,32**	1,67	**2,58**	**1,88**	1,95
Municipality C	**Preference structure** (weight)	**0,237**	0,198	0,136	**0,254**	**0,176**
	Decision matrix (impact)	**3,90**	2,11	1,68	**2,07**	**2,93**

4.3 Neutral weighting scenario and the role of impact structure

Building on the structural differences identified in Section 4.2, an additional scenario was considered in which all criteria were assigned equal weights. This configuration allows isolating the influence of the impact matrix by removing the differentiated preference structure derived from the participatory processes. While sensitivity to criteria weights is a well-known feature of MCDM methods, the results of this study indicate that ranking behaviour is not solely driven by variations in weights, but by how these weights interact with the distribution of impacts across alternatives, as reflected in the observed patterns of alignment and misalignment.

The results show that the equal-weight scenario leads to substantial changes in ranking outcomes for SAW and WPM across the three cases, with negative correlations relative to the baseline rankings in all municipalities. This behaviour is consistent with the compensatory nature of these methods, where the aggregation of weighted performances directly reflects the specified preference structure.

In contrast, TOPSIS exhibits consistently high correlations between the baseline and equal-weight scenarios across all cases. This contrast highlights that divergence across methods and sensitivity to weight structures are distinct dimensions of method behaviour. Under the observed configurations, ranking behaviour in TOPSIS is largely driven by the structure of the impact matrix rather than by the weight vector. The persistence of similar rankings suggests that the relative distances to ideal and anti-ideal solutions are dominated by the distribution of impacts across alternatives.

With respect to AHP, the equal-weight scenario produces a degenerate configuration due to the transformation of the original hierarchical structure into a single-level weighting scheme. When all criteria are assigned equal importance, the aggregation process no longer reflects the pairwise comparison logic inherent to AHP, resulting in limited or null differentiation between alternatives. For this reason, AHP results under the neutral weighting scenario are not directly comparable with those obtained for the other methods. This outcome highlights the structural dependency of AHP on its hierarchical formulation and reinforces the importance of preserving methodological coherence when constructing comparative scenarios.

Across the three municipalities, the magnitude of ranking changes varies, reflecting differences in the relationship between preference structures and impact distributions. In cases where impacts are more unevenly distributed across criteria, the removal of differentiated weights leads to stronger divergence in ranking outcomes, particularly for compensatory methods.

These results reinforce the interpretation that ranking behaviour in multicriteria decision-making reflects the interaction between aggregation logic and the structural configuration of the decision problem, particularly the degree of correspondence between preferences and impact distributions.

5. Discussion

The results highlight that differences in rankings across multicriteria methods cannot be understood solely as a consequence of their mathematical structure. Instead, they reflect how each method interacts with the underlying configuration of the decision problem, particularly the relationship between collective preferences and the distribution of impacts across alternatives. From a planning perspective, this finding is significant, as it suggests that variation in results should not be interpreted as inconsistency or methodological failure, but as an indication of structural tensions within the decision context.

The findings provide empirical support for the argument that variation in multicriteria decision outcomes cannot be fully explained by method-specific properties alone. While differences in aggregation logic contribute to divergence in rankings, the results indicate that such variation is also systematically associated with the structural characteristics of the decision problem.

This perspective extends existing work in the MCDM literature, which has predominantly focused on differences in methodological assumptions, aggregation procedures, and robustness properties (Athawale & Chakraborty, 2012; Cinelli et al., 2022; Wątróbski et al., 2019). While these contributions have clarified how methods behave under different modelling conditions, they tend to treat the decision problem as structurally neutral. The present findings suggest that this assumption is limiting, as method behaviour is also conditioned by how alternatives are positioned relative to the underlying preference structure.

In particular, the analysis highlights the role of preference–impact alignment as a relevant dimension for interpreting convergence and divergence across MCDM methods. Alternatives whose impact profiles are more closely aligned with the collective preference structure tend to

exhibit greater stability across methods, whereas alternatives with weaker or negative alignment show higher sensitivity to aggregation logic and greater variation in ranking outcomes. This result complements recent discussions on robustness and sensitivity in MCDA, which emphasise the dependence of outcomes on both modelling assumptions and input configurations (Greco et al., 2025; Więckowski & Sałabun, 2023).

These structural differences are reflected in the behaviour of the multicriteria methods. Additive and multiplicative models, such as SAW and WPM, tend to reinforce consistent ordering patterns when alternatives exhibit relatively balanced or well-aligned impact profiles, as their aggregation logic combines weights and performances in a direct and proportional manner. AHP, by incorporating pairwise comparisons and a hierarchical structuring of preferences, allows for greater differentiation across alternatives while still reflecting the underlying weighting scheme, particularly in cases where the relationship between criteria is more complex.

In contrast, TOPSIS, which evaluates alternatives based on their relative distance to ideal and anti-ideal solutions, appears more sensitive to the distribution of impacts and less directly anchored in the preference structure. This sensitivity becomes more pronounced in cases where preference–impact alignment is weaker, leading to greater variation in ranking outcomes across methods.

These differences do not indicate methodological superiority, but rather illustrate how each method responds to specific structural configurations of the decision problem. In this sense, variation in rankings reflects the interaction between aggregation logic and the degree of alignment between preferences and impacts, rather than purely methodological inconsistencies. This interpretation is consistent with recent arguments that MCDM outcomes should be analysed not only in terms of method selection but also in relation to the characteristics of the decision context (Cinelli et al., 2022).

More broadly, the findings suggest that divergence between methods should not necessarily be understood as a limitation of multicriteria approaches. Rather, it may reflect structural tensions within the decision problem itself, namely the extent to which the distribution of impacts across alternatives corresponds to the hierarchy of preferences. In this sense, variation in results can be interpreted as an informative feature of the decision context, rather than as a purely technical issue related to method selection.

From a practical perspective, this interpretation is particularly relevant in planning and policy contexts, where MCDM methods are often applied in participatory settings (Barbati et al., 2023; Kumar et al., 2017; Nogueira et al., 2017; Patel et al., 2017; Wolf et al., 2021). In such contexts, preference structures are derived from collective processes, and the relationship between those preferences and the distribution of impacts becomes a critical factor in interpreting results. The identification of alignment patterns may therefore support a more informed use of MCDM, helping decision-makers to distinguish between stable priorities and those that are more sensitive to methodological choices. In participatory planning contexts, low alignment may signal that consensus on preferences does not translate into coherent performance patterns across alternatives, suggesting the need for further deliberation rather than immediate decision.

At the same time, the results should be interpreted with caution. The analysis is based on a limited number of empirical cases and relies on a simplified operationalization of preference–impact alignment using rank-based correlation measures. This approach captures relative correspondence between weights and impacts but does not account for magnitude differences or more complex forms of interaction. In addition, the transformation of the original hierarchical structure into a single-level representation, while necessary for comparability across methods, implies a loss of

structural information, which may affect the sensitivity of certain methods, particularly those that rely explicitly on hierarchical relationships.

Despite these limitations, the analysis provides a consistent set of patterns across cases, suggesting that the relationship between preference–impact alignment and ranking behaviour is not incidental. Rather than establishing causal relationships, the findings offer an empirical basis for interpreting variation across methods as a function of problem structure, contributing to ongoing debates on robustness, interpretation, and method selection in applied multicriteria decision-making. These insights also reinforce the importance of interpreting multicriteria outputs within the broader context of planning processes. In participatory settings, where decision-making involves negotiation, deliberation, and the reconciliation of diverse perspectives, divergent rankings may signal the need for further discussion rather than the selection of a "more correct" method. In this sense, multicriteria analysis should be understood not only as a decision-support tool, but also as a device that reveals underlying inconsistencies or misalignments in how problems are structured and understood.

Importantly, these results challenge the assumption that participatory processes necessarily lead to convergence in preferences. Evidence from the empirical analysis indicates that while interaction does not significantly alter individual priorities, it enhances trust, perceived relevance, and legitimacy of the decision process. This suggests that the value of participatory planning lies less in producing consensus and more in structuring legitimate decision processes. In this context, analytical tools should be understood not only as instruments for optimisation, but as devices that shape how collective preferences are interpreted, negotiated, and translated into policy choices.

6. Conclusion

This paper examined how different multicriteria decision-making methods produce varying rankings of alternatives when applied to the same decision problem. Moving beyond a purely methodological comparison, the analysis showed that these variations are closely related to the structure of the decision context, particularly the relationship between collective preference structures and the distribution of impacts across alternatives. In line with recent contributions that call for a more contextualized interpretation of multicriteria outcomes (Cinelli et al., 2022; Greco et al., 2025; Nurmi, 1999; Więckowski & Sałabun, 2023), the findings suggest that divergence across methods should not be understood solely as a consequence of their mathematical properties, but also as a reflection of how decision problems are structured.

The results indicate that higher levels of correspondence between preferences and impacts are associated with greater convergence across methods, whereas lower correspondence leads to increased divergence and greater sensitivity to aggregation logic. This does not imply the superiority of any particular method. Rather, it highlights that ranking outcomes emerge from the interaction between method design and problem structure, extending previous work that has primarily addressed issues of consistency, robustness, and method selection from a methodological perspective (Athawale & Chakraborty, 2012; Wątróbski et al., 2019).

In participatory and multilevel governance contexts, this is particularly relevant. Decision-making processes are shaped by institutional complexity, actor diversity, and cognitive constraints (Kooiman, 2003; Cairney & Weible, 2017). In such settings, the interpretation of analytical outputs becomes as important as their computation. Multicriteria approaches are increasingly used in strategic and participatory planning processes to support the evaluation and prioritization of alternatives (Bana e Costa et al., 2014; Dias et al., 2018; Keeney, 2004; Polatidis et al., 2006; Wolf et al., 2021). In such contexts, where decision-making involves multiple actors, competing objectives, and uncertain impacts, divergent rankings should not be interpreted as methodological

inconsistency or failure. Instead, they may signal underlying tensions between what stakeholders collectively value and how alternatives are defined and evaluated. In this sense, multicriteria outputs can support not only prioritization, but also critical reflection and deliberation about the adequacy of the decision problem itself. By introducing preference-impact alignment as an explicit analytical dimension, this paper contributes to bridging the gap between technical decision-support methods and the interpretative challenges of policy-making. It highlights the need for a more reflexive use of analytical tools, where results are understood in relation to the broader process of policy problem structuring.

The analysis is based on a limited number of empirical cases and relies on a rank-based operationalization of the relationship between preferences and impacts. While this approach captures relative correspondence, it does not account for magnitude differences or more complex interactions between criteria. In addition, the transformation of hierarchical structures into a single-level representation, required for comparability across methods, entails a loss of structural information. These limitations mean that the findings should be interpreted as analytically grounded rather than statistically generalizable. Nevertheless, the consistency of the observed patterns across the cases suggests that the relationship between preferences and impacts constitutes a relevant dimension for interpreting multicriteria outcomes in applied decision contexts.

These findings also point to a broader line of inquiry concerning inconsistencies in the aggregation of individual preferences. In particular, situations where preference–impact alignment is weak may reflect underlying tensions similar to those identified in the literature on aggregation paradoxes, where individually consistent judgments can lead to collectively inconsistent outcomes (Arrow, 1963; Diss & Kamwa, 2020; Gonzales & Perny, 2021; List, 2022; Nitzan, 2010; Nurmi, 1983, 1999; Zhang, 2020). In complex decision contexts, such inconsistencies may also be reinforced by cognitive limitations and heuristic-based reasoning.

Further research could therefore explore how different representations of preferences, such as ordinal versus ratio-based approaches, affect the relationship between individual judgments and collective outcomes (Saaty, 2013; Vasiljev, 2011). This would provide a complementary perspective on the role of aggregation procedures in shaping decision results, particularly in participatory planning contexts.

Overall, the paper contributes to bridging the gap between the technical development of multicriteria methods and their use in planning and policy processes. By showing that the meaning of ranking convergence and divergence depends not only on the methods applied but also on the structure of the decision problem, it supports a more context-sensitive and reflexive use of multicriteria analysis. In practical terms, the key question is not simply which method to apply, but how the relationship between preferences, impacts, and alternatives shapes the results that decision-makers are asked to interpret.

Data Availability Statement

The datasets and analytical materials supporting the findings of this study are provided as supplementary material submitted with the manuscript.

The supplementary material is organised into two blocks. The first block contains the anonymised aggregated preference structures derived from the participatory AHP exercises conducted in the three empirical cases. The second block contains the analytical workflow and calculations used for the comparison of multicriteria decision-making methods, including transformed decision

matrices, normalization procedures, multicriteria outputs, ranking comparisons, Spearman correlation analysis, and sensitivity analysis under equal-weight scenarios.

Individual-level participant responses are not shared in order to preserve participant confidentiality and anonymity.

Ethical Approval

This study involved participatory planning processes based on voluntary stakeholder participation. No sensitive personal data were collected, and all data were analysed in aggregated and anonymised form in accordance with ethical standards for social science research.

This article does not contain any studies with human participants performed by any of the authors.

Informed Consent

Informed consent was obtained from all participants involved in the participatory processes from which the data used in this study were derived. This article does not contain any studies with human participants performed by any of the authors.

7. References

Afshari, A., Mojahed, M., & Yusuff, R. M. (2010). Simple Additive Weighting approach to Personnel Selection problem. *International Journal of Innovation, Management and Technology*, *1*(5), 511–515. https://doi.org/10.1061/9780784413265.043

Athawale, V. M., & Chakraborty, S. (2012). Material selection using multi-criteria decision-making methods: A comparative study. *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications*, *226*(4), 266–285. https://doi.org/10.1177/1464420712448979

Azhar, N. A., Radzi, N. A. M., & Wan Ahmad, W. S. H. M. (2021). Multi-criteria Decision Making: A Systematic Review. *(Recent Advances in Electrical & Electronic Engineering (Formerly Recent Patents on Electrical & Electronic Engineering)*, *14*(8), 779–801. https://doi.org/10.2174/2352096514666211029112443

Bana e Costa, C. A., Lourenço, J. C., Oliveira, M. D., & Bana e Costa, J. C. (2014). A Socio-technical Approach for Group Decision Support in Public Strategic Planning: The Pernambuco PPA Case. *Group Decision and Negotiation*, *23*(1), 5–29. https://doi.org/10.1007/s10726-012-9326-2

Barbati, M., Figueira, J. R., Greco, S., Ishizaka, A., & Panaro, S. (2023). A multiple criteria methodology for priority based portfolio selection. *Socio-Economic Planning Sciences*, 101595. https://doi.org/10.1016/J.SEPS.2023.101595

Belton, V., & Stewart, T. J. (2002). *Multiple Criteria Decision Analysis*. Springer US. https://doi.org/10.1007/978-1-4615-1495-4

Borges, M., Marques, J., & Castro, E. (2020). Decision-making as a socio-cognitive process. In V. Chkoniya, A. Madsen, & P. Bukhrashvili (Eds.), *Anthropological Approaches to Understanding Consumption Patterns and Consumer Behavior* (pp. 382–403). IGI Global. https://doi.org/10.4018/978-1-7998-3115-0.ch021

Cinelli, M., Kadziński, M., Miebs, G., Gonzalez, M., & Słowiński, R. (2022). Recommending multiple criteria decision analysis methods with a new taxonomy-based decision support system. *European Journal of Operational Research*, *302*(2), 633–651. https://doi.org/10.1016/j.ejor.2022.01.011

Dias, L. C., Antunes, C. H., Dantas, G., de Castro, N., & Zamboni, L. (2018). A multi-criteria approach to sort and rank policies based on Delphi qualitative assessments and ELECTRE TRI: The case of smart grids in Brazil. *Omega (United Kingdom)*, *76*, 100–111. https://doi.org/10.1016/j.omega.2017.04.004

Dodgson, J. S., Spackman, M., & Pearman, A. D. (2009). *Multicriteria Analysis - A manual*. Department for Communities and Local Government: London. http://eprints.lse.ac.uk/12761/1/Multi-criteria_Analysis.pdf

Figueira, J., Greco, S., & Ehrgott, M. (2005). *Multicriteria decision analysis. State of the art surveys*. Springer Science + Business Media, Inc. Print.

Fischer, F., Miller, G., & Sidney, M. (2007). *Handbook of Public Policy: Theory, Politics, and Methods*. SAGE Publications Ltd. https://doi.org/10.4135/9781848608054

Flyvbjerg, B., & Richardson, T. (2002). Planning and Foucault: in search of the dark side of planning theory. In Allmendinger & Tewdwr-Jones (Eds.), *Planning Futures: New Directions for Planning Theory* (pp. 44–64). Routledge.

Greco, S., Słowiński, R., & Wallenius, J. (2025). Fifty years of multiple criteria decision analysis: From classical methods to robust ordinal regression. *European Journal of Operational Research*, *323*(2), 351–377. https://doi.org/10.1016/j.ejor.2024.07.038

Healey, P. (1997). Collaborative Planning. *Collaborative Planning*. https://doi.org/10.1007/978-1-349-25538-2

Hwang, C. L., & Yoon, K. (1981). *Multiple Attribute Decision Making. Methods and Applications. A State-of-the-Art survey* (Vol. 1). Springer-Verlag.

Innes, J. E., & Booher, D. E. (2010). *Planning with complexity. An introduction to collaborative rationality for public policy*. Routledge.

Jann, W., & Wegrich, K. (2007). Theories of the Policy Cycle. In F. Fischer, G. Miller, & M. Sidney (Eds.), *Handbook of Public Policy Analysis: Theory, Politics and Methods* (pp. 43–62). Taylor & Francis.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: heuristics and biases*. Cambridge University Press.

Keeney, R. L. (2004). Framing public policy decisions. *International Journal of Technology, Policy and Management*, *4*(2), 95–115. https://doi.org/10.1504/IJTPM.2004.004815

Kooiman, J. (2003). *Governing as Governance*. SAGE Publications.

Kumar, A., Sah, B., Singh, A. R., Deng, Y., He, X., Kumar, P., & Bansal, R. C. (2017). A review of multi criteria decision making (MCDM) towards sustainable renewable energy development. *Renewable and Sustainable Energy Reviews*, *69*(June 2016), 596–609. https://doi.org/10.1016/j.rser.2016.11.191

Majone, G. (2006). Agenda setting. In M. Moran, M. Rein, & R. Goodin (Eds.), *The oxford handbook of public policy* (pp. 228–250). Oxford University Press.

Marttunen, M., Lienert, J., & Belton, V. (2017). Structuring problems for Multi-Criteria Decision Analysis in practice: A literature review of method combinations. *European Journal of Operational Research*, *263*(1), 1–17. https://doi.org/10.1016/j.ejor.2017.04.041

Moscovici, S. (1988). Notes towards a description of Social Representations. *European Journal of Social Psychology*, *18*(3), 211–250. https://doi.org/doi:10.1002/ejsp.2420180303

Moscovici, S. (2001). Why a Theory of Social Representation? In K. Deaux & G. Philogène (Eds.), *Representations of the social: Bridging theoretical traditions* (pp. 8–35). Oxford: Blackwell.

Munda, G. (2004). Social multi-criteria evaluation: Methodological foundations and operational consequences. *European Journal of Operational Research*, *158*(3), 662–677. https://doi.org/10.1016/S0377-2217(03)00369-2

Nogueira, F., Borges, M., & Wolf, J.-H. (2017). Collaborative Decision-Making in Non-formal Planning Settings. *Group Decision and Negotiation*, *Special Issue on "Making decisions in open communities: collective actions in the public realm,"* 1–16. https://doi.org/10.1007/s10726-016-9518-2

Nurmi, H. (1999). *Voting Paradoxes and How to Deal with Them*. Springer-VerlagBerlin Heidelberg.

Patel, M., Vashi, M. P., & Bhatt, B. V. (2017). SMART- Multi-criteria decision-making technique for use in planning activities. *NHCE*.

Pinzon Amorocho, J. A., & Hartmann, T. (2022). A multi-criteria decision-making framework for residential building renovation using pairwise comparison and TOPSIS methods. *Journal of Building Engineering*, *53*. https://doi.org/10.1016/j.jobe.2022.104596

Polatidis, H., Haralambopoulos, D. A., Munda, G., & Vreeker, R. (2006). Selecting an appropriate multi-criteria decision analysis technique for renewable energy planning. *Energy Sources, Part B: Economics, Planning and Policy*, *1*(2), 181–193. https://doi.org/10.1080/009083190881607

Rhodes, R. a. W. (1996). The New Governance: Governing without Government. *Political Studies*, *44*(4), 652–667. https://doi.org/10.1111/j.1467-9248.1996.tb01747.x

Rosszkowska, E. (2011). Multi-Criteria Decision Making Models By Applying the Topsis Method To Crisp and Interval Data. *Multiple Criteria Decision Making*, *6*, 200–230. https://mcdm.ue.katowice.pl/index.php?page=publications&id=Rosz11

Roy, B. (1996). *Multicriteria methodology for decision aiding*. Springer, Boston, MA. https://doi.org/https://doi.org/10.1007/978-1-4757-2500-1

Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, *1*(1), 83. https://doi.org/10.1504/IJSSCI.2008.017590

Sabatier, P. (2007). *Theories of the Policy Process*. Westview Press.

Schneider, A., & Ingram, H. (2008). Social Constructions in the Study of Public Policy. In J. Holstein (Ed.), *Handbook of Constructionist Research* (pp. 189–211). New York: Guilford Publications.

Simon, H. A. (1960). *The New Science of Management Decision*. New York: Harper and Row.

Surel, Y. (2000). The role of cognitive and normative frames in policy-making. *Journal of European Public Policy*, *7*(4), 495–512. https://doi.org/10.1080/13501760050165334

Teles, F. (2016). Local governance and intermunicipal cooperation. In *Local Governance and Intermunicipal Cooperation*. https://doi.org/10.1057/9781137445742

Triantaphyllou, E. (2000). *Multi-criteria decision making methods: a comparative study* (P. M. Pardalos & D. Hearn (eds.)). Springer Science+Business Media, B.V.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Tzeng, G.-H., & Huang, Z.-J. (2011). *Multi Attribute Decision Making. Methods and Applications*. Taylor & Francis Group.

Wang, J.-J., Jing, Y.-Y., Zhang, C.-F., & Zhao, J.-H. (2009). Review on multi-criteria decision analysis aid in sustainable energy decision-making. *Renewable and Sustainable Energy Reviews*, *13*(9), 2263–2278. https://doi.org/10.1016/j.rser.2009.06.021

Wątróbski, J., Jankowski, J., Ziemba, P., Karczmarczyk, A., & Zioło, M. (2019). Generalised framework for multi-criteria method selection. *Omega (United Kingdom)*, *86*, 107–124. https://doi.org/10.1016/j.omega.2018.07.004

Więckowski, J., & Sałabun, W. (2023). Sensitivity analysis approaches in multi-criteria decision analysis: A systematic review. *Applied Soft Computing*, *148*(September), 110915. https://doi.org/10.1016/j.asoc.2023.110915

Wolf, J. H., Nogueira, F., & Borges, M. (2021). A Collaborative Methodology for Local Strategic Planning: Insights from Four Plans in Portugal. *Planning Practice and Research*, *00*(00), 1–17. https://doi.org/10.1080/02697459.2020.1755138

Zanakis, S. H., Doukidis, G., & Zopounidis, C. (2000). Decision Making: Recent Developments and Worldwide Applications. In S. H. Zanakis, G. Doukidis, & C. Zopounidis (Eds.), *Environmental Science & Technology* (Vol. 45, Issue 12). Springer. https://doi.org/10.1021/es5027079

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the authors used ChatGPT (OpenAI) to support language editing and improve clarity of expression. All content was critically reviewed, revised, and validated by the authors, who take full responsibility for the final version of the manuscript.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- OriginalWeightsCaseA.xlsx
- OriginalWeightsCaseB.xlsx
- MCDMCaseA102021nivel0.xlsx
- MCDMCaseB102021nivel0.xlsx
- OriginalWeightsCaseC.xlsx
- MCDMCaseC102021nivel0.xlsx