

Against Artificial Complexification: Crisp vs. Fuzzy Information in the TOPSIS Method

David A. Pelta and María T. Lamata and José L. Verdegay and Carlos Cruz and Ana Salas

Models of Decision and Optimization Research Group
Department of Comp. Science and A.I, Universidad de Granada, Spain
{dpelta, mtl, verdegay, carlosacruz}@ugr.es

Abstract

The question of whether the use of crisp or fuzzy input information in the TOPSIS method produces a different ranking is explored. Using a basic representation of fuzziness through triangular fuzzy numbers, a set of randomly generated fuzzy and crisp multicriteria decision problems are solved. Then, the corresponding rankings are compared and variations in the top alternative are studied.

The results show that changes in the top alternative are minor. This situation, coupled with the fact that the “true” ranking is unknown and that more complex models of “fuzzy” information require a huge amount of *precise* information from the decision maker side, raise the discussion of whether in this specific context a “complexification” of the input data makes sense.

Keywords: decision making, TOPSIS, fuzzy sets, information representation

1 Introduction

The selection of an item from a set, giving their characteristics/criteria and our preferences is an ubiquitous problem: for example, replace “item” by mobile phone, or by an university to study, or by a hotel in the city of our next conference and a multicriteria decision making (MCDM) problem arises.

In its most basic setting, where all the involved information is precisely known, a MCDM problem can be represented using a decision matrix as the one shown in Table 1. There are m rows, each one associated with an alternative $\{A_1, A_2, \dots, A_m\}$. Every column (out of n) is associated with a set of criteria $\{C_1, C_2, \dots, C_n\}$. The value of the alternative A_i under criterion C_j is denoted as x_{ij} . Finally, it is assumed that a decision

MCDM	w_1	w_2	w_j	\dots	w_n
	C_1	C_2	C_j	\dots	C_n
A_1	x_{11}	x_{12}	\dots	\dots	x_{1n}
A_2	x_{21}	x_{22}	\dots	\dots	x_{2n}
\dots	\dots	\dots	\dots	\dots	\dots
A_i	\dots	\dots	x_{ij}	\dots	\dots
A_m	x_{m1}	x_{m2}	\dots	\dots	x_{mn}

Table 1: Decision matrix of a MCDM problem.

maker is able to reflect the importance of the criteria using a vector of n weights $W = \{w_1, w_2, \dots, w_n\}$.

Given such decision matrix A and the vector of weights W , a MCDM method or algorithm calculates a score of the alternatives, that is then used to produce a ranking.

The number of MCDM methods is large and it is not our intention here to review them. The interested reader can check recent books like [12] for further information.

There are two important facts about these methods: 1) given the same MCDM problem, different methods may lead to different results, as it have been shown in [2, 3], and 2) as long as we know, there is no “true ranking” to assess if a given method produces the right ranking.

The situation is worst in many real life situations, where the values x_{ij} could not be known precisely. In a situation where instead of saying the value of A_i under C_j is *exactly* x_{ij} , we are allowed to say (for example) *around* x_{ij} , then fuzzy numbers [9] would be suitable tools to model such kind of fuzziness leading to fuzzy multicriteria decision making problems [14, 16].

Several options are available to represent the values x_{ij} , so we can associate a specific type to such values, as “variables” are assigned a type in some programming languages. In our case, examples of data

types for x_{ij} could be integer, float, interval, triangular fuzzy number and so on.

Thus depending on the data type considered, different adaptations of a given MCDM method can be obtained. Here we focus in the TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) method because its wide adaptation to different input data types.

Since its proposal in [13], we can find TOPSIS for triangular fuzzy numbers [5], the interval-valued fuzzy TOPSIS [6], TOPSIS with linguistic variables [1], a proportional interval type-2 hesitant fuzzy TOPSIS approach [7] or a recent adaptation to deal with interval-valued spherical fuzzy sets [11].

Every new proposal for modeling the parameters of a MCDM problem can be associated with a new potential data type. As the type becomes more complex, it will require an extra effort to define a single value of such type. For example, a value of type trapezoidal fuzzy number requires more information than a value of type triangular fuzzy number. But the former requires less information than that of an interval-valued fuzzy number. If we consider type-2 fuzzy sets the situation becomes worst. Examples of data types for modeling fuzzy numbers are shown in Fig. 1. It should be noted that as the representations become more complex, there is an increasing need for *precise information* to define the corresponding values.

In general, despite the data type used for the input, the MCDM methods ends up with a score for every alternative from which a ranking is derived. This ranking, understood as a permutation over the alternatives, can be considered as a “crisp” object.

In this context, the aim of this paper is to make a step back and explore if the TOPSIS method operating on single numerical values and its adaptation to deal with fuzziness in its most basic form (i.e. using triangular fuzzy numbers, Fuzzy-TOPSIS), produces (or not) substantially different rankings of the alternatives.

In a previous research [4] we explore a similar situation using a different MCDM method (VIKOR) in a very basic setting. The main finding was that the rankings have not too much differences.

Now, using the TOPSIS method as an example, we pose the following question: *when considering fuzzy triangular numbers in the input of a MCDM problem, does the corresponding output substantially differs from that obtained when crisp values are considered?*

Besides this, we want to explore if the number of alternatives, criteria and level of spread (the width of

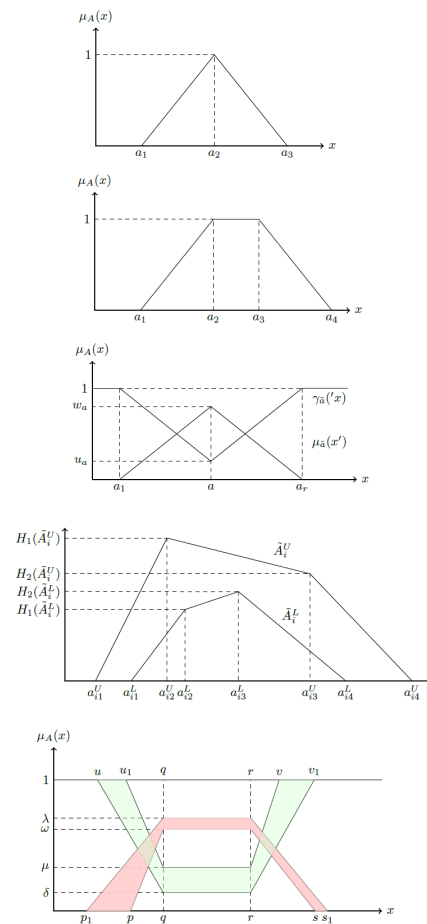


Figure 1: Representations of different fuzzy numbers. From top to bottom: triangular, trapezoidal, triangular intuitionistic, interval type-2 and interval valued intuitionistic. The more complex the representation is, a highest amount of “precise” information is required from the decision maker side.

the left/right values in the triangular fuzzy numbers) in the input data have any impact in the potential rankings’ discrepancies (these factors were not taken into account in [4]). In order to shed light into this topic, we design and perform a simulation based experiment to assess such impact.

This paper is organized as follows. Section 2 introduces the TOPSIS method and discuss its variants. In Section 3 we describe the computational experiments and the results obtained. Finally, Section 5 summarizes the conclusions of our research.

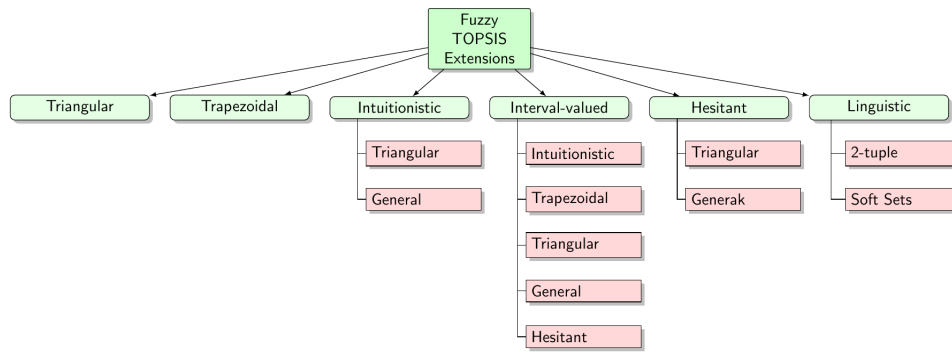


Figure 2: Taxonomy of TOPSIS extensions (adapted from [17]).

2 TOPSIS and its extensions

The TOPSIS method evaluates the alternatives in terms of their distance to the so-called “positive” and “negative” ideal solutions. Given a decision matrix and a vector of weights, the TOPSIS algorithm can be summarized as follows:

1. Normalize the decision matrix using some norm.
2. Calculate the weighted normalized decision matrix.
3. Determine the positive ideal solution (PIS), A^+ , and the negative ideal solution (NIS), A^- .
4. Calculate the Euclidean distance between every alternative A_i and A^+ (namely, d_i^+), and A^- (namely, d_i^-).
5. Calculate the “closeness” index of every alternative A_i to both ideal solutions:

$$CC_i = \frac{d_i^-}{d_i^+ + d_i^-}$$

where $i = 1, 2, \dots, m$. When $CC_i = 0$, then $d_i^- = 0$ and the alternative A_i is the worst one. On the other hand, when $CC_i = 1$, then $d_i^+ = 0$ then A_i is the best one. It holds that $0 \leq CC_i \leq 1$.

6. Rank the alternatives according to CC_i in descending order. The best alternative will have the highest CC_i .

From this simple scheme, more than a dozen of TOPSIS extensions were proposed to deal with different types of input values. Let’s call this data type \mathcal{F} and the corresponding extension \mathcal{F} -TOPSIS. In a recent survey paper [17], authors produced a very illustrative taxonomy about those different TOPSIS extensions in

MCDM	\tilde{w}_1	\tilde{w}_2	\dots	\tilde{w}_n
	C_1	C_2	\dots	C_n
A_1	\tilde{x}_{11}	\tilde{x}_{12}	\dots	\tilde{x}_{1n}
A_2	\tilde{x}_{21}	\tilde{x}_{22}	\dots	\tilde{x}_{2n}
\dots	\dots	\dots	\tilde{x}_{ij}	\dots
A_m	\tilde{x}_{m1}	\tilde{x}_{m2}	\dots	\tilde{x}_{mn}

Table 2: Fuzzy decision matrix: a decision matrix with values of type triangular fuzzy number.

terms of the data type used. An adaptation of such taxonomy is shown in Fig 2.

In order to produce a new TOPSIS extension designed to work with a new data type, the following steps (in very basic terms) should be taken.

1. Every value in the decision matrix and in the vector of weights is given as a type \mathcal{F} value.
2. The basic mathematical operations are adapted to deal with type \mathcal{F} operands.
3. Some distance function between vectors of type \mathcal{F} values (the alternatives) is provided.
4. Having obtained the CC_i values represented as a type \mathcal{F} value, two options can exist: 1) to provide a relational operator (\geq, \leq) to sort the CC_i values, or 2) to use a transformation function that maps a value of type \mathcal{F} to \mathfrak{R} (like a defuzzification operator)

Of course, one expects that all the operations, distance function and comparisons are properly defined from a mathematical point of view. If, for some type \mathcal{F} , any of the previous step failed, then TOPSIS method can not be adapted to deal with \mathcal{F} values.

In this paper, besides considering integer values at the input, we will consider $\mathcal{F} = TFN$ (triangular fuzzy number) leading to fuzzy decision making problems as shown in Table 2. From an algorithmic point of view, such *TFN-TOPSIS* (named as Fuzzy-TOPSIS in what follows) replaces the “classic arithmetic” with the triangular fuzzy numbers’ arithmetic. Let’s consider $\tilde{A} = (a_1, a_2, a_3)$ and $\tilde{B} = (b_1, b_2, b_3)$ two values of type triangular fuzzy number, with $a_i, b_i \geq 0, i = 1, 2, 3$. A potential set of basic operations is described in Table 3 (of course, other ones are possible).

3 Experiments and Results

Using a simulation based approach, we want assess if there are any noticeably differences in the rankings obtained when using either numerical single-value (crisp) or TFN (fuzzy) data types in the decision matrix. We will firstly solve a set of randomly generated fuzzy decision making problems and their derived crisp versions and secondly, we will compare the corresponding rankings using different measures.

Data Generation

We consider four factors to define a fuzzy MCDM problem: the number of criteria and their importance, the number of alternatives, and the level of spread in the TFNs.

For every factor, we detail below the levels considered:

- Criteria: $n = \{5, 10, 15\}$. All of them are given the same importance.
- Alternatives: $m = \{5, 10, 25, 50, 75, 100\}$
- Spread level: $\gamma = \{10, 20, 30\}$, to determine the maximum spread to the left/right of the triangular fuzzy number.

For each combination of criteria, alternatives and spread level, we construct 100 decision matrices following the steps detailed below.

We first define a fuzzy decision making problems dataset (*FDP*), where each element $\tilde{d}p_k \in FDP, k = 1, \dots, 100$ is a decision matrix. Each element in the matrix is a value of type TFN: triangular fuzzy number \tilde{x}_{ij} randomly generated as follows:

$$\tilde{x}_{ij} = (x_{ij1}, x_{ij2}, x_{ij3}) = \begin{cases} x_{ij2} = U(1, 100) \\ x_{ij1} = x_{ij2} - \Delta \\ x_{ij3} = x_{ij2} + \Delta \end{cases}$$

where $U(\min, \max)$ is a function that returns a

random integer number between $[\min, \max]$ and $\Delta = U(1, \gamma)$ (being γ the level of spread). As can be seen, the triangular fuzzy numbers are symmetric.

Regarding the criteria, and for the sake of simplicity, we assumed that all of them are equally important and should be maximized. They are defined as follows:

$$\tilde{w}_j = (w_{j1}, w_{j2}, w_{j3}) = \begin{cases} w_{j2} = 1/n \\ w_{j1} = w_{j2} - 10\% \\ w_{j3} = w_{j2} + 10\% \end{cases}$$

Departing from the set *FDP* we construct a crisp dataset *CDP*, where each $dp_k \in CDP, k = 1, \dots, 100$, is the “defuzzified” version of the corresponding $\tilde{d}p_k$. Recalling that $\tilde{x}_{ij} = (x_{ij1}, x_{ij2}, x_{ij3})$, we define the crisp value $x_{ij} = x_{ij2}$. In these crisp problems, the weights are defined as $w_j = 1/n$.

Finally, we end up with 5400 problems with fuzzy information modeled with triangular fuzzy numbers and their corresponding 5400 MCDM problems with crisp information modeled with integer values.

TOPSIS Algorithm

Then, the problems $(\tilde{d}p_k, dp_k), k = 1, \dots, 5400$ are solved by the following methods: a) Fuzzy-TOPSIS and b) TOPSIS, thus obtaining the rankings: r_k^f, r_k^c , respectively. Both TOPSIS methods use linear normalization.

The computational experiments have been done using the TOPSIS implementations provided in the MCDM [15] package, and Fuzzy TOPSIS implementation as provided in the FuzzyMCDM [10] package. Both packages are available on the CRAN repository [8].

Measures

So for each problem k we have two rankings (r_k^c, r_k^f) to be compared. Let’s denote as $r(t), t = 1 \dots m$, the alternative in position t of the ranking r . We consider the following three functions for ranking comparison:

The top alternative is equal

$$top(r_k^c, r_k^f) = \begin{cases} 1 & \text{if } r_k^c(1) = r_k^f(1) \\ 0 & \text{otherwise} \end{cases}$$

Top two alternatives are swapped

$$swap(r_k^c, r_k^f) = \begin{cases} 1 & \text{if } r_k^c(1) = r_k^f(2) \\ & \text{and } r_k^c(2) = r_k^f(1) \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned}
 \text{Addition : } \tilde{A} \oplus \tilde{B} &= (a_1 + b_1, a_2 + b_2, a_3 + b_3) \\
 \text{Subtraction : } \tilde{A} \ominus \tilde{B} &= (a_1 - b_3, a_2 - b_2, a_3 - b_1) \\
 \text{Multiplication : } \tilde{A} \otimes \tilde{B} &= (a_1 \times b_1, a_2 \times b_2, a_3 \times b_3) \\
 \text{Division : } \tilde{A} \oslash \tilde{B} &= (a_1/b_3, a_2/b_2, a_3/b_1) \\
 \text{Scalar Division : } \tilde{A}/k &= (a_1/k, a_2/k, a_3/k) \\
 \text{Maximum : } \text{MAX}(\tilde{A}, \tilde{B}) &= (\max(a_1, b_1), \max(a_2, b_2), \max(a_3, b_3)) \\
 \text{Minimum : } \text{MIN}(\tilde{A}, \tilde{B}) &= (\min(a_1, b_1), \min(a_2, b_2), \min(a_3, b_3)) \\
 \text{Distance : } d(\tilde{A}, \tilde{B}) &= \sqrt{\frac{1}{3}[(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2]}
 \end{aligned}$$

Table 3: Adaptation of basic arithmetic operations to operands $\tilde{A} = (a_1, a_2, a_3)$ and $\tilde{B} = (b_1, b_2, b_3)$ of triangular fuzzy number type

A “large change” in the rankings appeared if both previous functions are 0:

$$lc(r_k^c, r_k^f) = \begin{cases} 1 & \text{if } top(r_k^c, r_k^f) = 0 \\ & \text{and } swap(r_k^c, r_k^f) = 0 \\ 0 & \text{otherwise} \end{cases}$$

Departing from these basic functions, we calculate the following measures for the “level of change”:

$$\begin{aligned}
 \text{None} &= \sum_{k=1}^N top(r_k^c, r_k^f) \\
 \text{Minor} &= \sum_{k=1}^N swap(r_k^c, r_k^f) \\
 \text{Large} &= \sum_{k=1}^N lc(r_k^c, r_k^f)
 \end{aligned}$$

Recall that $N = 5400$.

4 Results

The results are analyzed in global terms first, and then considering the influence that the number of alternatives, criteria and the spread level have.

4.1 Global Analysis

Table 4 shows what happened with the top alternative in both rankings. We also indicate (in the last column) the number of cases were both rankings were exactly the same.

Level of Change			Same Ranking
None	Minor	Large	
4788	455	157	782
88%	9%	3%	16%

Table 4: Level of changes in the top alternative between rankings.

In 4788 (88%) cases¹ (out of 5400) the top alternative was the same either using crisp or fuzzy information as input data. The top alternative was not the same in 455 (9%)+ 157 (3%) cases (minor and large changes). In the former one, the top two alternatives were exchanged, which means that the corresponding scores were quite similar and then, we can consider that both alternatives are indistinguishable. From another point of view, we can say that in up to 97% of the cases (those with none or minor change), the resulting top alternative is the same.

There were just 157 (3%) cases where a large change happened: the top two alternatives in one ranking were different than the top two in the other one.

If we further analyze those cases where the top alternative did not change (right pie), we observe that 782 (16%) cases (out of those 4788), both rankings were exactly the same.

4.2 On the influence of different factors

In this part, we explore if the number of alternatives, criteria or spread level, have a relevant impact in the modification of the top alternative.

¹Percentage values are rounded here and in what follows.

1) Influence of the number of alternatives m

Table 5 shows the results according to number of alternatives considered. As we analyze one factor at a time, these values are for all the criteria and spread levels.

As the number of alternatives increased, the number of minor and large changes also increased. In turn, the number of cases where the top alternative does not change diminished from 816 (90%) to 775 (86%). When the number of alternatives is greater than 10, the rankings obtained when using fuzzy or crisp data in the input were never the same.

However, the number of cases with a large change is at most 44 out of 900 (5%) in problems with 100 alternatives. In smaller problems ($m \leq 50$), the large changes never reached more than 3%.

m	N	Level of Change			Same Ranking
		None	Minor	Large	
5	900	816	75	9	582
10	900	810	69	21	200
20	900	803	75	22	0
50	900	794	76	27	0
75	900	790	79	31	0
100	900	775	81	44	0

Table 5: Influence of the number of alternatives.

Influence of the number of criteria n

Table 6 shows the results according to the number of criteria considered. As we analyze one factor at a time, these values are for all the number of alternatives and spread levels.

As the number of criteria increased, the differences in the ranking diminished. The number of cases where the top alternative remain unchanged increased from 1541 to 1622, while the cases with a large change diminished from 67 (4%) to 43 (2%). When the number of criteria is small ($n = 5$), up to 96% (1541+192) of the cases were with none or minor level of change.

n	N	Level of Change			Same Ranking
		None	Minor	Large	
5	1800	1541	192	67	260
10	1800	1625	128	47	230
15	1800	1622	135	43	292

Table 6: Influence of the number of criteria.

2) Influence of the level of spread γ

Table 7 shows the results according to the levels of spread considered. As we analyze one factor at a time, these values are for all the alternatives and criteria.

Two elements can be highlighted as spread increases: the top alternative remains the same (the value of “no changes” increased), and the number of large changes diminished. With the lowest spread ($\gamma = 10$), these large changes occurred just in 75 cases (4%) (out of 1800), diminishing up to 30 (1.6%) when $\gamma = 30$ (the highest spread). In other words, the number of none or minor changes represents almost 96% of the cases with the lowest spread and 98% with the highest one.

This situation is very reasonable: as more spread is the input data, the less information is available to distinguish the quality of the alternatives.

α	N	Level of Change			Same Ranking
		None	Minor	Large	
10	1800	1547	178	75	269
20	1800	1585	162	53	254
30	1800	1656	114	30	259

Table 7: Influence of the spread level.

5 Discussion and Conclusions

In this work we wanted to shed light into the question of whether using crisp or fuzzy information in the input of a MCDM problem produces a different output (ranking).

We conducted a simulation based experiment where we randomly generated 5400 MCDM problems with triangular fuzzy numbers in the input and their 5400 corresponding crisp variants. Then, Fuzzy-TOPSIS and the standard TOPSIS were used to solve those problems respectively.

We explored if differences arise regarding which was the top alternative in each case and propose to consider three levels of change: none, minor and large.

Over 5400 problems, large differences in the rankings produced by a crisp or a fuzzy input happened in just 157 cases; a mere 3%. As the problems had a larger number of alternatives (from 5 to 50 to 100), the number of cases with large differences increased from 1% to 3% to 5%.

As the problems increased in the number of criteria considered, the cases with large changes diminished. From 3.7% having five criteria, to 2.6% with ten, and 2.4% with fifteen criteria.

An increase in the level of spread, also reduces the number of cases with large changes. When the allowed spread is the largest (up to 15 units), the large change cases reached just 2.4%

In the basic setting of the MCDM problems we posed, our results does not support the need for the inclusion of other than crisp information in the input data.

Although we consider just a kind of fuzziness that can be modeled as symmetrical triangular fuzzy numbers, it should be carefully considered if the use of more sophisticated extensions of fuzzy sets may provide any benefit. In the simplest case of an asymmetrical triangular fuzzy number, three values should be precisely defined per each entry in the decision matrix. Beyond that, the definition of a single datum in the decision matrix using a more complex extension of fuzzy sets, will require a *greater precise knowledge* from the decision maker side which looks like a contradiction: user needs more information to define an imprecise value.

Also, it needs to be discussed if making the input more complex, while keeping the output in the same terms (a ranking, in our context) is a reasonable approach. We consider that the output should be coupled with a sort of sensitivity analysis that allows to assess the level of fuzziness in the input that can be tolerated to not produce a change in the top alternative.

These issues are under investigation.

Acknowledgement

This work is partially supported by project TIN2017-86647-P from Ministerio de Economía, Industria y Competitividad, Spain (including FEDER funds, from the European Union).

References

- [1] E. Cables, M. S. García-Cascales, M. T. Lamata, The LTOPSIS: An alternative to TOPSIS decision-making approach for linguistic variables, *Expert Systems with Applications* 39 (2) (2012) 2119–2126.
- [2] B. Ceballos, M. T. Lamata, D. A. Pelta, A comparative analysis of multi-criteria decision-making methods, *Progress in Artificial Intelligence* 5 (4) (2016) 315–322.
- [3] B. Ceballos, M. T. Lamata, D. A. Pelta, Fuzzy multicriteria decision-making methods: A comparative analysis, *International Journal of Intelligent Systems* 32 (7) (2017) 663—753.
- [4] B. Ceballos, M. T. Lamata, D. A. Pelta, R. R. Yager, Crisp vs. fuzzy data in multicriteria decision making: The case of the VIKOR method, in: *Advances in Fuzzy Logic and Technology 2017*, Springer International Pub., 2017, pp. 455–465.
- [5] C.-T. Chen, Extensions of the TOPSIS for group decision-making under fuzzy environment, *Fuzzy Sets and Systems* 114 (1) (2000) 1–9.
- [6] T.-Y. Chen, C.-Y. Tsao, The interval-valued fuzzy TOPSIS method and experimental analysis, *Fuzzy Sets and Systems* 159 (11) (2008) 1410–1428.
- [7] Z.-S. Chen, Y. Yang, X.-J. Wang, K.-S. Chin, K.-L. Tsui, Fostering linguistic decision-making under uncertainty: A proportional interval type-2 hesitant fuzzy TOPSIS approach based on Hamacher aggregation operators and andness optimization models, *Information Sciences* 500 (2019) 229–258.
- [8] The comprehensive R archive network, <https://cran.r-project.org/> (2017).
- [9] D. Dubois, H. Prade, Fuzzy numbers: An overview, in: J. Bezdek (Ed.), *Analysis of Fuzzy Information*, Vol. 2, CRC Press, 1988, pp. 3–39.
- [10] FuzzyMCDM: Multi-criteria decision making methods for fuzzy data, <https://cran.r-project.org/package=FuzzyMCDM> (2017).
- [11] F. K. Gündoğdu, C. Kahraman, A novel fuzzy topsis method using emerging interval-valued spherical fuzzy sets, *Engineering Applications of Artificial Intelligence* 85 (2019) 307 – 323.
- [12] S. Greco, M. Ehrgott, J. R. Figueira (Eds.), *Multiple Criteria Decision Analysis: State of the Art Surveys*, Vol. 233 of Intern. Series in Operations Research & Management Science, 2016.
- [13] C.-L. Hwang, K. Yoon, *Multiple attribute decision making: methods and applications : a state-of-the-art survey*, Springer-Verlag, 1981.
- [14] C. Kahraman, S. C. Onar, B. Oztaysi, Fuzzy multicriteria decision-making: A literature review, *International Journal of Computational Intelligence Systems* 8 (4) (2015) 637–666.
- [15] MCDM: Multi-criteria decision making methods for crisp data, <https://cran.r-project.org/package=MCDM> (2017).
- [16] W. Pedrycz, P. Ekel, R. Parreiras, *Fuzzy Multi-criteria Decision-Making: Models, Methods and Applications*, Wiley, 2010.
- [17] M. M. Salih, B. Zaidan, A. Zaidan, M. A. Ahmed, Survey on fuzzy TOPSIS state-of-the-art between 2007 and 2017, *Computers & Operations Research* 104 (2019) 207–227.