

IMPROVING CONSISTENCY EVALUATION IN FUZZY MULTI-ATTRIBUTE PAIRWISE COMPARISON-BASED DECISION-MAKING METHODS

FARHAD SHAMS^{*‡}, SHERIF MOHAMED^{†§}, and AMINAH ROBINSON FAYEK^{*¶}

**Hole School of Construction Engineering
Department of Civil and Environmental Engineering
University of Alberta, Edmonton, Alberta
T6G 2W2 Canada*

*†Gold Coast Campus
Griffith University
QLD 4222, Australia*

*‡fshams@ualberta.ca
§s.mohamed@griffith.edu.au
¶ aminah.robinson@ualberta.ca*

Abstract

A typical approach to handle the complexity of multi-faceted decision-making problems is to use multi-attribute decision-making (MADM) methods based on pairwise comparisons. Fuzzy set theory has also been employed to cope with the uncertainty and vagueness involved in conducting the comparisons between components of a decision model. An important issue regarding the reliability of the output is the consistency of pairwise comparisons provided by the decision maker(s). Using the MADM method developed by Lu, Zhang, Ruan, and Wu (2007) as a foundation, this paper proposes an algorithm for evaluating the consistency level of pairwise comparison matrices where linguistic data are used. A crisp numeric scale has been introduced to consider the priority of linguistic data and to avoid the complexity of handling fuzzy calculations in consistency evaluation of pairwise comparison matrices. As an advantage, the proposed method of consistency evaluation is capable of assessing the degree of

[¶] Corresponding author.

inconsistency among the pairwise comparisons. Therefore, the acceptance or rejection of the pairwise comparisons can be determined based on the desired degree of tolerance in accepting inconsistent judgments. An application of a revised MADM method is then demonstrated in a case study involving flood mitigation project selection in Australia.

1. Introduction

Decision-making problems often involve different stakeholders whose opinions could conflict. Accordingly, quantifying the decision makers' preference level for one attribute over another in MADM methods has resulted in the development of a number of scales for this purpose (Triantaphyllou, Loostma, Pardalos, & Mann, 1994). As degree of preference is inherently vague and subjective, an interval or a linguistic term can be a better descriptor for such assessment, rather than a single point on a numeric scale. Therefore, using intervals or fuzzy numbers instead of scales with single numeric points has been suggested (Deng, 1999; Kahraman, 2008).

The Analytic Hierarchy Process (AHP), developed by Saaty in the 1980s, is one of the most applied MADM methods. AHP provides a framework in which pairwise comparisons between the attributes at one level and pairwise comparisons between alternatives with respect to each attribute at the second level effectively capture the preferences of participants by limiting the evaluations to two attributes (or alternatives) at a time. Since Saaty introduced AHP, a variety of methods to improve properties of AHP have been developed, as well as a number of hybrid AHP methods. Fuzzy AHP was initially suggested by Van Laarhoven and Pedrycz (1983) using triangular fuzzy numbers to conduct the pairwise comparisons and using the logarithmic least square method to obtain fuzzy weights. Many other fuzzy AHP methods developed afterwards are systematic decision-making methods using fuzzy sets and a hierarchical structure to model the decision problem.

Boender, Graan, and Lootsma (1989) used a logarithmic regression function to define the criteria weights and modified the Van Laarhoven and Pedrycz (1983) method by producing individual weights

when comparing alternatives and in the aggregation method. Chang (1996) used the extent analysis method for the synthetic extent of the pairwise comparisons. Kahraman, Ulukan, and Tolga (1998) used objective and subjective measures within a weighted method of AHP. Deng (1999) presented a fuzzy AHP that avoids ranking methods based on fuzzy utilities and used the fuzzy extent method and alpha-cuts to solve the fuzzy pairwise comparison matrices and to establish the interval matrix in accordance with the fuzzy comparison matrix.

One of the more recent MADM methods using the principle of pairwise comparisons and employing fuzzy sets theory and a linguistic basis for comparison was introduced by Lu, Zhang, Ruan, and Wu (2007); this method is a hybrid of AHP and the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) that leverages the advantages of both techniques. AHP provides pairwise comparison, limiting the comparison process to two attributes (or alternatives) at a time, while TOPSIS provides an inconsistency-free matrix at the alternative comparison level. TOPSIS uses a reliable algorithm in synthesising pairwise comparison matrices by modelling the problem as a geometric space. According to Zanakis, Solomon, Wishart, and Dublisch (1998), TOPSIS performs better than AHP in terms of coping with the rank reversal problem, so it can be considered as a more reliable method to provide the final ranking result. While fuzzy MADM methods such as Lu, Zhang, Ruan, and Wu (2007) represent a step forward in the field, there is still room for improvement in handling some aspects of MADM.

MADM methods in which data from decision makers are collected as pairwise comparisons between the attributes (or alternatives) of the model can be subject to inconsistency in the pairwise comparison matrices. Therefore, evaluating the consistency level is important regardless of whether the method uses linguistic terms or numeric scales in its pairwise comparisons.

This paper addresses the problem of consistency evaluation of linguistic pairwise comparisons that are typical inputs to fuzzy MADM methods. Given its advantages, the Lu, Zhang, Ruan, and Wu

(2007) fuzzy MADM method has been used to illustrate the applicability of the proposed method. However, the proposed method is not restricted to this specific MADM method and can be generalised to improve other methods that are based on linguistic data and pairwise comparisons. Finally, the proposed methodology is applied to the selection of a flood mitigation project in Australia to demonstrate its efficacy in a real-world decision-making scenario.

2. The Basic Components of Fuzzy MADM Methods Based on Pairwise Comparisons

In MADM methods based on pairwise comparisons (AHP and AHP-like methods) such as Saaty (1980), Van Laarhoven and Pedrycz (1983), and Lu et al. (2007), a problem is usually modelled in a hierarchy consisting of the main goal, a collection of attributes of the problem, and solution alternatives to be compared with respect to each attribute and ranked in the order of overall preference. Matrices are usually employed to structure the priority values of components of the model. In AHP, for instance, a matrix for attribute comparison and matrices for alternative comparisons regarding the attributes should be defined. Lu, Zhang, Ruan, and Wu (2007) introduced a relation matrix of aspects, attributes, alternatives, their weights, and their evaluation values. In this paper, pairwise comparisons between the attributes of the model are referred to as attribute level comparisons (level 1), and comparisons of alternative solutions are referred to as alternative level comparisons (level 2).

In order to establish the matrices, the preference levels based on the comparisons must be expressed by the use of a scale. The typical scales used in such methods are Likert-type scales with numeric points that vary in terms of the range and the structure of the scale. In fuzzy MADM methods, linguistic terms are typically used where each point on a linguistic preference scale is associated with a fuzzy number to handle the uncertainty and vagueness of human preferences. In this paper, both types of scales, with crisp and fuzzy numbers, are referred to as numeric scales. The linguistic terms can be ordered in terms of their significance, and each term can be viewed as a point of a scale that is referred to in this paper as a linguistic scale.

In AHP-like MADM methods, two levels of comparisons must be conducted. One type of matrix is typically used for capturing the pairwise preference of attributes (i.e., attribute comparison level), and one type of matrix is used for data on the pairwise degree of preference of alternatives with regard to each attribute (i.e., alternative comparison level). The concept of consistency evaluation applies to both types of matrices. For simplicity, the consistency of the attribute comparison matrix type (first level) is discussed in the following sections. However, the same concepts and computations apply to the alternative matrices in the second level, where alternatives are compared in a pairwise manner and in accordance to each of the attributes.

3. A Method for Evaluating the Consistency of Linguistic Pairwise Comparisons

In MADM methods, which are based on pairwise comparison, a relative logical priority among a few attributes is easy to retain. However, once the number of elements to be compared increases, the interrelationships between all the elements of a comparison matrix cannot be easily tracked by the human mind due to the complexity of the logic, and the matrix will be subject to inconsistency. Therefore, identifying the level of consistency of the pairwise comparison matrices is an important part of the method.

The output of an MADM method is typically a ranking of solution alternatives based on the attributes of the problem and the preferences of decision makers. This ranking has to be based on a single value that is a representative of all the involved factors. Different MADM methods have different principles for synthesising the preferences and providing a single basis for ranking the alternatives. For instance, AHP calculates a synthesised value for alternatives based on eigenvectors of the attribute-level comparison matrix and local priorities of the alternatives (Saaty & Kearns, 1985). In the Lu, Zhang, Ruan, and Wu (2007) method, the closeness coefficient borrowed from TOPSIS is used as the basis for ranking the alternatives.

3.1 An introduction to consistency evaluation in MADM methods

One of the most important tasks in MADM methods based on pairwise comparisons is evaluating the level of consistency in pairwise comparison matrices, as it influences the overall precision of the final result. Consistency evaluation is an unavoidable part of the process because a high level of inconsistency leads to an unreliable outcome.

In a consistent matrix of pairwise comparisons, the elements of the matrix must be reciprocal, and the preference of each attribute over itself must be equal to unity. These logical relationships are simple for decision makers to retain and can be easily tracked within the matrix. Another aspect of consistency is the principle of transitivity. This refers to the fact that when attribute a is preferred to b and b is preferred to c , then a needs to be expressed as preferred to c with a relative degree. The order of the preferences and the associated degree of preference must be retained in accordance with their logical relationships. For such a purpose, a consistency index (CI) was defined in the analytical hierarchy process (AHP) introduced by Saaty (1980), as shown in Equation 1:

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

In this definition, λ_{max} is the maximum eigenvalue of the pairwise comparison matrix, and n is the number of rows or columns of the square matrix.

Evaluating the level of transitivity of a pairwise comparison matrix is related to the type of comparison scale that is used in the MADM method. The types of scales vary in terms of the arithmetic that handles preference of one attribute over another. Some of the scales, such as Saaty's original AHP scale (Saaty, 1980), have a multiplicative basis (i.e., a ratio scale). In such a scale, the level of preference has to be expressed as a number that indicates how many times more or less important one attribute is relative to another. In other words, the preference of one attribute over another is a ratio; therefore, the following rules apply for each i and j and k as elements of the model when r_{ij} denotes the preference level of element i over element j :

- i. $r_{ij}=1/r_{ji}$
- ii. $r_{ii}=1$
- iii. $r_{ik} = r_{ij} \cdot r_{jk}$

Other scales have also been developed to improve the precision of quantifying the pairwise comparisons. For instance, Lootsma (1988) introduced a scale that replaces the multiplicative basis of preference with a logarithmic regression model. Another approach to quantification of priority between a pair of attributes is using an additive scale. In this kind of scale, the preference levels of attributes b over a is a value added to or subtracted from the value of attribute a on the scale. Ali, Cook, and Kress (1986), Harker and Vargas (1987), Lootsma (1993), Hochbaum, and Levin (2006), and Guh, Po, and Lou (2009) argued that an additive scale is a valid method for quantifying priorities in pairwise comparisons. If a symmetric additive scale with 0 as the neutral point of scale is used the following rules apply:

- i. $r_{ij} = -r_{ji}$
- ii. $r_{ii}=0$
- iii. $r_{ik} = r_{ij} + r_{jk}$

In a typical fuzzy AHP and many other fuzzy decision-making methods based on pairwise comparisons, the magnitude of preferences is usually expressed based on a linguistic scale. When using such scales, participants, in essence, think in an intuitive, additive manner. In other words, participants do not think of how many times more important one attribute is over another; rather, they think of how much more or less important (expressed as an interval) one attribute is as compared to another. According to Herrera-Viedma et al. (2004), consistency evaluation for fuzzy preference relations is associated with the

transitivity property, and additive transitivity is a relatively strong concept compared to other measures of transitivity.

Consistency and transitivity in pairwise comparison scales also apply to linguistic scales. The points of a linguistic scale (i.e. linguistic terms) can be ordered based on their level of significance, and a relative degree of significance is logically expected when comparing the attributes of a MADM model in a pairwise manner. In addition, points of a linguistic scale need to be quantified and mapped to a numeric scale, which can be also subject to inconsistency.

Arbel (1989) introduced the concept of defining a feasible region for local priorities based on the relation between interval judgment and linear constraints. The expressed pairwise comparison was stated to be inconsistent if the defined region was empty. Salo and Hämäläinen (1995) used interval judgements to deal with the uncertainty of priority judgments; they extended Arbel's (1989) method by the use of optimisation techniques to synthesise feasible regions that are not empty and provide dominance relations. The decision maker's involvement was required in describing the preferences to avoid inconsistency in the process.

Salo (1996) raised the issue of consistency evaluation in fuzzy AHP; in his approach, he defined a feasible region by the use of the extension principle and fuzzy ratios as constraints of the membership degrees of pairwise comparisons. Fuzzy weights, therefore, could be obtained as a result of a linear programming problem. Leung and Cao (2000) also used Salo's (1996) method but added a degree of tolerance for deviation from consistency, so that inconsistency could be tolerated to a determined degree based on the application and the decision maker's preference. In addition, Leung and Cao's (2000) method provides the capability of filtering inconsistent ratios of a pairwise comparison matrix.

All of the mentioned studies of consistency in fuzzy pairwise comparisons are restricted to certain conditions, such as having quasi-concave fuzzy comparison ratios, and work on the basis of multiplicative scales of assessment in which priorities are expressed as ratios. This paper introduces an algorithm,

presented in detail in Section 3.3, for measuring the consistency level of pairwise comparison matrices based on linguistic scales of priority through the use of an associated crisp additive scale. The proposed algorithm provides the advantage of indicating the set of triple elements that exhibits intransitivity so that only those pairwise comparisons and the affected pairwise comparisons need to be repeated or filtered—that is, repeating or filtering the set of elements of the pairwise comparison matrix that have common attributes with those of the intransitive set. Compared to other suggested methods of consistency evaluation, this advantage enables a large volume of information to be maintained, as an entire matrix is not necessarily subject to rejection in case of intransitivity. Moreover, the algorithm is flexible enough to be configured if a certain level of inconsistency is tolerable based on user preference and the criticality of accuracy of the decision problem.

Lu, Zhang, Ruan, and Wu (2007) also introduced a method for evaluating the consistency of linguistic pairwise comparisons. Table 1 provides the linguistic scale used in this method, as well as an associated scale that has been employed for consistency evaluation in this method. The proposed scale is a sample of a symmetric linguistic scale in which each point of scale is associated with a reverse point when the reverse priority is required to be considered in the pairwise comparison matrix. As an example, considering the linguistic scale presented in Table 1, if a is considered to be “Much less important” than b , b has to be considered as “much more important” than a .

Table 1: Linguistic terms and the suggested scale for consistency evaluation in Lu et al. (2007)

Linguistic Priority Scale	Absolutely less important	Much less important	Less important	Equally important	More important	Much more important	Absolutely more important
Lu et al. (2007) Associated Scale for Consistency Evaluation	a_1	a_2	a_3	a_4	a_5	a_6	a_7

Fuzzy Priority Scale (Triangular Fuzzy Numbers)	$(0,0,\frac{1}{6})$	$(0,\frac{1}{6},\frac{1}{3})$	$(\frac{1}{6},\frac{1}{3},\frac{1}{2})$	$(\frac{1}{3},\frac{1}{2},\frac{2}{3})$	$(\frac{1}{2},\frac{2}{3},\frac{5}{6})$	$(\frac{2}{3},\frac{5}{6},1)$	$(\frac{5}{6},1,1)$
--	---------------------	-------------------------------	---	---	---	-------------------------------	---------------------

Lu, Zhang, Ruan, and Wu (2007) suggested that using a set of linguistic inference rules corrects the inconsistency of each pairwise comparison matrix $[r_{ij}]_{m \times m}$. Considering Table 1, the rules are as follows (Lu et al., 2007):

- i. If $r_{ij} = a_s$, $s \in \{4,5,6,7\}$ and $r_{jp} = a_t$, $t \in \{4,5,6,7\}$; then $r_{ip} = a_{\max(s,t)}$.
- ii. If $r_{ij} = a_s$, $s \in \{1,2,3\}$ and $r_{jp} = a_t$, $t \in \{1,2,3\}$; then $r_{ip} = a_{\min(s,t)}$.
- iii. If $r_{ij} = a_s$, $s \in \{4,5,6,7\}$ and $r_{jp} = a_t$, $t \in \{1,2,3\}$; then $r_{ip} = a_i$ for any $t \leq i \leq s$.
- iv. If $r_{ij} = a_s$, $s \in \{1,2,3\}$ and $r_{jp} = a_t$, $t \in \{4,5,6,7\}$; then $r_{ip} = a_i$ for any $s \leq i \leq t$.

These rules check the consistency of pairwise comparison matrices to some extent but do not guarantee perfectly consistent pairwise comparison matrices. The following example presents one of the conditions that conform to these rules but deviate from perfect consistency:

- Consider attribute i to be “much less important” than attribute j , which is “absolutely more important” than attribute p . Based on these rules, $r_{ij} = a_2$ and $r_{jp} = a_7$; then, according to the Rule iv, r_{jp} can have any of the following values: $a_2, a_3, a_4, a_5, a_6, a_7$. This means that the degree of preference of attribute i as compared to attribute p can range from “less important” to “absolutely more important”; some of the points on the linguistic scale within this range have low degrees of accuracy in reflecting the degree of preference. In addition, it can be rationally deduced that attribute i is “less important” than attribute m based on the two related comparisons, and there is no need for an approximation.

3.2 A consistency evaluation method for fuzzy pairwise comparison-based MADM methods with linguistic scales

In this section, the development of an algorithm to check the level of consistency of pairwise comparison matrices that are based on linguistic scales of judgement is explained. An indicator for measuring the degree of consistency among sets of triple pairwise comparisons is introduced, and the use of this indicator to develop a consistency evaluation algorithm for linguistic pairwise comparison matrices is explained. The first step in applying the proposed algorithm requires designing a priority scale such as the one proposed in Table 1. In this paper, as previously mentioned, the Lu, Zhang, Ruan, and Wu (2007) linguistic scale is used to illustrate the application of the proposed algorithm. In association with the linguistic values describing a pairwise comparison (represented by a triangular fuzzy number), a crisp number is assigned, as indicated in Table 2. The one-to-one mapping between the linguistic values and the crisp numbers is specifically used for the purpose of consistency evaluation. Similar symmetric numeric scales with a neutral point equal to zero can be defined for other linguistic scales.

Table 2: A crisp numeric scale for consistency evaluation based on the defined linguistic variables and the associated fuzzy numbers (based on Lu et al. (2007) priority scale)

Linguistic Terms	Absolutely less important	Much less important	Less important	Equally important	More important	Much more important	Absolutely more important
Crisp Values for Consistency Evaluation	-3	-2	-1	0	+1	+2	+3
Fuzzy Priority Scale (Triangular Fuzzy Numbers)	$(0, 0, \frac{1}{6})$	$(0, \frac{1}{6}, \frac{1}{3})$	$(\frac{1}{6}, \frac{1}{3}, \frac{1}{2})$	$(\frac{1}{3}, \frac{1}{2}, \frac{2}{3})$	$(\frac{1}{2}, \frac{2}{3}, \frac{5}{6})$	$(\frac{2}{3}, \frac{5}{6}, 1)$	$(\frac{5}{6}, 1, 1)$

The additive nature of the scale can be observed by subtracting adjacent points on the scale. The subtraction values between the pairs of adjacent points on the scale are almost equal. Additive priority can be easily evaluated by crisp numerical scales that can be used in order to avoid the complexity of fuzzy calculations. In Table 2, a crisp number is assigned to each linguistic variable (and the associated fuzzy number). In other words, a crisp numerical scale with additive transitive nature has been designed and used in association with the linguistic and fuzzy numbers to facilitate consistency evaluation of the pairwise comparison matrices.

If $C_i, i = 1, \dots, m$ denote the model attributes and r_{ij} depicts the importance of C_i in comparison to C_j , the corresponding crisp numbers presented in Table 2 should be assigned to r_{ij} s. Figure 1 is a representation of assumptive pairwise comparisons between attributes i, j , and k . In these types of representations, each point on the vertical axis represents one level of preference with regard to the crisp numbers presented in Table 2. For example, in Figure 1, i is shown to be higher than j by 3 units. This means that i is “absolutely more important” than j ($r_{ij}=+3$). j is “less important” than k by 1 unit ($r_{jk} = -1$); therefore, i is expected to be much more important than k ($r_{ik}=+2$).

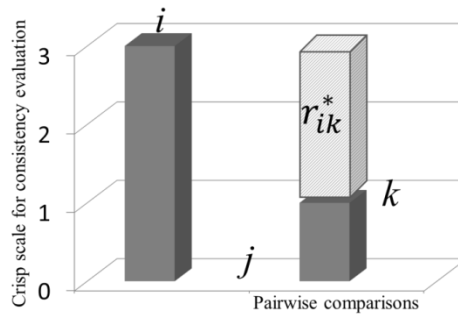


Figure 1: An illustration of comparison between i, j , and k based on Table 2

The value of r_{ik} as it is logically expected based on the expressed values for r_{ij} and r_{jk} is shown by r_{ik}^* . If the expressed value for r_{ik} is different from the logically expected value, its difference from r_{ik}^*

will be considered as a deviation from the perfect consistency situation. The deviation from the expected value of r_{ik} is an indicator of the level of inconsistency. Equation 2 shows the indicator for deviation from perfect consistency (IDPC) among any i, j , and k ($IDPC_{i,j,k}$) of m attributes, where r_{ik}^* denotes the value of r_{ik} as it is expected based on r_{ij} and r_{jk} , and r_{ik} is its value as it has been expressed. In other words, $IDPC_{i,j,k}$ indicates the degree to which the value that a decision maker has expressed for r_{ik} deviates from the value of r_{ik}^* , which is logically expected based on the decision maker's pairwise comparison between i and j , and j and k . IDPC is defined as the difference between r_{ik}^* and r_{ik} . The decision maker can express the pairwise comparisons in any order; the value of IDPC indicates the degree of deviation of a decision maker in providing consistent comparisons within a set of three attributes, regardless of the order of their expression.

$$IDPC_{i,j,k} = |r_{ik}^* - r_{ik}|; \forall i, j, k \in \{1, 2, \dots, m\}; i < j < k \quad (2)$$

In the given example, $r_{ik}^* = +2$; therefore, if $r_{ik} = +1$ is expressed by a decision maker, $IDPC_{i,j,k}$ will be equal to one, which means there is one unit deviation from the consistent judgment. This deviation might have happened when comparing any pair within the set of triple attributes (i.e., i, j , and k); however, the indicator shows that there is one unit error in pairwise comparisons among the three attributes. Therefore, the triple pairwise comparisons and the sets of triple pairwise comparisons that include one or more attributes in common with the inconsistent set of elements should be repeated or eliminated. Accepting, repeating, or eliminating the intransitive sets of elements depends on the decision maker's tolerance range and the degree to which the pairwise comparisons deviate from the consistent condition.

For each pairwise comparison matrix, the number of tests needed to assure the consistency of the matrix is equal to the number of subsets of the set of pairwise comparisons that have cardinality equal to

three . For each subset, an IDPC will be provided so that consistency of the whole matrix will be checked by these tests.

The suggested scale has limits that can cause inaccuracy in calculating IDPCs (any similar scale derived from a linguistic scale will have such limits). Therefore, possible situations that may be subject to error due to the scale limits are examined and a method of correcting errors based on the scale limits presented, prior to discussing the complete algorithm of consistency evaluation. These situations are summarised in Table 3 based on the positive or negative value of pairwise comparisons from the scale presented in Table 2. The neutral point “equally important” has been regarded as a positive value; thus, a positive value indicates a higher or equal priority of the first attribute in comparison to the second.

Table 3: Possible conditions of triple pairwise comparisons based on Table 2

Conditions:	1	2	3	4	5	6	7	8
r_{ij}	+	+	+	+	-	-	-	-
r_{jk}	+	-	+	-	+	-	+	-
r_{ik}	+	+	-	-	+	+	-	-

In this section, using (1) the explained indicator for deviation from perfect consistency (IDPC) presented in Equation 2, (2) the crisp scale of preference presented in Table 2, (3) the possible conditions of pairwise comparisons based on the crisp scale presented in Table 3, the basis of developing the proposed consistency evaluation algorithm is explained. Figures similar to Figure 1 are used to illustrate the conditions and examples, and the steps of the algorithm (presented in section 3.3) are provided in parenthesis to indicate the flow of the algorithm.

Conditions 3 and 6 are in conflict with the transitive property of inequality (i.e., if $a > b$ and $b > c$, then $a > c$). Based on the expressed value of r_{ij} and r_{jk} , r_{ik} contradicts its expected value, which presents a logical paradox that can be interpreted as having the maximum inconsistency (see Figures 2 and 3). In the proposed consistency evaluation algorithm, these two conditions have been recognised by

the rule below, and the maximum inconsistency value has been assigned to them, as a fundamental principle would be breached in these two conditions. This issue is identified in the algorithm, and the maximum inconsistency score is assigned to it by the use of the following rule (step 1), where n denotes the number of scale points:

If $\{r_{ij}, r_{jk}, r_{ik}\} \in$

$$\{r_{ij}, r_{jk}, r_{ik} \mid r_{ij} \cdot r_{jk} \geq 0 \wedge r_{jk} \cdot r_{ik} < 0; i > j > k\} \Rightarrow \text{IDPC}_{ijk} = \text{maximum Inconsistency} = n - 1$$

This rule, which is reflected in the first step of the proposed algorithm, considers zero tolerance to inconsistency in terms of conflicting priorities when the order of priorities is logically unacceptable. For example, when i is preferred to j and j is preferred to k , the preference of k over i will not be tolerated to any degree, and the maximum degree of inconsistency will be considered for this set of pairwise comparisons. This is an arbitrary decision depending on how important this deviation is to decision makers. If the user does not consider these conditions to be unacceptable, they can be handled like any other deviations; if so, the IDPC can be calculated using Equation 2.

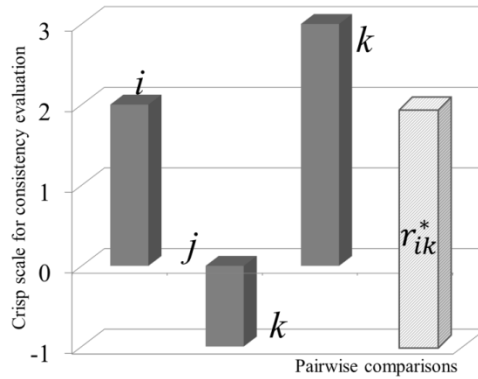


Figure 2: The third condition based on Table 3

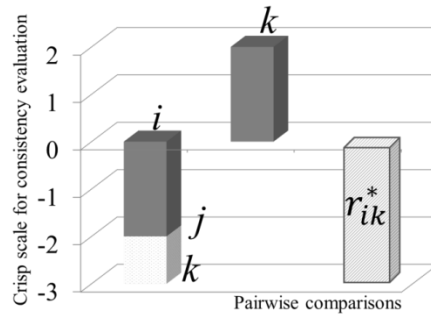


Figure 3: The sixth condition based on Table 3

Another problem that occurs when calculating the indicator using the scale presented in Table 2 is that, in some cases, the expected value of one of the pairwise comparisons (r_{ik} for example) exceeds the scale bounds, based on the other two pairwise comparisons in the set of triple elements (r_{ij}, r_{jk}, r_{ik}). This happens if and only if the following statement is true (n =number of scale points):

$$\{r_{ij}, r_{jk}, r_{ik}\} \in \left\{r_{ij}, r_{jk}, r_{ik} \mid |r_{ij}| + |r_{jk}| > \frac{n-1}{2} \vee |r_{jk}| + |r_{ik}| > \frac{n-1}{2} \vee |r_{ij}| + |r_{ik}| > \frac{n-1}{2}; i > j > k \right\}$$

In such a case (recognised in step 2), the decision maker has no choice but to adapt the expressed value to the scale if the choice is limited to the linguistic terms on a predetermined linguistic scale. In other words, since the scale is limited and it is not possible to choose expressions with lower or greater degrees of significance than the scale points, the decision maker's choice of expression may not be completely consistent with his/her prior pairwise comparisons. In this case, r_{ik}^* is not necessarily equal to the summation of r_{ij} and r_{jk} . In such a situation, if condition 1 or 8 applies (so premise in step 3 is rejected), r_{ik}^* still follows the $r_{ik}^* = r_{ij} + r_{jk}$ rule but has to be adapted to the scale. Figures 4 and 5 illustrate such situations: $|r_{ik}|$ is required to be larger than the scale after pairwise comparisons between i and j , and j and k ; however, the required value could not be assigned, as the scale points are limited.

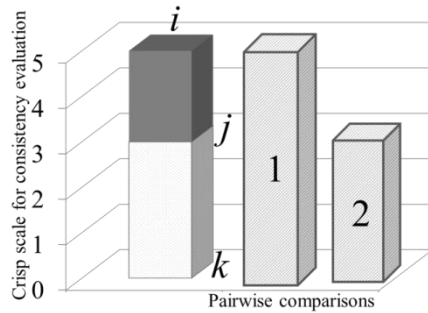


Figure 4: Adaptation to scale under the first condition

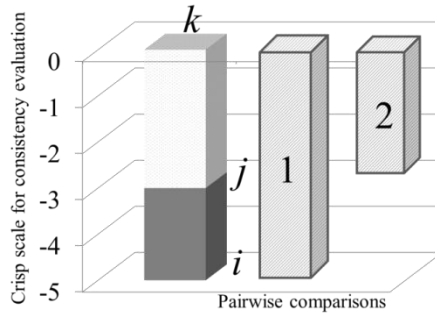


Figure 5: Adaptation to scale under the eighth condition

In the situation presented in Figure 4, r_{ik}^* is equal to +5 as shown in column 1; however, a corresponding linguistic term does not exist in the proposed scale and the decision maker is limited to choose from the provided linguistic terms. Hence, the decision maker has no option but to express this value as “absolutely more important” ($r_{ik}=+3$), as shown in column 2. However, as the consistency test needs r_{ik}^* to be compared with the real value of r_{ik} (i.e., value of r_{ik} as it would be expressed if the scale was not limited), such an expression cannot be considered as a deviation from consistency on the part of the decision maker given the absence of a suitable choice due to the limits of the scale. Therefore, $r_{ik}^*=+5$ should be used and the expressed value of r_{ik} should be replaced by r_{ik}^* as the basis of the comparison ($r_{ik} = r_{ik}^* = +5$), and the comparison should be regarded as consistent (so premise in step 4 is rejected and r_{ik}^* will be modified in step 5). Figure 5 indicates a similar situation in the negative area of scale. For both cases, the algorithm considers $r_{ik}^* = r_{ij} + r_{jk}$ (step 4) so that, regardless of the expressed value of r_{ik} , the real and non-modified value is considered in the calculation of IDPC (steps 6 and 7). With such a strategy, IDPC shows the level of inconsistency among the three elements, regardless of the order of conducting the pairwise comparisons. For conditions 2 and 7, the pairwise comparisons will never exceed the scale bounds. Figures 6 and 7 show the extreme values in these two conditions; the values have remained within the scale bounds (premise rejected in step 2, applies in step 4). When conditions 4 or 5 apply in cases that exceed the scale bound, the value of r_{jk} can be subject to error based on the scale. Pairwise comparisons are not necessarily conducted in a certain order; thus, in conditions 4 or 5, if a decision maker determines the values of r_{ij} and r_{ik} prior to r_{jk} , the decision maker has to modify the value of r_{jk} based on the available options on the scale. Figure 8 gives an example of such a situation: $r_{ij}=+3$ and $r_{ik}=-1$ are expressed by a decision maker, but the decision maker cannot evaluate r_{jk} such that $r_{jk} = -4$, as the value does not exist on the scale. Figure 9 also indicates a similar example under condition 5. In both cases, the proposed algorithm determines and uses the non-modified values of r_{jk} . Regardless of the order in which the pairwise comparisons are conducted, the algorithm determines the condition and avoids the possible error caused by the limitation of the scale (premise applies in step 3 and step 4).

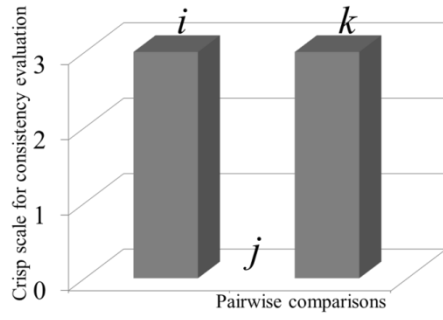


Figure 6: The second condition with maximum comparison degrees

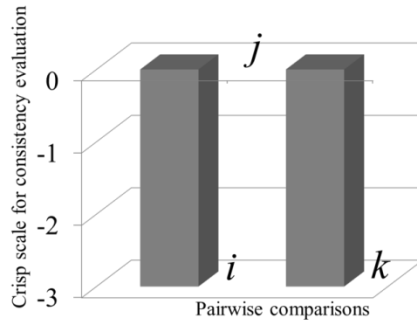


Figure 7: The seventh condition with maximum comparison degrees

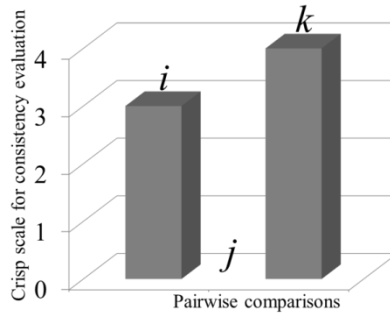


Figure 8: Excess from the scale bounds under the fourth condition

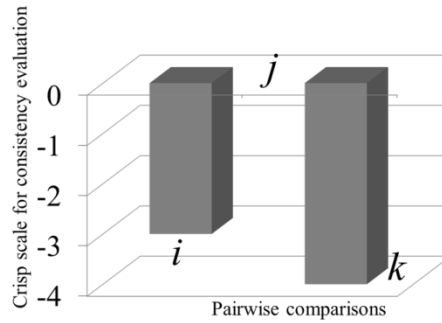


Figure 9: Excess from the scale bounds under the fifth condition

Once all of the conditions of exceeding the scale bounds are recognised and correction methods are applied, the data are ready to be processed and IDPCs calculated. These steps have been summarised in an algorithm shown in its general format in Figure 10 and in more detail in Section 3.3.

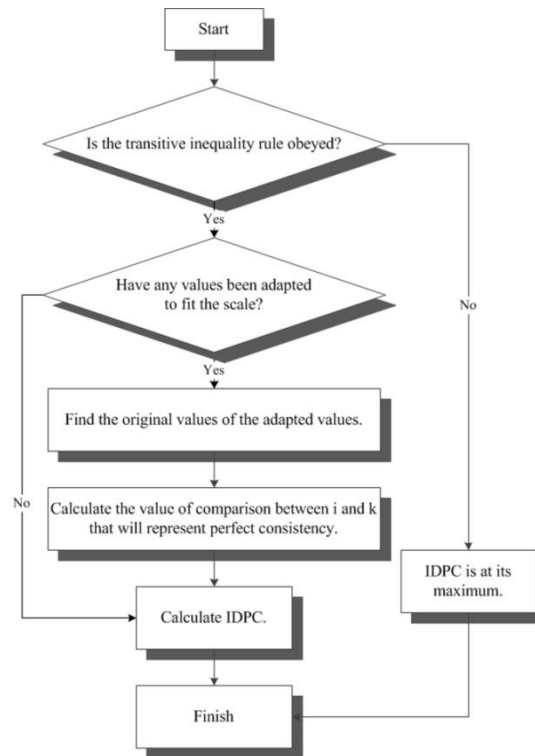


Figure 10: The general consistency evaluation algorithm for each set of independent triple pairwise comparisons

The outcome of the proposed algorithm is a set of IDPCs for each matrix that measures the level of consistency of pairwise comparisons and indicates the sets of pairwise comparisons that are intransitive. By defining an acceptable tolerance threshold, users can choose a desired level of consistency. However, while a strict tolerance threshold may increase the accuracy of the output of the MADM method, it can result in more data that is subject to elimination or recollection. The proposed method enables the capability of measuring the degree of deviation from perfect consistency and identifying the source of inconsistency. Thus, decision makers can be aware of the extent to which the pairwise comparison matrix is inconsistent and which comparison has to be repeated or eliminated. Once the intransitive comparisons are identified, pairwise comparisons that do not include the elements involved in the intransitive comparisons can be accepted as reliable data, while the intransitive comparisons have to be repeated or rejected.

When applied to the example previously mentioned in section 3.1, the proposed algorithm does not produce the errors that result from the application of the consistency evaluation rules of Lu, Zhang, Ruan, and Wu (2007). In this example, i is anticipated to be “more important” than p . Any other representation of this value will be considered as an error to some degree. For example, if i is expressed to be “less important” than p , then, under condition 5, $IDPC_{i,j,p} = |r_{ip}^* - r_{ip}| = |+1 - (-1)| = 2$. This value would be considered as consistent based on Lu et al. (2007), though logically it is not.

3.3 The Proposed Algorithm for Consistency Evaluation in Pairwise Comparison-Based

MADM methods

The complete proposed algorithm for consistency evaluation is presented in this section. Steps 1 to 7 must be followed for all sets of triple elements of the comparison matrix (i.e., $\forall i, j, \text{ and } k \in \{1, 2, \dots, m\}; i \neq j \neq k; m = \text{number of attributes of the MADM model}$).

Step 1:

$n = \text{number of scale points}$

If $\{r_{ij}, r_{jk}, r_{ik}\} \in$

$\{r_{ij}, r_{jk}, r_{ik} | r_{ij} \cdot r_{jk} \geq 0 \wedge r_{jk} \cdot r_{ik} < 0; i > j > k\} \Rightarrow IDPC_{ijk} = \text{maximum Inconsistency} = n - 1$

Otherwise: Go to step 2

Step 2:

$$\text{If } \{r_{ij}, r_{jk}, r_{ik}\} \in \left\{ r_{ij}, r_{jk}, r_{ik} \left| \begin{array}{l} |r_{ij}| + |r_{jk}| > \frac{n-1}{2} \vee |r_{jk}| + |r_{ik}| > \frac{n-1}{2} \vee |r_{ij}| + |r_{ik}| > \frac{n-1}{2}; \\ i > j > k \end{array} \right. \right\}$$

⇒ Go to step 3

Otherwise: Go to step 4

Step 3:

$$\text{If } \{r_{ij}, r_{jk}, r_{ik}\} \in \{r_{ij}, r_{jk}, r_{ik} \mid r_{ij} \cdot r_{jk} < 0 \wedge r_{jk} \cdot r_{ik} \geq 0; i > j > k\}$$

⇒ $r_{jk} = r_{ik} - r_{ij}$; Go to step 4

Otherwise: Go to step 4

Step 4:

$$\text{If } -\frac{n-1}{2} \leq r_{ij} + r_{jk} \leq \frac{n-1}{2}$$

⇒ $r_{ik}^* = r_{ij} + r_{jk}$; Go to step 6

Otherwise: Go to step 5

Step 5:

$$\text{If } r_{ij} + r_{jk} > \frac{n-1}{2}$$

⇒ $r_{ik}^* = \frac{n-1}{2}$; Go to step 6

Otherwise: $r_{ik}^* = -\frac{n-1}{2}$; Go to step 6

Step 6:

If $r_{ik} \cdot r_{ik}^* \geq 0$

$\Rightarrow t = |r_{ik} - r_{ik}^*|$; Go to step 7

Otherwise: $t = |r_{ik} + r_{ik}^*|$; Go to step 7

Step 7:

If $t \leq n - 1$

$\Rightarrow IDPC_{ijk} = t$

Otherwise: $IDPC_{ijk} = \text{maximum Inconsistency} = n-1$

4. Application of the Proposed MADM Method to a Flood Mitigation Project Selection

Case

The proposed consistency evaluation method and the Lu et al. (2007) MADM method were applied to a flood mitigation project selection case on the Gold Coast, Australia. This study was conducted in 2009 as research exploring the applicability and functionality of the proposed methodology to decision problems of this nature. A previous study conducted by the Gold Coast City Council was used as the baseline for this research. The Council's study considered a variety of aspects regarding to this decision problem, and applicable options based on budget availability and technical investigations were identified. The decision of which option to implement was made throughout a number of meetings and consultation with area residents and flood mitigation experts.

In the case study conducted for this paper, given the variety of aspects, attributes, and alternatives to be considered, as well as the varied interests of the different stakeholders, using an MADM method was considered a valid approach.

This case study involved developing a hierarchical decision model to assist in selecting a flood mitigation option. Development of the decision-model was based on a literature review of flood management, the previous study conducted by the Gold Coast City Council, and interviews with three experienced subject experts in flood mitigation studies (modelling process). The study also captured different viewpoints of stakeholders regarding a variety of aspects of the problem and provided an overall ranking of options with regard to these considerations. The result of the case study was consistent with

the result of the city council's initial study and was assessed as satisfactory by the three experts who were interviewed for modelling the problem.

4.1 The decision model and data collection

The Gold Coast is one of the areas in Australia that has a major likelihood of flooding. Therefore, the Gold Coast City Council studied the possible flood mitigation methods that suit the specific condition of the Nerang River catchment. Based on technical investigations, four options were suggested as feasible options: raising the Hinze dam, river dredging, modification of the flood channel, and bridge modification. The same options have been considered as the proposed alternatives of the decision model in this case study (Gold Coast City Council website, 2009).

In this case study, the goal was defined as reducing the level of vulnerability of the society and environment to flood hazards. Based on the previous study by the Gold Coast City Council and interviews with flood mitigation experts, the four classes of attributes were considered with respect to the defined goal that are: financial efficiency, time (project duration), protection level, and sustainability. In order to distinguish the interests more explicitly, protection level was further divided into two attributes: community protection and environmental protection, and sustainability was subdivided into two basic concerns: flexibility and durability. These divisions allowed consideration of possible conflicts of interests among groups of decision makers.

In this application, preferences at the attribute comparison level could be determined by all the decision makers. However, at the alternative comparison level some determinations, such as those involving financial efficiency, required a certain level of expertise and technical knowledge. Therefore, two types of questionnaires were distributed to capture information on these two different levels. At the attribute comparison level, one questionnaire was distributed among participants in three groups: residents, business owners, and city council managers; at the alternative comparison level a second questionnaire, which contained items requiring technical knowledge or expertise, was distributed among flood mitigation technical experts.

The questionnaires were distributed among 25 residents, 15 business owners, 7 city council managers, and 5 flood mitigation technical experts; 19, 11, 6, and 5 questionnaires were completed and returned from each group, respectively, and 16, 9, 5, and 5 results from each group, respectively, were recognised as adequately consistent and reliable data.

4.2 The decision-making process

After collecting data, the first step in applying the proposed method is to check the consistency level of the pairwise comparisons. After the proposed consistency evaluation algorithm was applied, six matrices were recognised as not adequately consistent. As a result, the sets of triple attributes producing pairwise comparisons that showed intransitivity and the sets of triple attributes affected by the intransitivity were recognised and eliminated from the inconsistent matrices. In other words, only the consistent part of each matrix was kept as input data. Another option would have been to re-administer those questions in the questionnaire that contain one of the elements of the intransitive sets; this can be an important alternative in situations in which the information provided by the removed data is critical, but that was not the case in this study. For the consistency evaluation phase, the tolerance threshold was assumed to be equal to two points of the comparison scale. In other words, a deviation from the perfect consistency condition (all IDPCs=0) of equal or less than two points of the scale was tolerated and considered to be adequately consistent (tolerance threshold is subject to how critical the accuracy of pairwise comparisons is in a decision problem and will vary depending on the nature of the problem). The pairwise comparison matrices were based on the results of the questionnaires, which collected linguistic data in a pairwise manner.

As an example, Table 4 shows one of the pairwise comparison matrices using the crisp scale provided in Table 2. Applying the algorithm presented in Section 3.3 to the attributes of financial efficiency, time, and flexibility, measuring the inconsistency of this matrix is as follows:

$$r_{12} = 0, r_{25} = 3, r_{15} = 0 \Rightarrow IDPC_{125} = 3$$

While financial efficiency is considered to be “equally important” compared to time, time is considered to be “absolutely more important” (three points of scale) than flexibility; financial efficiency is expected to be “absolutely more” important (three points of scale) than flexibility, while “equally important” has been expressed. $IDPC_{125} = 3$ indicates that the decision maker’s comparisons between elements 1, 2, and 5—financial efficiency, time, and flexibility, respectively—deviates from a perfect transitive situation by three points on the scale. Since this measure is beyond the tolerance threshold, one of three options must be taken: (i) reject the matrix; (ii) filter the inconsistent data in the matrix by removing all the intransitive pairwise comparisons that include any of the intransitive elements; or (iii) repeat the pairwise comparisons that include any of these elements. In this case, any of the pairwise comparisons that contain attributes 1, 2 and 5 are subject to corrective repetition or elimination from the matrix. If some of the assessments are repeated, the entire matrix must be checked for consistency again. To ensure that the matrix is adequately consistent, the rest of the IDPCs for the matrix should also be calculated.

Table 4: An example of an inconsistent comparison matrix

	Financial Efficiency	Time	Community Protection	Environmental Protection	Flexibility	Durability
Financial Efficiency	0	0	-3	-2	0	+3
Time	+2	0	-1	0	+3	-1
Community Protection	+3	+1	0	+2	+2	+1
Environmental Protection	+2	0	-2	0	+1	+2
Flexibility	-2	-2	-2	-1	0	-2
Durability	-3	+1	-1	-2	+2	0

The next step of the study was quantifying the linguistic data by assigning fuzzy numbers to the components of the adequately consistent linguistic data. For each group of participants, fuzzy numbers provided by participants for each pairwise comparison were aggregated into one fuzzy number representing the overall judgement of the group on the particular pairwise comparison. In this case study, arithmetic mean was used to aggregate the results. However, the suitability of using other methods of aggregating fuzzy numbers is the subject of future research.

The final result, given in Figure 12, shows that raising the Hinze Dam ranked best in satisfying the considered attributes, followed by river dredging, Benowa flood channel modification, and bridge modifications, respectively. The closeness coefficients are the values of the TOPSIS index for each alternative. TOPSIS index is a synthesised value for each alternative and indicates the overall level of satisfaction in vulnerability reduction for each alternative. Hence, the closeness coefficients provide a basis for investigating the degree of preference for each alternative over the others.

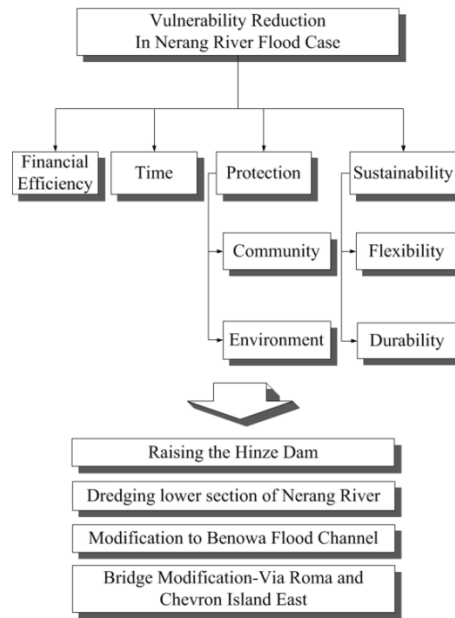


Figure 11: The hierarchical decision model for the flood mitigation project selection case

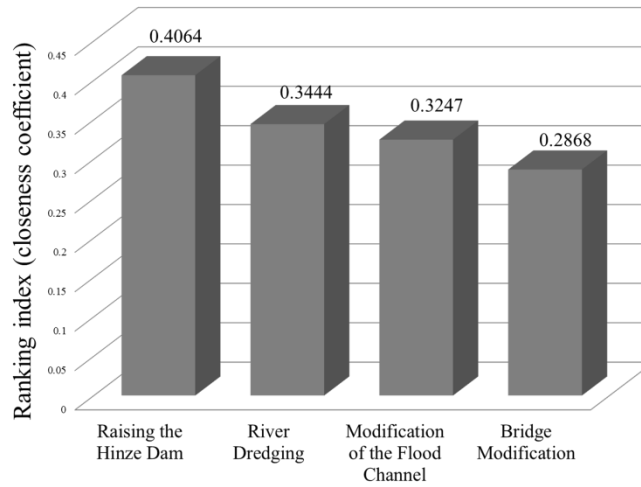


Figure 12: Ranking of the flood mitigation options

5. Conclusion

In this paper, fuzzy multi-attribute decision-making methods based on pairwise comparisons have been reviewed, and an algorithm for evaluating the consistency level of linguistic pairwise comparisons has been proposed. This algorithm determines the source of inconsistency by identifying the intransitive set of triple elements of a matrix; thus only pairwise comparisons involving these elements can be repeated or rejected, instead of rejecting or repeating the entire matrix.

Finally, the modified method has been applied to a flood mitigation project selection case on the Gold Coast, Australia. A brief review of the process and the results of the application of the method have been provided. The result was consistent with the Gold Coast City Council's initial study and improved on the original study by including a wider range of attributes, and producing a ranking of all options. Future work will explore the possibility of improving the priority scale in order to capture the opinions of decision makers more efficiently. As well, the development of a software tool that will dynamically evaluate consistency of data as it is being entered and prompt the user to correct problematic (i.e., intransitive) comparisons is being explored.

Acknowledgements

The authors would like to thank Professor Witold Pedrycz from the University of Alberta for his valuable input to this paper.

References

- Ali, I and W D Cook and M Kress (1986). Ordinal ranking and intensity of preference: A linear programming approach. *Management Science*, 32, 1642-1647.
- Arbel, A (1989). Approximate articulation of preference and priority derivation. *European Journal of Operational Research*, 43, 317-326.
- Boender, C G E and J G Graan and F A Lootsma (1989). Multi-criteria decision analysis with fuzzy pairwise comparisons. *Fuzzy Sets and Systems*, 29, 133-143.
- Chang, D Y (1996). Applications of the extent analysis method on fuzzy AHP. *European Journal of Operational Research*, 95, 649-655.
- Deng, H (1999). Multicriteria analysis with fuzzy pairwise comparison. *International Journal of Approximate Reasoning*, 21, 215-231.
- Flood Mitigation Fact sheet, Gold Coast City Council Website, Accessed August 2009, available at http://www.goldcoast.qld.gov.au/attachment/flood_mitigation_infosheet_reduce_risk.pdf.
- Guh, Y Y and R W Po and K R Lou (2009). An additive scale model for the analytical hierarchy process. *International Journal of Information and Management Science*, 20, 71-88.
- Harker, P T and L G Vargas (1987). The theory of ratio scale estimation: Saaty's analytic hierarchy process. *Management Science*, 33, 1383-1403.
- Herrera-Viedma, E, F Herrera, F Chiclana and M Luque (2004). Some issues on consistency of fuzzy preference relations. *European Journal of Operational Research*, 154, 98-109.
- Hochbaum, D S and A Levin (2006). Methodologies and algorithms for group-rankings decision. *Management Science*, 52, 1394-1408.
- Hwang, C L and P K Yoon (1981). *Multiple attribute decision making: methods and applications*. Berlin: Springer-Verlag.
- Kahraman, C (2008). *Fuzzy multi-criteria decision-making: Theory and applications with recent developments*. New York, USA: Springer Science+Business Media.

- Kahraman, C and Z Ulukan and E Tolga (1998). *A fuzzy weighted evaluation method using objective and subjective measures*. International ICSC Symposium on Engineering of Intelligent Systems (EIS'98), University of La Laguna Tenerife, Spain, 1998.
- Leung, L C and D Cao (2000). On consistency and ranking of alternatives in fuzzy AHP. *European Journal of Operational Research*, 124, 102-113.
- Lootsma, F A (1988). Numerical scaling of human judgement in pairwise-comparison methods for fuzzy multi-criteria decision analysis, In H J Greenberg and F A Lootsma and M J Rijckaert and H J Zimmermann and G Mitra (Eds.), *Mathematical models for decision support*. New York, NY: Springer-Verlag.
- Lootsma, F A (1993). Scale sensitivity in the multiplicative AHP and SMART. *Journal of Multi-Criteria Decision Analysis*, 2, 87-110.
- Lu, J and G Zhang and D. Ruan and F Wu (2007). *Multi-objective group decision making: Methods, software and applications with fuzzy set techniques*. London: Imperial College Press.
- Saaty, T L (1980). *The analytical hierarchy process*. New York, NY: McGraw-Hill.
- Saaty, T L and P K Kearns (1985). *Analytical planning: the organization of systems*, Oxford: Pergamon Press.
- Salo, A A and R P Hämäläinen (1995). Preference programming through approximate ratio comparisons. *European Journal of Operational Research*, 82, 458-475.
- Salo, A A (1996). On fuzzy ratio comparisons in hierarchical decision models. *Fuzzy Sets and Systems*, 84, 21-32.
- Triantaphyllou, E and F A Lootsma, P M Pardalos and S H Mann (1994). On the evaluation and application of different scales for quantifying pairwise comparisons in fuzzy sets. *Journal of Multi-Criteria Decision Analysis*, 3(3), 133-155.
- Van Laarhoven, P J M and W Pedrycz (1983). A fuzzy extension of Saaty's priority theory. *Fuzzy Sets and Systems*, 11, 199-227.
- Zanakis, S H, A Solomon, N Wishart and S Dublisch (1998). Multi-attribute decision making: A simulation comparison of select methods. *European Journal of Operational Research*, 107, 507-529.