



# Comparative Analysis of Fuzzy Analytic Hierarchy Process and Interval Analytic Hierarchy Process for Multi-Stakeholder Evaluation of Public Transport Supply Quality

Hiba Solieman<sup>1,\*</sup>, Szabolcs Duleba<sup>1,2</sup>

<sup>1</sup> Budapest University of Technology and Economics, Department of Transport Technology and Economics, Budapest, Hungary

<sup>2</sup> University of Nyíregyháza, Institute of Mathematics and Informatics, Nyíregyháza, Hungary

## ARTICLE INFO

### Article history:

Received 29 Novembre 2025

Received in revised form 26 January 2026

Accepted 14 February 2026

Available online 16 February 2026

### Keywords:

Multi-Criteria Decision Making (MCDM), Fuzzy Analytic Hierarchy Process (FAHP), Interval Analytic Hierarchy Process (IAHP), Stakeholder Analysis, Public Transport Quality, Kendall's Coefficient of Concordance (W), Urban Mobility

## ABSTRACT

Evaluating public transport service quality requires balancing diverse stakeholder perspectives, a complex task essential for sustainable urban mobility. The efficiency of Multi-Criteria Decision-Making (MCDM) techniques, such as the Fuzzy Analytic Hierarchy Process (FAHP), in fostering consensus among diverse groups is still up for debate. This study compares FAHP and Interval AHP (IAHP) to determine which method better achieves stakeholder agreement in public transport quality assessment. Using a three-level criteria hierarchy, data were collected from regular passengers, potential passengers, and government officials in Mersin, Turkey. Both FAHP and IAHP derived priority weights, with consensus measured by Kendall's coefficient of concordance (W). Results revealed that consensus was strongest for strategic goals and weakened for operational details. Notably, IAHP consistently outperformed FAHP, producing higher concordance across all levels. The findings demonstrate that IAHP is a more effective tool for reconciling divergent preferences, offering transport planners a superior method for fostering legitimate, consensus-driven policies and improvements.

## 1. Introduction

Urban mobility systems are integral to sustainable city development, with public transport playing a pivotal role in ensuring accessibility, reducing congestion, and promoting environmental quality. Also, they are fundamental to economic vitality, social equity, and environmental sustainability [1]. Assessing the quality of public transportation's supply is intrinsically complicated; it goes beyond simple technical performance measurements to take into account the various and frequently divergent opinions of its customers, future users, and providers [2]. Multi-Criteria Decision-Making (MCDM) methodologies have emerged as powerful tools to navigate this complexity, providing structured frameworks to evaluate qualitative and quantitative factors simultaneously [3].

\* Corresponding author.

E-mail address: [hiba.solieman@edu.bme.hu](mailto:hiba.solieman@edu.bme.hu)

<https://doi.org/10.59543/a5ctzq55>

Among these, the Analytic Hierarchy Process (AHP), which was first presented by Saaty [4], has been extensively used in transportation planning due to its capacity to break down complicated issues and determine ratio-scale priorities. The Fuzzy Analytic Hierarchy Process (FAHP) was created by integrating fuzzy set theory [5] with AHP to address the inherent subjectivity and ambiguity in human judgment. FAHP effectively captures linguistic vagueness (e.g., "slightly more important") through fuzzy numbers, making it a popular choice for transport service evaluation [6]. More recently, the Interval Analytic Hierarchy Process (IAHP) has been proposed as an alternative, representing uncertainty through interval judgments rather than fuzzy sets, which some argue offers greater transparency in preserving the range of expert opinions [7].

A critical yet underexplored dimension of transport planning is the multi-stakeholder nature of decision-making. The success of any public transport system hinges on its ability to meet the needs of a heterogeneous group of stakeholders, including regular passengers, potential passengers, and government authorities [8]. Each group possesses unique priorities: regular users value reliability and comfort, potential users are influenced by accessibility and cost, while government planners must balance user satisfaction with fiscal constraints and regulatory requirements. Consequently, a significant challenge lies not only in incorporating these diverse perspectives but also in measuring the degree of consensus among them. Without a robust measure of agreement, MCDM results may reflect a false compromise, leading to policies that lack broad support and legitimacy [9].

While numerous studies have applied FAHP and IAHP in transportation contexts [10], a conspicuous gap exists in the literature concerning the comparative efficacy of these methods in building consensus among fundamentally different stakeholder groups. It remains unclear whether the "smoothing" effect of FAHP's defuzzification or the "transparent" representation of IAHP's intervals is more effective in harmonising divergent preferences. This gap is critical because the choice of method itself may significantly influence the perceived consensus and, ultimately, the policy recommendations derived from the analysis.

This study, therefore, seeks to address this research gap by conducting a rigorous comparative analysis of FAHP and IAHP within a multi-stakeholder framework. We employ Kendall's coefficient of concordance ( $W$ ) not merely as a secondary statistic but as a primary metric to evaluate the consensus-building performance of each MCDM method. The investigation is grounded in an empirical case study of Mersin, Turkey, evaluating its public bus transport supply quality through the lenses of three distinct stakeholder groups.

The three stakeholder groups were selected based on the representative principle:

- Regular Passengers (RP) ( $\geq 12$  monthly trips),
- Potential Passengers (PP) (non-users open to using public transport if service improves), and
- Government Officials (Gov) from the Mersin Transportation Department.

This division addresses both retention and attraction strategies, ensuring policy-relevant insights across user experience, latent demand, and institutional feasibility.

The research is guided by two central objectives:

- To evaluate the relative importance of public transport supply quality criteria from the perspectives of regular passengers, potential passengers, and government officials.
- To determine which MCDM method, FAHP or IAHP, provides a more efficient and reliable approach for achieving and measuring consensus among these heterogeneous evaluator groups.

We apply FAHP and IAHP to a hierarchical framework comprising three levels: Level 1 (broad dimensions: Service Quality, Transport Quality, Tractability), Level 2 (11 sub-criteria as tactical indicators, such as Accessibility, Directness, Time Availability, Speed, Reliability), and Level 3 (10 operational indicators, such as Stops Safety, Need for Transfer). To assess inter-group consensus, Kendall's Coefficient of Concordance (W) was calculated separately for each combination of method and hierarchical level.

The study's conclusions make theoretical and practical contributions. Theoretically, we provide evidence-based insights into the performance of different uncertainty-handling methods in MCDM under conditions of stakeholder heterogeneity. Practically, we offer policymakers a validated methodological framework for participatory transport planning and specific, actionable recommendations for improving the public transport system in Mersin.

Following this introduction, the paper is structured as follows. A comprehensive literature review examines the evolution of public transport quality metrics, the application of MCDM methods, and the critical role of multi-stakeholder perspectives. The methodology section details the case study context, stakeholder selection process, criteria hierarchy development, and the computational steps for FAHP, IAHP, and Kendall's W. The results section presents the priority weights and consensus levels for each hierarchical level, comparing the outcomes of both methods. The discussion interprets these findings, explores their theoretical and practical implications, and addresses study limitations. Finally, the conclusion summarises the key findings, underscores the paper's contributions, and provides actionable recommendations for policymakers and researchers.

## **2. Literature Review**

The pursuit of sustainable urban mobility has placed public transport at the forefront of urban policy, making the robust evaluation of its service quality a critical academic and practical endeavour. This review synthesises the literature across four thematic areas: the conceptualisation of public transport service quality, the role of Multi-Criteria Decision-Making (MCDM) frameworks, the imperative of multi-stakeholder engagement, and the evolution of methods to handle evaluative uncertainty.

### *2.1. Conceptualising and Measuring Public Transport Service Quality*

The assessment of public transport has evolved from purely efficiency-based metrics (e.g., vehicle load factors, cost per kilometre) towards a richer understanding of service quality from the user's perspective [11]. This shift acknowledges that ridership is not only a function of availability but also of perceived value and experience [12]. Early frameworks, such as the European Standard EN 13816, defined quality in public transport through multiple dimensions, including availability, accessibility, and customer care, providing a structured yet broad basis for evaluation [13].

Empirical research has consistently demonstrated that the importance of service attributes is not uniform. For example, [14] identified reliability, frequency, and travel time as paramount across various cultural contexts. Conversely, factors like aesthetics often rank lower in importance, though they can significantly influence the overall perception of service [5]. This establishes that service quality is a latent construct, best measured through a hierarchy of interlinked criteria whose weights must be determined empirically, not assumed a task for which MCDM methods are particularly well-suited.

## 2.2. MCDM Frameworks in Transportation Evaluation

Multiple-Criteria Decision-Making (MCDM) methodologies have been widely adopted in the transport sector to navigate complex choices involving operational, environmental, and social trade-offs. A comprehensive review by Yannis et al. (2020) highlights that MCDM methods, especially AHP, are the most commonly used in transport-related evaluations due to their structured approach to conflicting criteria [15]. Similarly, sustainable transport planning frequently leverages comparative MCDM frameworks including FAHP, VIKOR, and TOPSIS to balance multi-dimensional objectives [16].

AHP's strength lies in its use of pairwise comparisons to derive precise priority weights, making it a popular choice for transport studies. It has been effectively used to select public transport technologies [17], evaluate transit policy alternatives [18], and prioritise infrastructure projects [19]. However, a recurring critique of standard AHP is its reliance on crisp numerical values, which struggles to capture the imprecise and linguistic nature of human judgments about quality [20]. This limitation has driven the adoption of extensions that can model uncertainty more effectively.

## 2.3. Fuzzy and Interval Extensions of AHP

The Fuzzy Analytic Hierarchy Process (FAHP) enhances classical AHP by accounting for the ambiguity of human judgment through fuzzy set theory. This extension is valuable for linguistic assessments such as "highly important" or "slightly less important," although it risks smoothing over stakeholder differences during defuzzification [21], [22]. The Interval Analytic Hierarchy Process (IAHP), in contrast, represents weights as intervals to reflect uncertainty and divergence among evaluators a capability that preserves stakeholder-specific opinions during aggregation [23].

## 2.4. The Imperative of Multi-Stakeholder Perspectives

A transport system serves a diverse ecosystem of users and providers, each with distinct priorities. Regular commuters may prioritise predictability and speed [24], while occasional users might value comprehensibility and cost [25, p.]. Meanwhile, authorities and operators must balance user satisfaction with financial viability, regulatory compliance, and environmental goals [26].

Failing to account for this plurality can lead to technically sound projects that fail socially due to a lack of legitimacy [27]. This has spurred the development of participatory MCDM frameworks. The Multi-Actor Multi-Criteria Analysis (MAMCA) methodology, for instance, is explicitly designed to showcase how different stakeholder groups evaluate projects based on their own value systems [28]. Studies applying such approaches have revealed significant divergences in priorities; for example, residents may prioritise environmental impacts, while businesses focus on accessibility [29]. This underscores a key challenge: the need for methodologies that not only capture these divergent preferences but also possess the analytical rigour to quantify the level of consensus or discord among groups, an area where statistical measures like Kendall's W are underutilised.

## 2.5. Advancements in Handling Uncertainty: Fuzzy and Interval Approaches

The integration of fuzzy set theory [30] with MCDM marked a significant advancement in capturing the vagueness of subjective judgments. Fuzzy AHP (FAHP) replaces crisp numbers in pairwise comparisons with fuzzy numbers (e.g., triangular, trapezoidal) to model linguistic assessments [31]. This approach has been widely adopted in transportation, such as for evaluating traffic safety quality [32] and selecting sustainable urban transport alternatives [33].

However, a noted complication in FAHP is the choice of defuzzification method, which can influence the final ranking of alternatives and potentially introduce bias [34]. As an alternative, Interval AHP (IAHP) represents uncertainty through interval numbers, preserving a range of possible values for each weight [35]. This method offers superior transparency for group decision-making, as

it maintains the diversity of opinions without compressing them into a single value through an irreversible defuzzification step [36]. While applied in fields like environmental risk assessment [37] and supplier selection [38], its application in comparative multi-stakeholder public transport evaluations remains scarce.

### 2.6. Synthesis and Identification of the Research Gap

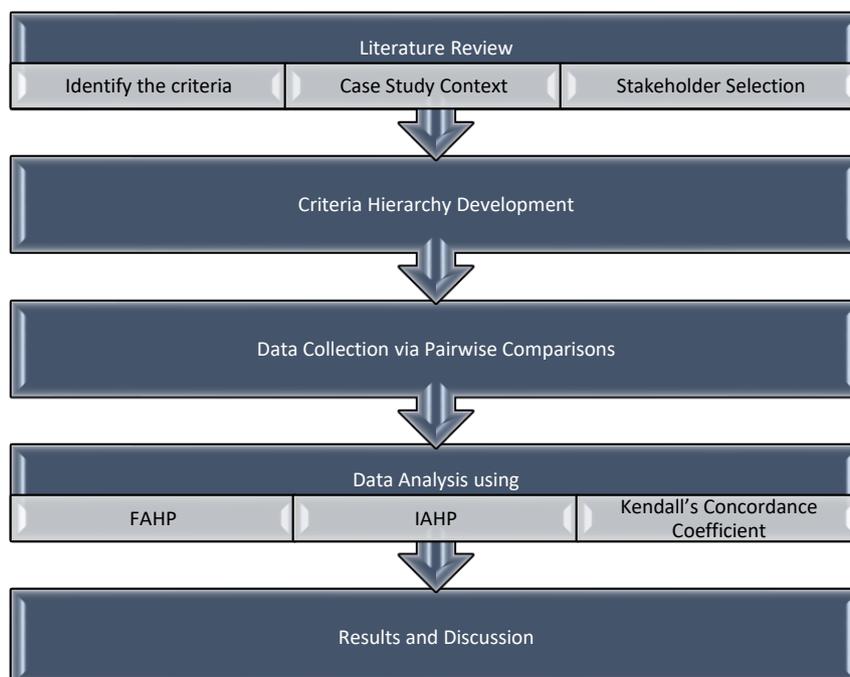
The extant literature confirms that:

1. Evaluating public transport quality requires a multi-dimensional framework.
2. MCDM methods, particularly AHP and its derivatives, provide the necessary structured approach.
3. Capturing the perspectives of multiple stakeholders is essential for legitimate and sustainable outcomes.
4. Uncertainty in human judgment necessitates methods like FAHP or IAHP.

Despite this, a critical gap persists. There is a lack of direct comparative studies that evaluate the efficacy of FAHP versus IAHP specifically within a multi-stakeholder public transport context. Furthermore, the success of such participatory approaches hinges on achieving consensus, yet there is a paucity of research that uses robust statistical measures of agreement (e.g., Kendall's W) to evaluate which MCDM method is more effective at reconciling divergent preferences. This study aims to fill this gap by conducting a rigorous head-to-head comparison of FAHP and IAHP, using Kendall's W as a key performance metric, with incorporating three stakeholder groups (regular users, potential users, and officials), to determine the superior method for achieving consensus in evaluating public transport supply quality.

## 3. Methodology

The effectiveness of the Fuzzy Analytic Hierarchy Process (FAHP) and the Interval AHP (IAHP) in reaching agreement among various stakeholder groups assessing the quality of public transportation supply is compared in this study using a multi-method research approach. The methodological framework is illustrated in Fig. 1.



**Fig. 1. Research Workflow**

**3.1 Case Study Context and Stakeholder Selection**

The study was conducted in Mersin, Turkey, a representative Mediterranean port city with a population exceeding one million. Its public transport system, primarily bus-based, faces typical challenges of a growing city, including congestion, increasing demand, and the need for service quality improvements.

Adopting a multi-stakeholder approach is critical for legitimate and sustainable transport planning [39]. Participants were recruited based on the representative principle to ensure the three primary perspectives of the public transport ecosystem were included:

- Regular Passengers (RP; n=50): Individuals using public transport ≥3–4 days/week (≥12 trips/month) over the preceding three months. This threshold captures habitual users with direct, recent experience of service quality.
- Potential Passengers (PP; n=50): Individuals using public transport <1 time per week but expressing a willingness to increase usage if service improved. This group provides insights into barriers to adoption.
- Government Officials (GOV; n=10): Professionals from the Mersin Municipality Transportation Department involved in policy-making, strategic planning, or project evaluation. Their perspective ensures practical constraints and regulatory frameworks are considered.

This structure ensures the triangulation of demand-side (users/non-users) and supply-side (planners) perspectives, mitigating the bias of consulting a single homogenous group.

**3.2 Criteria Hierarchy Development**

A three-level hierarchical model was developed to structure the evaluation of public transport supply quality, moving from strategic goals to operational indicators (Table 1). The model was constructed based on a comprehensive review of existing service quality frameworks and adapted to the local context of Mersin.

**Table 1**  
 The Three-Level AHP Hierarchy for Evaluation

Level	Category	Description	Criteria
1	Strategic Goals	Overarching objectives of the transport system	Service Quality, Transport Quality, Tractability
2	Performance Dimensions	Key areas of service delivery	Accessibility, Directness, Time Availability, Speed, Reliability, Perspicuity, Pre-Trip Information, En-Route Information, Physical Comfort, Mental Comfort, Travel Safety
3	Operational Indicators	Specific, measurable attributes	Frequency of Lines, Limited Time of Use, Journey Time, Awaiting Time, Time to Reach Stops, Directness to stops, Stops Safety, Comfort in Stops, Need for Transfer, Fit Connection

Each indicator was explicitly defined for respondents during survey administration to ensure a common understanding (e.g., "Reliability" was defined as punctuality and consistency of service).

### 3.3 Data Collection: Pairwise Comparison Surveys

Data was collected through structured questionnaires designed for pairwise comparisons. For each level of the hierarchy, respondents were asked to compare criteria in pairs and judge their relative importance using the fundamental Saaty scale [40].

To capture uncertainty in judgments, the surveys were designed for both FAHP and IAHP data collection:

- For FAHP, participants used linguistic terms (e.g., "equally important," "moderately more important") which were later converted to triangular fuzzy numbers (e.g., (2, 3, 4) for "moderately more important").
- For IAHP, participants provided their judgments by giving a range (lower and upper bound) for the relative importance of each criterion pair, directly creating interval comparison matrices.

The surveys were administered in person to ensure clarity and to achieve a high response rate. The consistency of each respondent's pairwise comparisons was checked using the Consistency Ratio (CR) [40]; responses with a  $CR > 0.10$  were excluded from the analysis to ensure data reliability.

### 3.4 Data Analysis

#### 3.4.1. FAHP Calculations

Fuzzy logic and fuzzy set theory provide a mathematical framework that closely reflects human reasoning in situations of vagueness and partial truth. Instead of the strict true/false assignment of classical Boolean logic, fuzzy sets allow elements to belong to a set with a graded membership between 0 and 1; as a result, an element can partially belong both to a set and to its complement to different degrees. Zadeh's seminal paper on fuzzy sets formalised this idea and showed how basic operations such as union, intersection and complement can be extended to classes with continuous membership functions [30]. Since then, the fuzzy set framework has been widely applied to address the limitations of traditional set and logic theory when dealing with vagueness. Because many geographic or spatial phenomena are described using imprecise or linguistic terms, fuzzy set theory is often integrated with MCDA and machine-learning models for spatial problem-solving [41]. Numerous studies have integrated fuzzy sets with AHP and other MCDA techniques in GIS environments to model landslide susceptibility, natural hazard risk, and broader spatial planning problems, showing that fuzzy-AHP combinations can yield more reliable and nuanced maps than crisp approaches [42].

In fuzzy AHP applications, the preferences of decision makers are the primary source for determining the relative importance (weights) of criteria. These preferences are usually elicited as linguistic variables (such as "equally important", "moderately more important", or "extremely more important") rather than as precise numbers. To process these qualitative judgments, fuzzy sets are used to map each linguistic term to a corresponding fuzzy number via an appropriate membership function. Triangular fuzzy numbers are especially popular, because they provide a simple yet flexible way to capture the lower, most likely and upper values associated with a verbal judgment [43]. The resulting fuzzy pairwise comparison matrices are then used to derive fuzzy weights, which better reflect the uncertainty and subjectivity in human judgments. This type of fuzzy-AHP framework has been successfully used in domains such as landslide susceptibility mapping, site-selection problems and sustainable transport planning, where spatial criteria and stakeholder preferences are inherently uncertain and often expressed linguistically.

The linguistic judgments are then mapped into fuzzy numbers via suitable membership functions. In the referenced study, a triangular fuzzy number (TFN) representation is used for converting linguistic variables into quantitative values [44].

Certain linguistic factors serve as qualitative representations of these judgments. At this point, the judgments must be quantified using a fuzzy set and the corresponding membership function. In this investigation, the linguistic variables were converted to quantitative values using a triangular fuzzy set (see Fig. 2). Table 2 illustrates the connection between language characteristics and quantitative values.

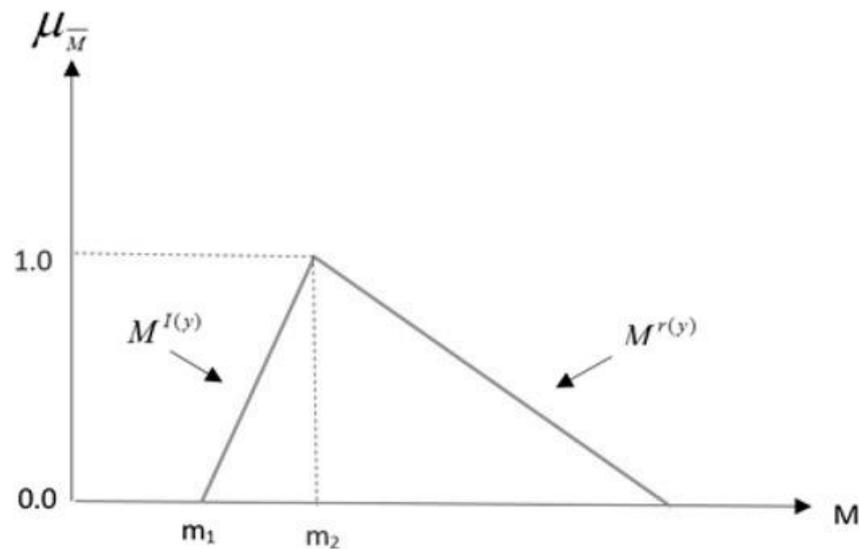


Fig. 2. Triangular fuzzy membership function.

Table 2.  
 Triangular fuzzy parameters corresponding to linguistic variables.

Linguistic Variables	Triangular Fuzzy Numbers	Reciprocal Triangular Fuzzy Numbers
Extremely strong	(9, 9, 9)	(1/9, 1/9, 1/9)
Very strong	(6, 7, 8)	(1/8, 1/7, 1/6)
Strong	(4, 5, 6)	(1/6, 1/5, 1/4)
Moderately strong	(2, 3, 4)	(1/4, 1/3, 1/2)
Equally strong	(1, 1, 1)	(1, 1, 1)
Intermediates	(7, 8, 9), (5, 6, 7), (3, 4, 5), (1, 2, 3)	(1/9, 1/8, 1/7), (1/7, 1/6, 1/5), (1/5, 1/4, 1/3), (1/3, 1/2, 1)

Fuzzy triangular numbers are used to create pairwise comparison (PC) matrices. Assume the following fuzzy pairwise comparison matrix:

$$\tilde{A} = \begin{bmatrix} (1,1,1) & (l_{12}, m_{12}, u_{12}) & \dots & (l_{1n}, m_{1n}, u_{1n}) \\ (l_{21}, m_{21}, u_{21}) & (1,1,1) & \dots & (l_{2n}, m_{2n}, u_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ (l_{n1}, m_{n1}, u_{n1}) & (l_{n2}, m_{n2}, u_{n2}) & \dots & (1,1,1) \end{bmatrix} \quad (1)$$

where each element must meet the condition that if  $\tilde{a} = (l_{ij}, m_{ij}, u_{ij})$  then  $\tilde{a}_{ij}^{-1} = (\frac{1}{l_{ij}}, 1/m_{ij}, 1/u_{ij})$  for  $1, j, \dots, n$  and  $i \neq j$ .

As a result, the fuzzy numbers can be used to specify every element of the matrix. Equation (2) was utilized for the normalized summing of the matrix's rows using the triangular fuzzy numbers:

$$S_i = \sum_{j=1}^m M_{gi}^j \otimes \left[ \sum_{i=1}^n \sum_{j=1}^m M_{gi}^j \right]^{-1} \tag{2}$$

where  $\otimes$  is the extended multiplication of two fuzzy triangular numbers and  $i$  and  $j$  were the number of rows and columns, respectively. The triangular fuzzy numbers in paired matrices are denoted by  $M_{gi}^j$ . The following formulas can be used to get the values of  $[\sum_{i=1}^n \sum_{j=1}^m M_{gi}^j]^{-1}$ :

$$\sum_{j=1}^m M_{gi}^j = \left( \sum_{j=1}^m l_j, \sum_{j=1}^m m_j, \sum_{j=1}^m u_j \right) \tag{3}$$

and

$$\sum_{i=1}^n \sum_{j=1}^m M_{gi}^j = \left( \sum_{i=1}^n l_j, \sum_{i=1}^n m_j, \sum_{i=1}^n u_j \right) \tag{4}$$

$$\left[ \sum_{i=1}^n \sum_{j=1}^m M_{gi}^j \right]^{-1} = \left( \frac{1}{\sum_{i=1}^n u_i}, \frac{1}{\sum_{i=1}^n m_i}, \frac{1}{\sum_{i=1}^n l_i} \right) \tag{5}$$

where the first through third components of fuzzy numbers are denoted by  $l_i, m_i$  and  $u_i$ . Equation (8) defines the degree of possibility  $M_1$  in relation to  $M_2$  if  $M_1 = (l_1, m_1, u_1)$  and  $M_2 = (l_2, m_2, u_2)$  are regarded as two triangular fuzzy numbers (see Figure 3) [45]:

$$V(M_2 \geq M_1) = hgt(M_1 \cap M_2) = \mu_{M_2}(d) \begin{cases} 1 & \text{if } m_2 \geq m_1 \\ 0 & \text{if } l_1 \geq u_2 \\ \text{otherwise} & \frac{l_1 - u_2}{(m_2 - u_2) - (m_1 - l_1)} \end{cases} \tag{6}$$

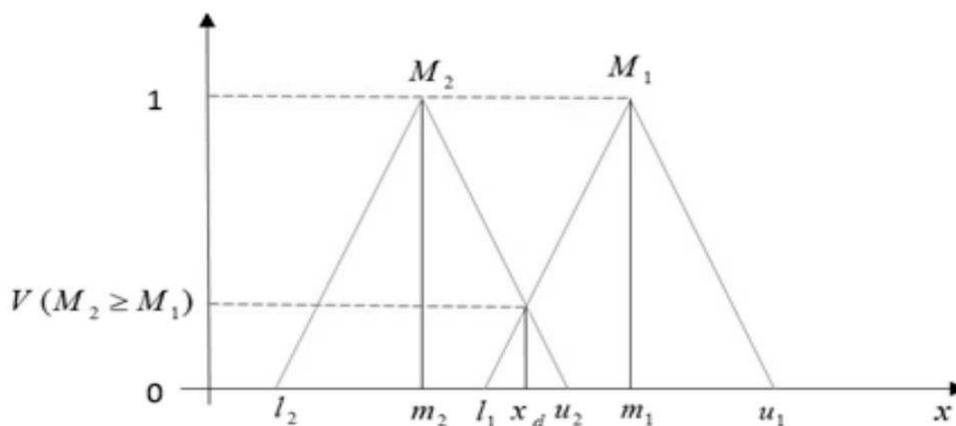


Fig. 3. The degree of possibility  $V(M_2 \geq M_1)$ .

On the other hand, the following equation's k fuzzy number makes a triangular number possible:  

$$V(M \geq M_1, M_2, \dots, M_k) = V[(M \geq M_1) \text{ and } (M \geq M_2) \text{ and } \dots \text{ and } (M \geq M_k)]$$

$$= \text{Min}V(M \geq M_i), i = 1, 2, \dots, k \quad (7)$$

Equation (8) can be used to determine the weight of the criterion and alternatives in the paired comparative matrices:

$$d'(A_i) = \text{Min}V(S_i \geq S_k) \quad k = 1, 2, \dots, n, k \neq i \quad (8)$$

Consequently, the normalised vector will look like this:

$$W' = (d'(A_1), d'(A_2), \dots, d'(A_n))^T \quad A_i (i = 1, 2, \dots, n) \quad (9)$$

Equation (10) was used to determine the final weight vector:

$$W = (d(A_1), d(A_2), \dots, d(A_n))^T \quad (10)$$

### 3.4.2. Interval AHP Calculations

The use of the Analytic Hierarchy Process (AHP) in decision-making can become problematic when experts provide heterogeneous or uncertain judgments, especially in systems where multiple stakeholder groups interact. Interval AHP (IAHP) offers a suitable enhancement to the classical AHP framework because it allows pairwise comparisons to be expressed as ranges rather than fixed numerical values. This interval representation captures uncertainty, variability, and disagreement among experts more realistically than single crisp judgments. As noted in methodological studies, IAHP is particularly effective when the decision context involves multiple domains of expertise or when subjective perceptions differ substantially among evaluators (e.g., planners, operators, policymakers) [46].

For complex public systems such as urban bus transportation, stakeholder diversity is inherent. Transport authorities, operators, passengers, and urban planners often prioritise different criteria, reflecting the multifaceted nature of service quality and system performance. Such diversity is typical in large metropolitan environments, where the transport network evolves as a complex system with emergent properties that arise from socio-economic, technical, and behavioural interactions [47]. Under these conditions, classical AHP which relies on single, unified comparison matrices struggles to integrate divergent viewpoints. Attempting to aggregate all stakeholder judgments into a single crisp matrix often leads to high inconsistency ratios (CRs), unstable eigenvalues, and weights that fail to represent stakeholder preferences adequately [48].

Moreover, even when only one stakeholder group is considered, variations in individual judgments typically produce pairwise matrices with distinct eigenvalues, limiting the reliability of a conventional AHP evaluation. IAHP overcomes these shortcomings by permitting each comparison to be expressed as an interval that reflects the range of stakeholder opinions, allowing the final prioritisation to incorporate uncertainty explicitly. This makes IAHP particularly suitable for analysing transport-related decision problems where expert opinions differ and where system complexity requires flexible, uncertainty-aware modelling approaches [10].

We employed interval pairwise comparison matrices in the AHP computation procedure to get over these drawbacks of the traditional AHP.

As a result, numerical intervals were used for any pairwise comparison matrix.

An interval's lower bound was defined as its lower value, and its upper bound as its higher value.

Let  $a$  serve as the interval's lower limit, while  $b$  corresponds to its upper limit; thus, it holds that  $a \leq$

b. The range utilized in the pairwise comparison matrices can then be represented by the interval, which can be written as  $x = [a, b] = \{y \in \mathbb{R}: a \leq y \leq b\}$ .  $\mathbb{IR}$  stands for the collection of all such intervals. Binary operations, such as multiplication, can be readily defined on  $\mathbb{IR}$  under this paradigm. If  $A$  is a pairwise comparison matrix with intervals as its elements, then  $A = (r_{ij})_{n \times n} \subset X \times X$  as (11):

$$\begin{cases} r_{ij} \times r_{ji} = 1, \forall i, j = 1, 2, \dots, n \\ r_{ii} = 1 \\ r_{ij} \geq 0 \end{cases} \quad (11)$$

is regarded as consistent when it satisfies the transitivity condition stated in Equation (12):

$$r_{ij} = r_{ik}/r_{jk} \quad \forall i, j, k = 1, 2, \dots, n \quad (12)$$

This shows that the criterion  $x_i$  is rated between  $a_{ij}$  and  $b_{ij}$  times more desirable than the criterion  $x_j$  for an interval represented as  $x = [a_{ij}, b_{ij}]$ . Thus, Equation (13) can be used to create the relevant interval-based pairwise comparison matrix:

$$A = \begin{bmatrix} 1 & [a_{12}, b_{12}] & \dots & [a_{1i}, b_{1i}] & \dots & [a_{1j}, b_{1j}] & \dots & [a_{1n}, b_{1n}] \\ [a_{21}, b_{21}] & 1 & \dots & [a_{2i}, b_{2i}] & \dots & [a_{2j}, b_{2j}] & \dots & [a_{2n}, b_{2n}] \\ \dots & \dots & 1 & \dots & \dots & \dots & \dots & \dots \\ [a_{i1}, b_{i1}] & [a_{i2}, b_{i2}] & \dots & 1 & \dots & [a_{ij}, b_{ij}] & \dots & [a_{in}, b_{in}] \\ \dots & \dots & \dots & \dots & 1 & \dots & \dots & \dots \\ [a_{j1}, b_{j1}] & [a_{j2}, b_{j2}] & \dots & [a_{ji}, b_{ji}] & \dots & 1 & \dots & [a_{jn}, b_{jn}] \\ \dots & \dots & \dots & \dots & \dots & \dots & 1 & \dots \\ [a_{n1}, b_{n1}] & [a_{n2}, b_{n2}] & \dots & [a_{ni}, b_{ni}] & \dots & [a_{nj}, b_{nj}] & \dots & 1 \end{bmatrix} \quad (13)$$

Given that  $a_{ij} \leq b_{ij}$  and both  $a_{ij}$  and  $b_{ij} \geq 0$  for all  $i, j = 1, 2, \dots, n$ , the matrix  $A$  qualifies as a reciprocal and well-defined matrix, as outlined in Equation (14):

$$a_{ij} = \frac{1}{b_{ji}}, b_{ij} = \frac{1}{a_{ji}}, \forall i, j = 1, 2, \dots, n \quad (14)$$

The matrix  $A$  must be reciprocal in order to evaluate the consistency ratio (CR) of this matrix and continue with the ensuing AHP calculations. It may be broken down into two clear pairwise comparison matrices under this condition:  $P = (p_{ij})_{n \times n}$  and  $Q = (q_{ij})_{n \times n}$ . Equation (15) defines the transformation  $D(\alpha)$ , which can then be created using these matrices:

$$D(\alpha) = (d_{ij}(\alpha))_{n \times n} = (p_{ij}^\alpha q_{ij}^{1-\alpha})_{n \times n} \quad \forall \alpha \in [0, 1] \quad (15)$$

Every entry of the matrix  $D(\alpha)$  represents a convex combination of the corresponding elements in matrices  $P$  and  $Q$ . Furthermore,  $D(\alpha)$  forms a monotonic and continuous matrix-valued function with respect to  $\alpha$ . As a result,  $D(1) = P$  and  $D(0) = Q$ . For any  $\alpha \in [0, 1]$  and for all  $i, j = 1, 2, \dots, n$ , each element  $d_{ij}(\alpha)$  lies within the interval  $[a_{ij}, b_{ij}]$ , allowing  $D(\alpha)$  to be interpreted as an interval-based reciprocal pairwise comparison matrix for all  $\alpha$  in this range. Consequently, matrix  $A$  is reciprocal, and the interval bounds  $a_{ij}$  and  $b_{ij}$  are non-negative and satisfy the relationships  $a_{ij} \leq b_{ij}$ ,  $a_{ij} = 1/b_{ij}$ , and  $b_{ij} = 1/a_{ij}$ . For such an interval reciprocal matrix  $A$  to be considered acceptably consistent, one of the following four conditions must hold: (i)  $CR_P = 0$  and  $CR_Q = 0$ ; (ii)  $CR_P = 0$

and  $0 < CR_Q < 0.1$ ; (iii)  $0 < CR_P < 0.1$  and  $CR_Q = 0$ ; or (iv)  $0 < CR_P < 0.1$  and  $0 < CR_Q < 0.1$ , where  $CR_P$  and  $CR_Q$  denote the consistency ratios of matrices  $P$  and  $Q$ , respectively. If none of these conditions is satisfied, the matrix  $A$  is deemed inconsistent.

To determine the interval-based relative weights, let  $w(\alpha)$  denote the weight vector obtained from the matrix  $D(\alpha)$  using the geometric mean approach, as presented in Equation (16).

$$w_i(\alpha) = \left( \prod_{j=1}^n d_{ij}(\alpha) \right)^{\frac{1}{n}} \quad i = 1, 2, \dots, n \text{ and } \alpha \in [0, 1] \quad (16)$$

This condition holds when  $\prod_{i=1}^n w_i(\alpha) = 1$ . Equation (2) can then be applied in defining  $d_{ij}(\alpha)$ . Then

$$\begin{aligned} w_i w_i(\alpha!) &= \left( \prod_{j=1}^n p_{ij}^\alpha q_{ij}^{(1-\alpha)} \right)^{\frac{1}{n}} = \left( \left[ \prod_{j=1}^n p_{ij}(\alpha) \right]^{\frac{1}{n}} \right)^\alpha \left( \left[ \prod_{j=1}^n q_{ij}(\alpha) \right]^{\frac{1}{n}} \right)^{(1-\alpha)} \\ &= w_i^\alpha(P) w_i^{(1-\alpha)}(Q) \end{aligned}$$

where  $w(P)$  and  $w(Q)$  represent the weighting vectors corresponding to matrices  $P$  and  $Q$ , respectively. The vector  $w(\alpha)$  is then used to construct the interval weights  $w_i$  as follows:

$$w_i = [\underline{w}, \bar{w}], \underline{w} = \min\{w_i(\alpha) \mid \alpha \in [0, 1]\} \text{ and } \bar{w} = \max\{w_i(\alpha) \mid \alpha \in [0, 1]\}$$

As for any  $\alpha \in [0, 1]$ ,  $w_i(\alpha)$  (in which  $i = 1, 2, \dots, n$ ), the  $\underline{w}$  and  $\bar{w}$  can be defined as follows [16,20]:

$$w_i = [\min\{w_i(P), w_i(Q)\}, \max\{w_i(P), w_i(Q)\}]$$

A straightforward statistical sampling technique, the MCS was then applied to derive the final weights. Using the interval weights  $w_i = [\underline{w}, \bar{w}]$  for  $i = 1, 2, \dots, n$ , MCS generates repeated samples within these bounds to estimate the resulting distribution of weights. As MCS is widely employed for analysing complex systems, it enables the model to capture variability and uncertainty effectively. The average value obtained from the simulated weight distribution for each interval was subsequently taken as the final weight, representing the mean of  $\underline{w}$  and  $\bar{w}$  over all simulations.

### 3.4.3. Measuring Consensus: Kendall's Coefficient of Concordance (W)

To address the core research objective of measuring consensus, Kendall's  $W$  was calculated for each level of the hierarchy and for each MCDM method. This non-parametric statistic measures the agreement among  $k$  stakeholders (the three groups) on the ranking of  $N$  criteria [7].

The mean ranks  $R_i$  for each criterion, derived from the priority weights, were used to compute  $W$  as follows:

$$R_i = \sum_{j=1}^m r_{ij} \quad (17)$$

In this context,  $i$  denotes each individual criterion being evaluated, while  $m$  represents the number of rater groups assessing the  $n$  factors (see Equations (18) and (19)):

$$R = m(n + 1)/2 \tag{18}$$

$$S = \sum_{i=1}^n (R_i - R)^2 \tag{19}$$

Here,  $S$  is the sum-of-squares statistic that captures the deviation of the cumulative ranks  $R_i$  from their mean, and  $R$  denotes the average of the  $R_i$  values.

Kendall’s coefficient of concordance  $W$ , which varies between 0 and 1, is then calculated using:

$$W = \frac{12S}{m^2(n^3 - n)} \tag{20}$$

Applying this formula yields a value that indicates the overall level of agreement among the different rater groups, as summarised in Table 3.

**Table 3**

Interpretive scale for Kendall’s coefficient of concordance ( $W$ ).

<b>W</b>	<b>Interpretation</b>
0	No agreement
0.10	Weak agreement
0.30	Moderate agreement
0.60	Strong agreement
1	Perfect agreement

## 4. Results

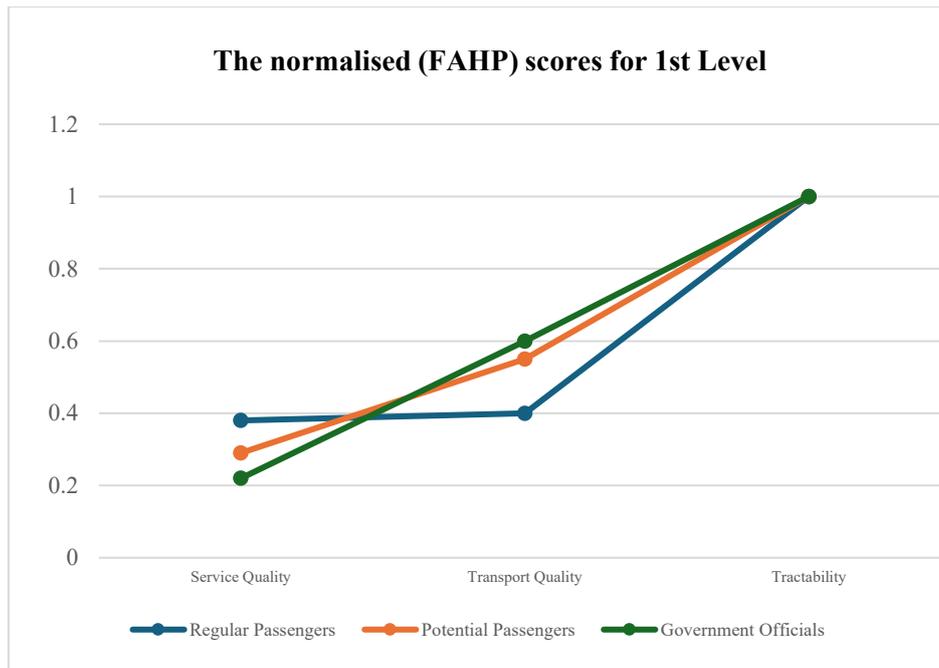
The study’s results are presented in this section, with an emphasis on the priority weights for the three stakeholder groups across a three-level criteria hierarchy that were obtained using the FAHP and IAHP methodologies. Kendall’s Coefficient of Concordance ( $W$ ) is used to objectively assess the level of agreement between these groupings.

### 4.1 Consensus on Strategic Goals (Level 1)

The top-level criteria encompass the broad strategic goals of the public transport system: Service Quality, Transport Quality, and Tractability.

#### 4.1.1 FAHP Results:

As shown in Fig. 4, the FAHP analysis revealed that Tractability was the highest priority, followed by Transport Quality and Service Quality. This ranking was consistent across all three stakeholder groups.



**Fig. 4.** Comparison of normalised FAHP scores among evaluator groups at Level 1.

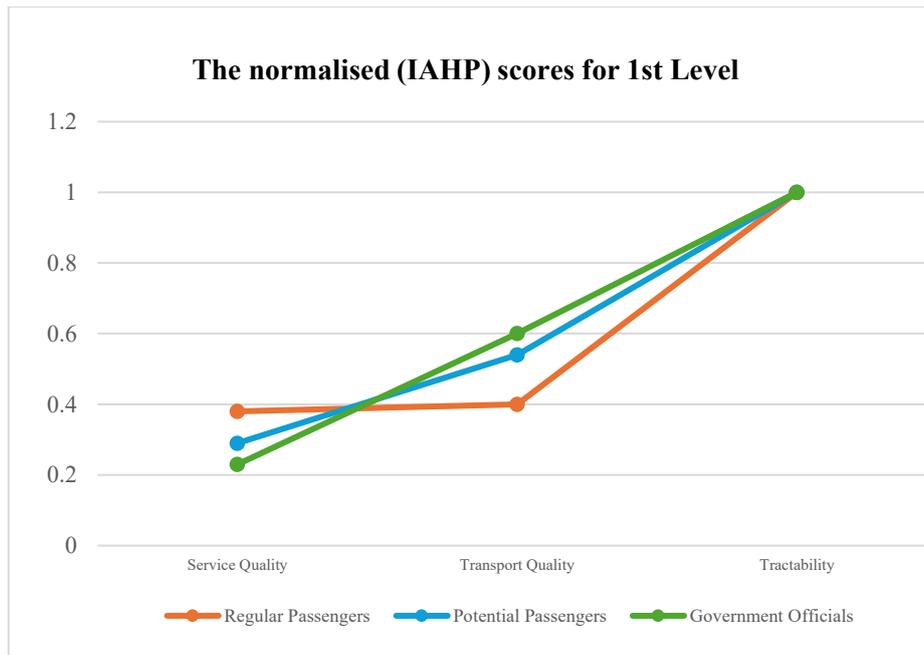
The agreement on this ranking, measured by Kendall’s W, was 0.778 (Table 4), indicating a very strong degree of consensus among regular passengers, potential passengers, and government officials.

**Table 4**  
 Agreement assessment using Kendall’s W across evaluator groups for Level 1 with the FAHP method.

	Regular users	Potential users	Government	Ri	(Ri-R)^2
Service Quality	2	3	3	8	4
Transport Quality	3	2	2	7	1
Tractability	1	1	1	3	9
<b>N = 3</b>	<b>M = 3</b>	<b>S = 14</b>	<b>R = 6</b>	<b>W = 0.778</b>	

**4.1.2 IAHP Results:**

The IAHP method, shown in Fig. 5, produced an even more unified result. All three stakeholder groups assigned identical ranks: 1st to Tractability, 2nd to Transport Quality, and 3rd to Service Quality.



**Fig. 5.** Comparison of normalised IAHP scores among evaluator groups at Level 1.

Because of this perfect alignment, the evaluator groups' Kendall's W value was 1.000 (Table 5), indicating perfect agreement on the strategic priorities.

**Table 5**  
 Agreement assessment using Kendall's W across evaluator groups for Level 1 with the IAHP method.

	Regular users	Potential users	Government	Ri	(Ri-R)^2
Service Quality	3	3	3	9	9
Transport Quality	2	2	2	6	0
Tractability	1	1	1	3	9
<b>n = 3</b>	<b>m = 3</b>	<b>S = 18</b>	<b>R = 6</b>	<b>W = 1</b>	

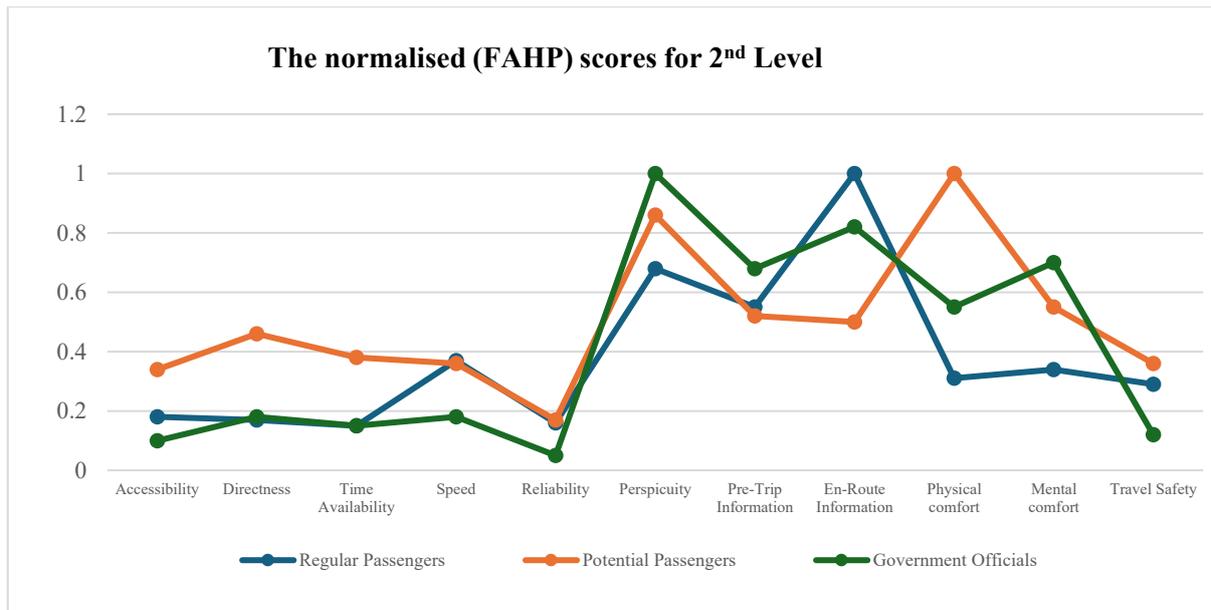
In both methods, Tractability consistently emerged as the highest priority dimension, followed by Transport Quality and Service Quality. These findings suggest strong cross-group alignment on the broad categories that define public transport supply quality.

#### 4.2. Level 2 Results: Performance Dimensions

This level consisted of 11 sub-criteria defining the key dimensions of service performance.

##### 4.2.1 FAHP Results:

The final FAHP scores are shown in Fig. 6. With the highest priority weight, the most important dimension during travel was found to be perspicuity and information. Information before to travel and physical comfort came next.



**Fig. 6.** Final normalised FAHP scores among evaluator groups at Level 2.

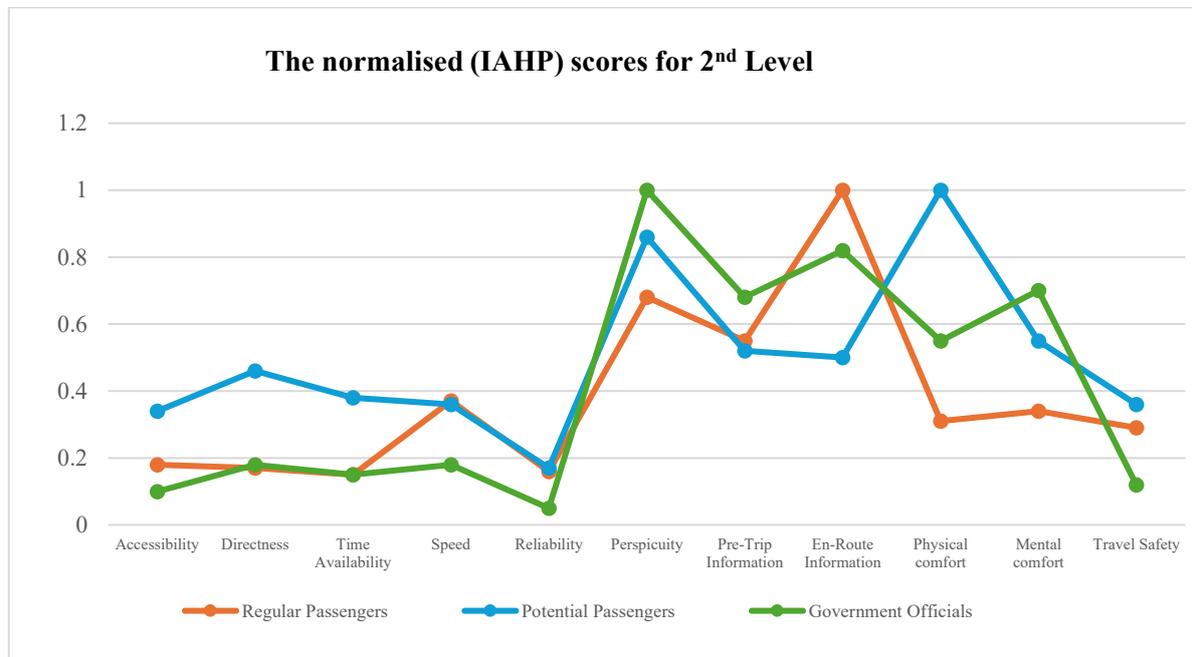
Despite the larger number of criteria, a strong consensus was maintained. Kendall’s W was calculated at 0.838 (Table 6), confirming that the groups strongly agreed on the relative importance of these 11 performance dimensions.

**Table 6**  
 Agreement assessment using Kendall’s W across evaluator groups for Level 2 with the FAHP method.

	Regular users	Potential users	Government	Ri	(Ri-R)^2
Accessibility	9	9	10	28	100
Directness	8	5	7	20	4
Time Availability	10	7	8	25	49
Speed	4	8	6	18	0
Reliability	11	11	11	33	225
Perspicuity	2	2	1	5	169
Pre-Trip Information	3	4	3	10	64
En-Route Information	1	6	2	9	81
Physical comfort	5	1	5	11	49
Mental comfort	6	3	4	13	25
Travel Safety	7	10	9	26	64
<b>N = 11</b>	<b>M = 3</b>	<b>S = 830</b>	<b>R = 18</b>	<b>W = 0.838</b>	

#### 4.2.2 IAHP Results:

The IAHP results, illustrated in Fig. 7, reinforced the findings from the FAHP analysis. Information during travel was again ranked as the most important dimension by all groups. The order of the top five criteria (Information during travel, Perspicuity, Information before travel, Physical comfort, Mental Comfort) showed strong alignment with the FAHP results.



**Fig. 7.** Final normalised IAHP scores among evaluator groups at Level 2.

The IAHP method yielded a marginally higher level of agreement, with a Kendall’s W of 0.861 (Table 7). This suggests that the interval-based approach slightly improved the harmonisation of stakeholder preferences at this level.

**Table 7**  
 Agreement assessment using Kendall’s W across evaluator groups for Level 2 with the IAHP method.

	Regular users	Potential users	Government	Ri	(Ri-R)^2
Accessibility	10	9	10	29	121
Directness	8	5	7	20	4
Time Availability	9	8	8	25	49
Speed	4	7	6	17	1
Reliability	11	11	11	33	225
Perspicuity	2	2	1	5	169
Pre-Trip Information	3	4	3	10	64
En-Route Information	1	6	2	9	81
Physical comfort	6	1	4	11	49
Mental comfort	5	3	5	13	25
Travel Safety	7	10	9	26	64
<b>n = 11</b>	<b>m = 3</b>	<b>S = 852</b>	<b>R = 18</b>	<b>W = 0.861</b>	

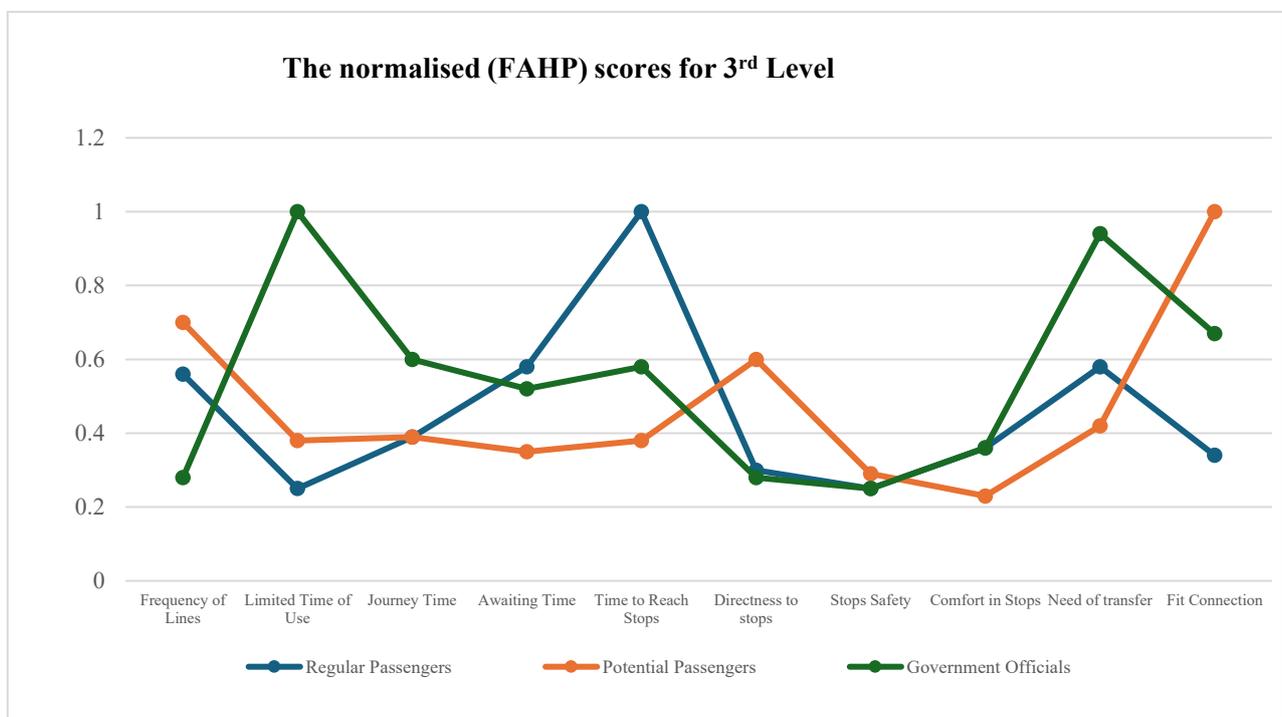
Across both methods, Information during travel was unanimously ranked as the most important sub-criterion. Perspicuity, Information before travel followed closely. Criteria such as Reliability and Accessibility consistently received the lowest weights across all evaluator groups. Thus, stakeholder consensus was not only high at this level but also reflected shared prioritisation of operational qualities most directly linked to everyday travel experience.

### 4.3. Level 3 Results: Operational Indicators

The most granular level included 10 specific, measurable indicators of service quality.

#### 4.3.1 FAHP Results:

As shown in Fig. 8, Stops Safety and Awaiting Time were identified as the two least important operational indicators. In contrast, Need of transfer and Time to Reach Stops were among the highest priorities.



**Fig. 8.** Final normalised FAHP scores among evaluator groups at Level 3.

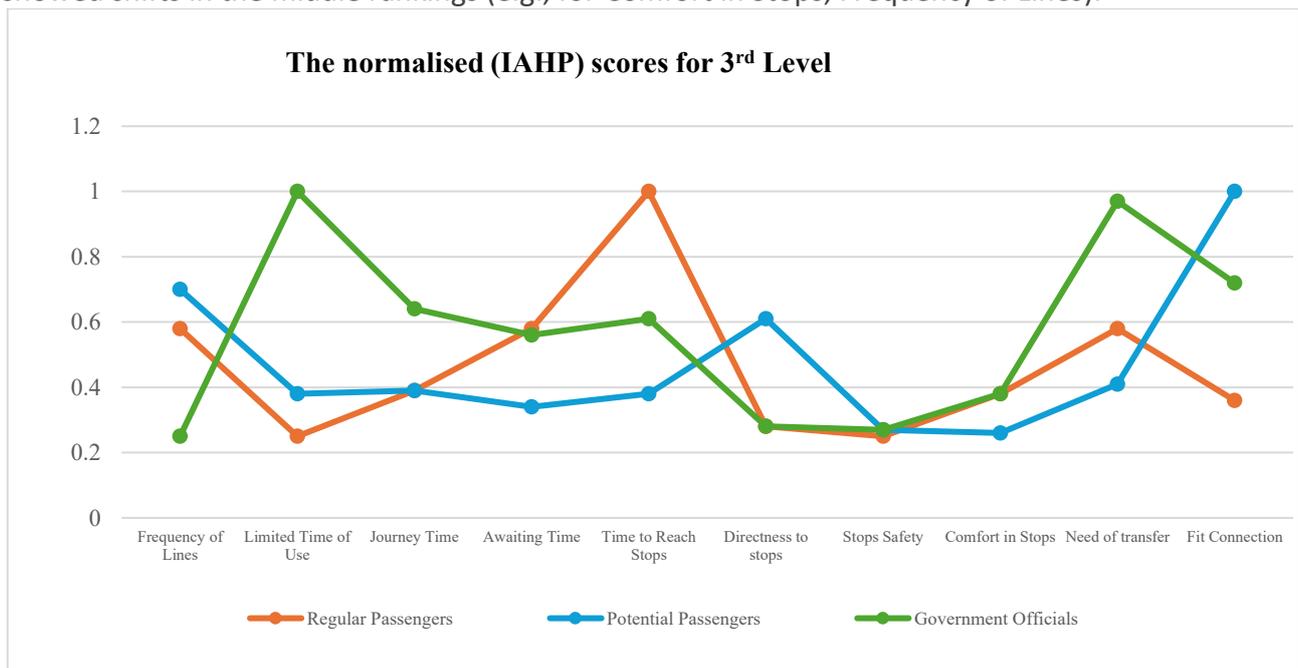
A significant drop in consensus was observed at this operational level. Kendall’s W decreased to 0.437 (Table 8), indicating only a moderate level of agreement. This reveals that while stakeholders agree on high-level goals, their priorities diverge significantly on specific, practical details.

**Table 8**  
 Agreement assessment using Kendall’s W across evaluator groups for Level 3 with the FAHP method.

	Regular users	Potential users	Government	Ri	(Ri-R)^2
Frequency of Lines	4	2	9	15	2.25
Limited Time of Use	10	8	1	19	6.25
Journey Time	5	5	5	15	2.25
Awaiting Time	3	7	6	16	0.25
Time to Reach Stops	1	6	4	11	30.25
Directness to stops	8	3	8	19	6.25
Stops Safety	9	9	10	28	132.25
Comfort in Stops	6	10	7	23	42.25
Need of transfer	2	4	2	8	72.25
Fit Connection	7	1	3	11	30.25
<b>n = 10</b>	<b>M = 3</b>	<b>S = 324.5</b>	<b>R = 16.5</b>	<b>W = 0.437</b>	

**4.3.2 IAHP Results:**

The IAHP results for Level 3 as shown in Fig. 9, confirmed the same top and bottom priorities: Time to Reach Stops (highest) and Stops Safety (lowest). However, the interval method showed shifts in the middle rankings (e.g., for Comfort in Stops, Frequency of Lines).



**Fig. 9.** Final normalised IAHP scores among evaluator groups at Level 3.

Crucially, the IAHP method achieved a higher consensus measure than FAHP at this level. The Kendall’s W value was 0.504 (Table 9), which falls into a moderate-to-strong agreement range. This demonstrates IAHP's superior capability to find common ground among stakeholders even where opinions are most divided.

**Table 9**  
 Agreement assessment using Kendall's W across evaluator groups for Level 3 with the IAHP method.

	Regular users	Potential users	Government	Ri	(Ri-R)^2
Frequency of Lines	4	2	9	15	2.25
Limited Time of Use	10	8	1	19	6.25
Journey Time	5	5	5	15	2.25
Awaiting Time	3	7	6	16	0.25
Time to Reach Stops	1	6	4	11	30.25
Directness to stops	8	4	8	20	12.25
Stops Safety	9	9	10	28	132.25
Comfort in Stops	7	10	7	24	56.25
Need for Transfer	2	3	2	7	90.25
Fit Connection	6	1	3	10	42.25
<b>n = 10</b>	<b>m = 3</b>	<b>S = 374.5</b>	<b>R = 16.5</b>	<b>W = 0.504</b>	

The level of stakeholder agreement is inversely related to the specificity of the criteria. Consensus was strongest on broad Strategic Goals (Level 1, W: 0.778-1.000), strong on Performance Dimensions (Level 2, W: 0.838-0.861), and weakest on detailed Operational Indicators (Level 3, W: 0.437-0.504).

The Interval AHP (IAHP) method consistently outperformed the Fuzzy AHP (FAHP) method across all three hierarchical levels, yielding higher Kendall's W values (Level 1: 1.000 > 0.778; Level 2: 0.861 > 0.838; Level 3: 0.504 > 0.437). This indicates that IAHP is a more effective tool for achieving and measuring consensus in a multi-stakeholder environment.

## 5. Discussion

This study set out to achieve two primary objectives: to evaluate public transport supply quality from a multi-stakeholder perspective and to determine the relative efficacy of FAHP versus IAHP in achieving consensus. The results present a clear narrative: while stakeholders share a strong consensus on high-level strategic goals, significant divergence emerges at the operational level, and the IAHP method consistently proves superior in harmonising these diverse perspectives. These results are interpreted, their theoretical and practical consequences are examined, and the limits of the study are acknowledged.

The inverse relationship between the level of criteria detail and the degree of consensus (Kendall's W) is a fundamental and expected finding, yet its magnitude is highly instructive. The perfect consensus (W = 1.000) on Level 1 strategic goals indicates that all stakeholders Regular users, Potential users and Government side share a unified vision for the public transport system in Mersin. This aligns with broader sustainable mobility paradigms that emphasise accessibility, reliability, and user satisfaction as universal goals [1]. It confirms that disagreement is not about the ultimate objectives but rather about how to achieve them.

The high consensus on Level 2 performance dimensions (W ≈ 0.85) suggests that stakeholders also largely agree on what constitutes good service. Attributes like Reliability, Accessibility, and Travel Safety are recognised as critical pillars by all parties, reflecting findings from previous studies on service quality determinants [49] & [50].

The significant drop in consensus at Level 3 ( $W \approx 0.47$ ) is where the most valuable insights for planners emerge. This divergence reveals the "devil in the details." While all groups value "Safety," their operational focus differs: passengers may prioritise "Stops Safety" (lighting, security), while officials might focus on "Travel Safety" (accident rates). Similarly, the low priority assigned to "Need of transfer" by passengers contrasts with its potential value to planners for creating efficient network structures. This fragmentation directly reflects the groups' divergent lived experiences and institutional roles, a classic challenge in participatory planning [51]. It underscores that the most intense negotiations and trade-offs in policy-making will occur at this operational level.

This study represents the outperformance of IAHP over FAHP across all hierarchical levels is the core methodological contribution of this study. We posit that this is not a random outcome but a direct consequence of the fundamental differences in how each method handles uncertainty and aggregates group preferences.

FAHP, through its defuzzification process, effectively compresses a range of opinions into a single, compromised value. While efficient, this process can act as a "black box," smoothing over nuanced disagreements and potentially creating an artificial veneer of agreement that masks the true extent of underlying divergence. This is likely why FAHP reported lower concordance.

In contrast, IAHP preserves the diversity of input judgments through its interval outputs. It does not force an artificial early compromise but instead identifies the bounds of agreement and disagreement transparently. This allows the aggregation process to more authentically find the "common ground" between groups. The interval for a criterion like "Reliability" would be wide, accurately reflecting conflict, while for "Need of transfer" it would be narrow, reflecting true consensus. This inherent transparency makes IAHP a more honest broker in group decision-making, leading to a higher authentic consensus as measured by Kendall's  $W$ . This finding resonates with arguments that methods preserving uncertainty lead to more robust and legitimate decisions in complex environments [10].

Therefore, IAHP is not merely an alternative but a more appropriate tool for contexts where understanding the spectrum of stakeholder opinion is paramount. It provides planners with a superior diagnostic tool, highlighting both zones of easy agreement and areas requiring targeted public engagement or negotiated trade-offs.

Regarding the above, the results provide a clear, evidence-based action plan for Mersin Municipality. The strong agreement on Strategic Goals (Level 1) and Performance Dimensions (Level 2) provides a solid mandate for investment. Resources should be prioritised immediately towards enhancing Reliability (e.g., via real-time tracking and priority lanes), Accessibility (e.g., improving pedestrian access to stops), and Time Availability (e.g., extending service hours). These investments have the highest probability of universal support and perceived success.

Also, we can assure that the disagreements at Level 3 are not a failure but a map for necessary dialogue. The notably low priority given to "Need of transfer" by passengers signals strong resistance to transfer-based solutions. Therefore, any policy aimed at network optimisation through hubs must be accompanied by robust citizen co-design workshops and compelling communication of the benefits (e.g., higher frequency on trunk lines). Similarly, improving "Stops Safety" (a high-priority indicator) would be a universally supported step that could build trust for more complex changes.

The Municipality should institutionalise the IAHP method for future planning consultations. Its ability to transparently document the degree of agreement provides a more legitimate foundation for tough decisions and helps avoid policies that are technically sound but lack social license.

This study has limitations that also present opportunities for future research. First, the focus on a single city limits the generalizability of the specific priority weights. Replicating this framework in

cities of different sizes and cultures would help distinguish universal preferences from context-specific ones. Second, the stakeholder sample, though representative, had asymmetric group sizes. Future studies could aim for greater balance or explore within-group heterogeneity. Finally, this study focused on perceived quality.

This research demonstrates that the choice of MCDM method has a real and measurable impact on participatory outcomes. By moving from a fuzzy to an interval approach, planners can achieve a more genuine consensus, leading to public transport policies that are not only smarter but also more broadly supported.

## **6. Conclusion**

This study undertook a rigorous comparative analysis of two Multi-Criteria Decision-Making (MCDM) methods, Fuzzy AHP (FAHP) and Interval AHP (IAHP), to evaluate public transport supply quality in Mersin, Turkey, through a multi-stakeholder lens. The integration of Kendall's Coefficient of Concordance ( $W$ ) as a primary performance metric provided a novel means to quantify the consensus-building efficacy of each method.

The findings lead to two overarching conclusions. First, the research confirms that stakeholder consensus is inherently hierarchical. While a strong, unified agreement exists on strategic goals and performance dimensions, significant divergence emerges at the level of specific, operational indicators. This illustrates that while all parties agree on the ends of a high-quality public transport system, they often disagree on the precise means to achieve them. For policymakers, this underscores the necessity of moving beyond high-level consultations to engage in detailed, participatory dialogues around operational trade-offs.

Second, and most significantly, the study demonstrates the clear methodological superiority of IAHP over FAHP in multi-stakeholder settings. The consistent achievement of higher Kendall's  $W$  values across all hierarchical levels indicates that IAHP is more effective at capturing and preserving the diversity of stakeholder preferences. By utilising interval judgments instead of forcing a compromise through defuzzification, IAHP provides a more transparent, honest, and ultimately robust aggregation of group inputs. This makes it a more reliable and legitimate tool for participatory transport planning.

The practical implications of this work are immediate. Planners in Mersin and similar cities are provided with a validated framework for action: they can confidently invest in high-consensus areas like reliability and accessibility while using IAHP to identify and strategically engage with stakeholders on more contentious operational issues like network design and transfer policies.

In closing, this research contributes to the fields of transport planning and MCDM by providing empirical evidence that the choice of methodological approach directly influences participatory outcomes. The IAHP method, paired with Kendall's  $W$ , offers a superior pathway to developing public transport policies that are not only technically sound but also socially legitimate and sustainable. Future work should focus on validating this framework in diverse urban contexts and integrating these subjective stakeholder priorities with objective performance data to create a comprehensive decision-support system for the cities of tomorrow.

## **Conflicts of Interest**

Declare conflicts of interest or state "The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper." Authors must identify and declare any personal circumstances or interests that may

be perceived as inappropriately influencing the representation or interpretation of reported research results. Any role of the funders in the design of the study; in the collection, analysis or interpretation of data; in the writing of the manuscript, or in the decision to publish the results must be declared in this section. If there is no role, please state "The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results".

### Author Contributions

Conceptualization, Hiba Solieman and Szabolcs Duleba; methodology, Hiba Solieman and Szabolcs Duleba; software, Hiba Solieman; validation, Hiba Solieman and Szabolcs Duleba; formal analysis, Hiba Solieman and Szabolcs Duleba; investigation, Hiba Solieman; resources, Hiba Solieman and Szabolcs Duleba; data curation, Hiba Solieman and Szabolcs Duleba; writing original draft preparation, Hiba Solieman; writing review and editing, Szabolcs Duleba; visualization, Hiba Solieman; supervision, Szabolcs Duleba; project administration, Hiba Solieman and Szabolcs Duleba; funding acquisition, Hiba Solieman and Szabolcs Duleba. All authors have read and agreed to the published version of the manuscript."

### Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request. Due to privacy and expert confidentiality agreements, the dataset is not publicly available.

### Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. No funders had a role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

### Acknowledgement

This research was not funded by any grant.

### References

- [1] T. Litman, "Evaluating Transportation Equity: Guidance for Incorporating Distributional Impacts in Transport Planning," *ITE Journal*, vol. 92, no. 4, Apr. 2022, Accessed: Aug. 22, 2025. [Online]. Available: <https://trid.trb.org/View/2006247>
- [2] E. Nathanail, "Measuring the quality of service for passengers on the hellenic railways," *Transportation Research Part A: Policy and Practice*, vol. 42, no. 1, pp. 48-66, Jan. 2008. <https://doi.org/10.1016/j.tra.2007.06.006>
- [3] E. K. Zavadskas, Z. Turskis, and S. Kildienė, "State of art surveys of overviews on MCDM/MADM methods," *Technological and Economic Development of Economy*, vol. 20, no. 1, pp. 165-179, Jan. 2014. <https://doi.org/10.3846/20294913.2014.892037>
- [4] T. L. Saaty, "Analytic Hierarchy Process," in *Encyclopedia of Operations Research and Management Science*, Springer, Boston, MA, 2013, pp. 52-64. [https://doi.org/10.1007/978-1-4419-1153-7\\_31](https://doi.org/10.1007/978-1-4419-1153-7_31)
- [5] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338-353, Jun. 1965. [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X)
- [6] M. Sarbast and D. Szabolcs, "Sustainable Urban Transport Development by Applying a Fuzzy-AHP Model: A Case Study from Mersin, Turkey," *Urban Science (Web)*, vol. 3, no. 2, p. 55, 2019, Accessed: Aug. 22, 2025. [Online]. Available: [https://jglobal.jst.go.jp/en/detail?JGLOBAL\\_ID=202102261091072115](https://jglobal.jst.go.jp/en/detail?JGLOBAL_ID=202102261091072115)  
<https://doi.org/10.3390/urbansci3020055>

- [7] O. Ghorbanzadeh, S. Moslem, T. Blaschke, and S. Duleba, "Sustainable Urban Transport Planning Considering Different Stakeholder Groups by an Interval-AHP Decision Support Model," *Sustainability*, vol. 11, no. 1, p. 9, Jan. 2019. <https://doi.org/10.3390/su11010009>
- [8] Y. Shiftan, S. Kaplan, and S. Hakkert, "Scenario building as a tool for planning a sustainable transportation system," *Transportation Research Part D: Transport and Environment*, vol. 8, no. 5, pp. 323-342, Sep. 2003. [https://doi.org/10.1016/S1361-9209\(03\)00020-8](https://doi.org/10.1016/S1361-9209(03)00020-8)
- [9] S. Duleba, F. Kutlu Gündoğdu, and S. Moslem, "Interval-Valued Spherical Fuzzy Analytic Hierarchy Process Method to Evaluate Public Transportation Development," *Informatica*, vol. 32, no. 4, pp. 661-686, Aug. 2021. <https://doi.org/10.15388/21-INFOR451>
- [10] S. Moslem, O. Ghorbanzadeh, T. Blaschke, and S. Duleba, "Analysing stakeholder consensus for a sustainable transport development decision by the fuzzy AHP and interval AHP," *Sustainability (Switzerland)*, vol. 11, no. 12, p. 3271, Jan. 2019. <https://doi.org/10.3390/su11123271>
- [11] D. A. Hensher, P. Stopher, and P. Bullock, "Service quality--developing a service quality index in the provision of commercial bus contracts," *Transportation Research Part A: Policy and Practice*, vol. 37, no. 6, pp. 499-517, Jul. 2003. [https://doi.org/10.1016/S0965-8564\(02\)00075-7](https://doi.org/10.1016/S0965-8564(02)00075-7)
- [12] D. van Lierop and A. El-Geneidy, "Enjoying loyalty: The relationship between service quality, customer satisfaction, and behavioral intentions in public transit," *Research in Transportation Economics*, vol. 59, pp. 50-59, Nov. 2016. <https://doi.org/10.1016/j.retrec.2016.04.001>
- [13] L. Eboli and G. Mazzulla, "Customer Satisfaction as a Measure of Service Quality in Public Transport Planning," *International Encyclopedia of Transportation*, pp. 220-224, Jan. 2021. <https://doi.org/10.1016/B978-0-08-102671-7.10643-8>
- [14] L. Redman, M. Friman, T. Gärling, and T. Hartig, "Quality attributes of public transport that attract car users: A research review," *Transport Policy*, vol. 25, pp. 119-127, Jan. 2013. <https://doi.org/10.1016/j.tranpol.2012.11.005>
- [15] G. Yannis, A. Kopsacheili, A. Dragomanovits, and V. Petraki, "State-of-the-art review on multi-criteria decision-making in the transport sector," *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 7, no. 4, pp. 413-431, Aug. 2020. <https://doi.org/10.1016/j.jtte.2020.05.005>
- [16] E. Broniewicz and K. Ogródnik, "A Comparative Evaluation of Multi-Criteria Analysis Methods for Sustainable Transport," *Energies*, vol. 14, no. 16, p. 5100, Jan. 2021. <https://doi.org/10.3390/en14165100>
- [17] K. De Brucker, C. Macharis, and A. Verbeke, "Multi-criteria analysis and the resolution of sustainable development dilemmas: A stakeholder management approach," *European Journal of Operational Research*, vol. 224, no. 1, pp. 122-131, Jan. 2013. <https://doi.org/10.1016/j.ejor.2012.02.021>
- [18] L. Shen, L. Du, X. Yang, X. Du, J. Wang, and J. Hao, "Sustainable Strategies for Transportation Development in Emerging Cities in China: A Simulation Approach," *Sustainability*, vol. 10, no. 3, p. 844, Mar. 2018. <https://doi.org/10.3390/su10030844>
- [19] S. Duleba and S. Moslem, "Examining Pareto optimality in analytic hierarchy process on real Data: An application in public transport service development," *Expert Systems with Applications*, vol. 116, pp. 21-30, Feb. 2019. <https://doi.org/10.1016/j.eswa.2018.08.049>
- [20] R. A. Aliev, B. G. Guirimov, O. H. Huseynov, and R. R. Aliyev, "A consistency-driven approach to construction of Z-number-valued pairwise comparison matrices. | EBSCOhost." Accessed: Aug. 23, 2025. [Online]. Available: <https://openurl.ebsco.com/contentitem/gcd:150458045?sid=ebsco:plink:crawler&id=ebsco:gcd:150458045>
- [21] G. Demir and S. Moslem, "Chapter 11 - Fuzzy multicriteria decision making for assessing public transportation systems for sustainable transportation," in *Intelligent Urban Mobility*, M. Deveci, Ed., Academic Press, 2025, pp. 229-258. <https://doi.org/10.1016/B978-0-443-34160-1.00013-4>
- [22] H. Solieman and S. Duleba, "Enhancing urban transportation by shared mobility solutions surveyed with fuzzy ZE-based decision-making," in *Reliable Decision-Making for Sustainable Transportation*, Academic Press, 2026, pp. 169-201. <https://doi.org/10.1016/B978-0-443-33740-6.00007-4>
- [23] M. Keshavarz-Ghorabae, M. Amiri, E. K. Zavadskas, Z. Turskis, and J. Antuchevičienė, "MCDM approaches for evaluating urban and public transportation systems: a short review of recent studies," *Transport*, vol. 37, no. 6, pp. 411-425, Dec. 2022. <https://doi.org/10.3846/transport.2022.18376>
- [24] L. dell'Olio, A. Ibeas, and P. Cecin, "Modelling user perception of bus transit quality," *Transport Policy*, vol. 17, no. 6, pp. 388-397, Nov. 2010. <https://doi.org/10.1016/j.tranpol.2010.04.006>
- [25] G. Currie and I. Wallis, "Effective ways to grow urban bus markets - a synthesis of evidence," *Journal of Transport Geography*, vol. 16, no. 6, pp. 419-429, Nov. 2008. <https://doi.org/10.1016/j.jtrangeo.2008.04.007>

- [26] K. Gwilliam, "Bus transport: Is there a regulatory cycle?," *Transportation Research Part A: Policy and Practice*, vol. 42, no. 9, pp. 1183-1194, Nov. 2008. <https://doi.org/10.1016/j.tra.2008.05.001>
- [27] D. Banister, "The sustainable mobility paradigm," *Transport Policy*, vol. 15, no. 2, pp. 73-80, Mar. 2008. <https://doi.org/10.1016/j.tranpol.2007.10.005>
- [28] G. Baudry, C. Macharis, and T. Vallée, "Range-based Multi-Actor Multi-Criteria Analysis: A combined method of Multi-Actor Multi-Criteria Analysis and Monte Carlo simulation to support participatory decision making under uncertainty," *European Journal of Operational Research*, vol. 264, no. 1, pp. 257-269, Jan. 2018. <https://doi.org/10.1016/j.ejor.2017.06.036>
- [29] V. S. Krishnan, H. Solieman, S. Moslem, and F. Pilla, "Inclusive sustainable transportation planning: A qualitative analysis involving diverse stakeholders perspectives," *Case Studies on Transport Policy*, vol. 19, p. 101393, Mar. 2025. <https://doi.org/10.1016/j.cstp.2025.101393>
- [30] L. A. Zadeh, G. J. Klir, and B. Yuan, *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers*. World Scientific, 1996. <https://doi.org/10.1142/2895>
- [31] S. Moslem, H. Solieman, L. Oubahman, S. Duleba, T. Senapati, and F. Pilla, "Assessing Public Transport Supply Quality: A Comparative Analysis of Analytical Network Process and Analytical Hierarchy Process," *Journal of Soft Computing and Decision Analytics*, vol. 1, no. 1, Art. no. 1, Aug. 2023. <https://doi.org/10.31181/jscda11202311>
- [32] S. Moslem et al., "A Fuzzy Decision-Making Support Model for Traffic Safety Analysis," *Expert Systems*, vol. 42, no. 9, p. e70104, 2025. <https://doi.org/10.1111/exsy.70104>
- [33] A. Awasthi and S. S. Chauhan, "A hybrid approach integrating Affinity Diagram, AHP and fuzzy TOPSIS for sustainable city logistics planning," *Applied Mathematical Modelling*, vol. 36, no. 2, pp. 573-584, Feb. 2012. <https://doi.org/10.1016/j.apm.2011.07.033>
- [34] Y.-M. Wang, Y. Luo, and Z. Hua, "On the extent analysis method for fuzzy AHP and its applications," *European Journal of Operational Research*, vol. 186, no. 2, pp. 735-747, Apr. 2008. <https://doi.org/10.1016/j.ejor.2007.01.050>
- [35] S. Moslem, M. Mohammadi, K. Ismael, and D. Esztergár-Kiss, "Fostering sustainable urban mobility via stakeholder engagement: A novel analytic hierarchy process and half-quadratic programming," *Research in Transportation Business & Management*, vol. 59, p. 101291, Mar. 2025. <https://doi.org/10.1016/j.rtbm.2025.101291>
- [36] K. Sugihara and H. Tanaka, "Interval Evaluations in the Analytic Hierarchy Process By Possibility Analysis," *Computational Intelligence*, vol. 17, no. 3, pp. 567-579, 2001. <https://doi.org/10.1111/0824-7935.00163>
- [37] R. M. Darbra, E. Eljarrat, and D. Barceló, "How to measure uncertainties in environmental risk assessment," *TrAC Trends in Analytical Chemistry*, vol. 27, no. 4, pp. 377-385, Apr. 2008. <https://doi.org/10.1016/j.trac.2008.02.005>
- [38] T.-C. Wang and Y.-H. Chen, "Applying consistent fuzzy preference relations to partnership selection," *Omega*, vol. 35, no. 4, pp. 384-388, Aug. 2007. <https://doi.org/10.1016/j.omega.2005.07.007>
- [39] M. Le Pira, M. Ignaccolo, G. Inturri, A. Pluchino, and A. Rapisarda, "Modelling stakeholder participation in transport planning," *Case Studies on Transport Policy*, vol. 4, no. 3, pp. 230-238, Sep. 2016. <https://doi.org/10.1016/j.cstp.2016.06.002>
- [40] T. L. Saaty, "The Analytic Hierarchy and Analytic Network Processes for the Measurement of Intangible Criteria and for Decision-Making," in *Multiple Criteria Decision Analysis: State of the Art Surveys*, S. Greco, M. Ehrgott, and J. R. Figueira, Eds., New York, NY: Springer, 2016, pp. 363-419. [https://doi.org/10.1007/978-1-4939-3094-4\\_10](https://doi.org/10.1007/978-1-4939-3094-4_10)
- [41] M. J. Hill, R. Braaten, S. M. Veitch, B. G. Lees, and S. Sharma, "Multi-criteria decision analysis in spatial decision support: the ASSESS analytic hierarchy process and the role of quantitative methods and spatially explicit analysis," *Environmental Modelling & Software*, vol. 20, no. 7, pp. 955-976, Jul. 2005. <https://doi.org/10.1016/j.envsoft.2004.04.014>
- [42] B. Feizizadeh, M. Shadman Roodposhti, P. Jankowski, and T. Blaschke, "A GIS-based extended fuzzy multi-criteria evaluation for landslide susceptibility mapping," *Computers & Geosciences*, vol. 73, pp. 208-221, Dec. 2014. <https://doi.org/10.1016/j.cageo.2014.08.001>
- [43] Z. Xu, "Group Decision Making with Triangular Fuzzy Linguistic Variables," in *Intelligent Data Engineering and Automated Learning - IDEAL 2007*, H. Yin, P. Tino, E. Corchado, W. Byrne, and X. Yao, Eds., Berlin, Heidelberg: Springer, 2007, pp. 17-26. [https://doi.org/10.1007/978-3-540-77226-2\\_3](https://doi.org/10.1007/978-3-540-77226-2_3)
- [44] L. Abdullah and P. Goh, "Decision making method based on Pythagorean fuzzy sets and its application to solid waste management," *Complex Intell. Syst.*, vol. 5, no. 2, pp. 185-198, Jun. 2019. <https://doi.org/10.1007/s40747-019-0100-9>

- [45] J. Żak and M. Kruszyński, "Application of AHP and ELECTRE III/IV Methods to Multiple Level, Multiple Criteria Evaluation of Urban Transportation Projects," *Transportation Research Procedia*, vol. 10, pp. 820-830, Jan. 2015. <https://doi.org/10.1016/j.trpro.2015.09.035>
- [46] M. Brunelli, "A survey of inconsistency indices for pairwise comparisons," *International Journal of General Systems*, vol. 47, no. 8, pp. 751-771, Nov. 2018. <https://doi.org/10.1080/03081079.2018.1523156>
- [47] P. Cilliers, *Complexity and Postmodernism: Understanding Complex Systems*. London: Routledge, 2002 <https://doi.org/10.4324/9780203012253>
- [48] M. Brunelli, *Introduction to the Analytic Hierarchy Process*. Springer, 2014. <https://doi.org/10.1007/978-3-319-12502-2>
- [49] L. Eboli and G. Mazzulla, "Performance indicators for an objective measure of public transport service quality," Mar. 2012, Accessed: Nov. 03, 2025. [Online]. Available: <http://hdl.handle.net/10077/6119>
- [50] L. dell'Olio, A. Ibeas, and P. Cecin, "The quality of service desired by public transport users," *Transport Policy*, vol. 18, no. 1, pp. 217-227, Jan. 2011, <https://doi.org/10.1016/j.tranpol.2010.08.005>
- [51] C. Macharis and A. Bernardini, "Reviewing the use of Multi-Criteria Decision Analysis for the evaluation of transport projects: Time for a multi-actor approach," *Transport Policy*, vol. 37, pp. 177-186, Jan. 2015, <https://doi.org/10.1016/j.tranpol.2014.11.002>