Big Data and Computing Visions

www.bidacv.com

Big. Data. Comp. Vis. Vol. 3, No. 4 (2023) 137–145.

Paper Type: Original Article

A New Multi-Attribute Decision-Making Method for Interval Data Using Support Vector Machine

Ghassem Farajpour Khanaposhtani*

Department of Industrial Engineering, Parand Branch, Islamic Azad University, Parand, Iran; ghfarajpour44@gmail.com.

Citation:

Farajpour Khanaposhtani, Gh. (2023). A new multi-attribute decision-making method for interval data using support vector machine. *Big data and computing visions, 3*(4), 137-145.

Received: 16/05/2023 Reviewed: 19/06/2023 Revised: 09/07/2023 Accepted: 14/08/2023

Abstract

There are numerous and various methods for solving the Multi-Attribute Decision-Making (MADM) problems in the literature, such as Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), Elimination and Choice Expressing Reality (ELECTRE), Analytic Hierarchy Process (AHP), etc. We have explored Support Vector Machine (SVM) as an efficient method for solving MADM problems. The SVM technique was proposed for classifying data at first. At the same time, in the current research, this popular method will be used to sort the preference alternatives in a MADM problem with interval data. The accuracy of the proposed technique will be compared with a popular extended method for interval data, say interval TOPSIS. Numerical experiments showed that admissible results can be obtained by the new method.

Keywords: Multi-criteria decision making, Multi-attribute decision making, Support vector machine, TOPSIS, Interval data.

1 | Introduction

Decision-making is the process of finding the best option from all of the feasible alternatives [1]. In the literature on decision-making, many scholars and researchers proposed their model [2]–[6]. In almost such problems, the multiplicity of attributes for judging the alternatives is pervasive. These attributes usually conflict, so there may be no solution satisfying all attributes simultaneously. Therefore, a Multi-Attribute Decision-Making (MADM) model should identify the best alternative by considering all criteria. The general format of the MADM problem is illustrated in *Table 1*, in which there are m possible alternatives $A_1, A_2, ..., A_m$ among which decision-makers have to choose and n criteria (also say attributes) $C_1, C_2, ..., C_n$. Also, x_{ij} is the rate of alternative i with respect to criterion j, and w_j is the weight of criterion j.

There are various methods for solving MADM problems [7]. Some of the traditional methods are designed and developed for crisp data cases. However, in real problems, there may be other types, such as fuzzy, ordinal, and interval data. In other words, the decision maker would prefer to present

Corresponding Author: ghfarajpour44@gmail.com

https://doi.org/10.22105/bdcv.2023.190406

Farajpour Khanaposhtani | Big. Data. Comp. Vis. 3(4) (2023) 137-145

their judgment in these forms rather than a crisp number because of the uncertainty and the lack of certain data. Therefore, it seems that other cases are essential to be developed for decision-making [8]–[10].

One of the special and interesting cases in MADM problems that has received much attention is that the decision matrix elements are interval numbers. There are many contributions in the mentioned area for solving these types of problems [11]–[13]. Also, it is admissible that our proposed method is considered an interval decision-making method. This paper will concentrate on interval data, which is more useful for real-life problems.

In the next sections, a popular method for solving interval MADM problems, interval Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), will be introduced. Then, a description of the SVM classifier will be given. The procedure of using SVM to solve interval MADM problems will be discussed in the section on the research method. The results section contains numerical results and comparisons; finally, this research is finished with concluding remarks.

We have explored the Support Vector Machine (SVM) as an efficient and flexible distance-based technique for use in this area. Therefore, the base contribution of the current research is to use SVM for solving MADM problems with interval data by ranking the alternatives according to a decision function derived from the SVM model. Numerical experiments and comparisons investigate the performance of the proposed technique. The next sections of the paper are organized as follows:

2 | Interval TOPSIS

One of the popular methods in MADM problems is TOPSIS (technique for order preference by similarity to an ideal solution), presented by Hwang and Yoon [14]. This method investigates alternatives according to their distance from ideal and negative-ideal, which, as an extended procedure, Olson [15] used the weights and some other norms to measure these distances. Many articles have extended the TOPSIS method in terms of the theory and application. For example, Lai et al. [16] applied the concept of TOPSIS to Multi-Objective Decision Making (MODM) problems. Abo-Sinna and Amer [17] extended the TOPSIS method for solving multi-objective large-scale non-linear programming problems. Also, Kuo et al. [18] and Shis et al. [19] have extended TOPSIS for group decision-making, and after some years, their model was extended by Yue [20] for the case of interval data. Many other novelties can be found in this area. One can follow the research of Behzadian et al. [21] and Salih et al. [22] for more details. Jahanshahloo et al. [1] proposed an algorithmic scheme for the TOPSIS method with interval data.

Table 1. MADM scheme.

Criteria / Alternative	C_1	C_2	…	C_n
A_1	X_{11}	X_{12}	…	X_{1n}
A_2	X_{21}	X_{22}	…	X_{2n}
⋮	⋮	⋮	⋱	⋮
A_m	X_{m1}	X_{m2}	…	X_{mn}

In the case of interval data, the decision matrix may be modified as *Table 2*, in which each alternative has an upper and lower rate concerning any criterion. The procedure of interval TOPSIS proposed by Jahanshahloo et al. [1] will be used in our research to compare it to SVM.

Table 2. Multi-attribute decision matrix for interval data.

Alternative \ Criteria	C_1	C_2	…	C_n
A_1	$[x_{11}^l, x_{11}^u]$	$[x_{12}^l, x_{12}^u]$	…	$[x_{1n}^l, x_{1n}^u]$
A_2	$[x_{21}^l, x_{21}^u]$	$[x_{22}^l, x_{22}^u]$	…	$[x_{2n}^l, x_{2n}^u]$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
A_m	$[x_{m1}^l, x_{m1}^u]$	$[x_{m2}^l, x_{m2}^u]$	…	$[x_{mn}^l, x_{mn}^u]$

The procedure of interval TOPSIS method can be outlined as follows:

I. Calculating the lower and upper normalized values and the elements of the normalized matrix.
II. Constructing the lower and upper ideal alternatives as well as the negative alternatives.
III. Determine the distance measure for investigating the alternatives. Note that in the case of interval data, the distance measure consists of lower and upper distances. Regarding this fact, all ranks will be in the form of intervals, leading to a comparison using interval data.

By considering m and h as the mid-point and half-width of interval data, two approaches exist for comparing the interval data:

Senguta`s approach [23]: for comparing two interval data E and D*,* calculate the below index:

$$A = \frac{m_D - m_E}{h_D + h_E}. \tag{1}$$

Then, for benefit interval, it can be concluded that if:

I. $A > 0 \rightarrow$ The interval D is inferior to the interval E.
II. $A < 0 \rightarrow$ The interval D is superior to the interval E.
III. $A = 0 \rightarrow$ The lower h is more acceptable.

The above statements can be generalized for cost intervals.

Delgado`s approach [24]: interval data can be represented as a trapezoidal fuzzy number by considering the value V and ambiguous A as the mid-point and half-width, respectively. So, the next two steps should be followed to determine the excellent one:

Step 1. Compare V_E and V_D. If they are approximately equal, then go to Step 2. Otherwise, rank the data concerning their value.

Step 2. If A_E and A_D are approximately equal, conclude that the interval data are not different. Otherwise, the data with less A may be preferred by an optimistic decision maker and the other by a pessimistic viewpoint.

In this paper, two approaches will be used to rank the alternatives.

3 | Support Vector Machine

SVM is a machine learning technique introduced by Cortes and Vapnik [25] for classifying data. Classification is a task for assigning predefined labels to each data. SVM tries to find a separator hyperplane with the largest margin for classifying data into two groups. A graphical example of a SVM is illustrated in

Fig. 1 in the case of linearly separable data; however, there may be some situations in which data are linearly nonseparable.

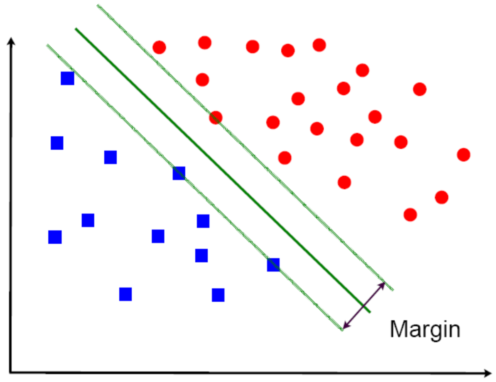

Fig. 1. separating hyperplane with maximum margin.

The nearest points to the separator hyperplane are support vectors. For constructing the model of SVM in the case of linearly separable, assume that $X=\{x_1,x_2,...,x_n\}$ are training datasets and $y_i\in\{-1,1\}$ are class variables refer to the label of each training data. Therefore, by considering the above notations, the optimization model of the linearly separable case SVM will be as follows:

$$\begin{aligned} \min \quad & \frac{1}{2}\|\omega\|^2, \\ & y_i\left(\omega^T x_i + b\right) \geq 1, \\ & y_i \in \{-1,1\}, \end{aligned} \tag{2}$$

where ω and b are decision variables (normal vector and bias of the separator hyperplane, respectively).

The SVM model can be modified easily for use in the case of linearly nonseparable. *Fig. 2* shows an example of a nonseparable case. In the mentioned situation, some of the constraints will be unfeasible because of the nonseparable data and should be adjusted by introducing variables ξ_i, say error variables, in the literature. This modification leads to the following model [25]:

$$\begin{aligned} \min \quad & \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}\xi_i, \\ & y_i\left(\omega^T x_i + b\right) \geq 1-\xi_i, \\ & y_i \in \{-1,1\}, \\ & \xi_i \geq 0, \end{aligned} \tag{3}$$

where C is the misclassification cost and ξ_i determines the misclassified data and can be interpreted as follows:

I. $\xi_i = 0 \to$ i[th] data is classified correctly (out of margin).

II. $0 \leq \xi_i \leq 1 \to$ i[th] data is classified correctly (in margin).

III. $\xi_i > 1 \to$ i[th] data is misclassified.

The currently mentioned method can be used to rank the alternatives in a MADM problem. The next section is dedicated to describing this new procedure.

4 | Research Method

A separator hyperplane can be used for ranking the alternatives in a MADM problem. This hyperplane may be found by using a SVM.

Farajpour Khanaposhtani | Big. Data. Comp. Vis. 3(4) (2023) 137-145

Assume that $X=\left[x_{ij}^l,x_{ij}^u\right]$ $\left(i=1,...,m,\ j=1,...,n\right)$ is an interval decision matrix contains m alternatives with n criteria. The following procedure should be followed to find the best alternative(s):

Step 1. Construct the weighted decision matrix as $\upsilon_{ij}^l=w_ix_{ij}^l$ and $\upsilon_{ij}^u=w_ix_{ij}^u$. Where υ_{ij}^l and υ_{ij}^u are the lower and upper elements, respectively.

Step 2. Set all the alternatives in their best situation and define the upper ideal alternative as follows:

$$A_k^{+u}=\left\{\left(\upsilon_1^{+u},\upsilon_2^{+u},\ldots,\upsilon_n^{+u}\right)\right\}=\left\{\left(\max\upsilon_{ij}^u\mid i\in B\right),\left(\min\upsilon_{ij}^l\mid i\in C\right)\right\}, \tag{4}$$

where according to the previous sections, B and C are the subsets of benefit and cost criterion, respectively.

Step 3. The lower negative ideal alternative will be calculated after setting all of the alternatives in their worst situation:

$$A_k^{-l}=\left\{\left(\upsilon_1^{-l},\upsilon_2^{-l},\ldots,\upsilon_n^{-l}\right)\right\}=\left\{\left(\min_{j\neq k}\left(\upsilon_{ij}^l\right)\mid i\in B\right),\left(\max_{j\neq k}\left(\upsilon_{ij}^u\right)\mid i\in C\right)\right\}. \tag{5}$$

After calculating both ideal and negative alternatives, a classification problem can be described. In other words, a separator hyperplane can be found using SVM, and the mentioned plane will be used as a decision function that ranks alternatives according to their scores. *Table 3* shows the classification problem to be solved by SVM.

Table 3. Conversion of MADM problem to a classification problem.

Criterion / Data	1	2	…	n	Class Label
1	υ_1^{+u}	υ_2^{+u}	$\cdots$	υ_n^{+u}	1
2	υ_1^{-l}	υ_2^{-l}	$\cdots$	υ_n^{-l}	-1

Step 4. The optimization model derived from *Table 3* will be as follows:

$$\begin{aligned}\min\quad & \frac{1}{2}\left(\omega_1^2+\omega_2^2+\cdots+\omega_n^2\right)+C\left(\xi_1+\xi_2\right)\\ & +1\left(\omega_1\upsilon_1^{+u}+\omega_2\upsilon_2^{+u}+\cdots+\omega_n\upsilon_n^{+u}+b\right)\geq 1-\xi_1\\ & -1\left(\omega_1\upsilon_1^{-l}+\omega_2\upsilon_2^{-l}+\cdots+\omega_n\upsilon_n^{-l}+b\right)\geq 1-\xi_2\\ & \xi_i\geq 0\ .\end{aligned} \tag{6}$$

Lemma 1. *Model (6)* is a linearly separable case that will be feasible without using ξ_1,ξ_2.

Proof: it is sufficient to prove that a feasible solution can be found for this model with $\xi_1=\xi_2=0$. So, assume that these two variables are equal to 0. In this situation, if two constraints of the model are summed together, the below nonequality will be obtained:

$$\sum_{i=1}^{n}\omega_i\left(\upsilon_i^{+u}-\upsilon_i^{-l}\right)\geq 1.$$

Note that the term $\upsilon_i^{+u}-\upsilon_i^{-l}$ is positive for all i. So, many feasible solutions can be introduced for *Model (6)* by considering the previous result. For example, by substituting the solution $\omega_i=\dfrac{1}{\upsilon_i^+-\upsilon_i^-}$, the above nonequality is held, and it can be concluded that the introduced solution is feasible.

Considering this argument, ξ_1, ξ_2 can be eliminated from *Model (6)*, which leads us to the below model:

$$\begin{aligned} \min \quad & \frac{1}{2}\left(\omega_1^2+\omega_2^2+\cdots+\omega_n^2\right), \\ & +1\left(\omega_1\upsilon_1^{+u}+\omega_2\upsilon_2^{+u}+\cdots+\omega_n\upsilon_n^{+u}+b\right)\geq 1, \\ & -1\left(\omega_1\upsilon_1^{-l}+\omega_2\upsilon_2^{-l}+\cdots+\omega_n\upsilon_n^{-l}+b\right)\geq 1. \end{aligned} \tag{7}$$

The above problem is a quadratic optimization model which many efficient algorithms can solve.

Step 5. Define the upper and lower distance for alternative i as follows:

$$f_u(\omega,b)=\sum_{j=1}^{n}\left(\omega_j\upsilon_{ij}^{u}\right)+b. \tag{8}$$

$$f_l(\omega,b)=\sum_{j=1}^{n}\left(\omega_j\upsilon_{ij}^{l}\right)+b. \tag{9}$$

Step 6. When the separator hyperplane is found, the decision function $f_A(\omega,b)$ will be used for ranking the alternatives as follows:

$$f_A(\omega,b)=\frac{f_l(\omega,b)+f_u(\omega,b)}{2}, \tag{10}$$

where the greater value of the decision function refers to a better rank.

The introduced index for ranking the alternatives is the average of the distances from the separator hyperplane for each alternative. If this value for alternative *i* is greater than the others; it can be concluded that the alternative i is more preferable. $f_A(\omega,b)$ can be rewritten as follows:

$$f(\omega,b)=\frac{\sum_{j=1}^{n}\left[\omega_j\left(\upsilon_{ij}^{l}+\upsilon_{ij}^{u}\right)\right]+2b}{2}, \tag{11}$$

5 | Results and Discussion

The numerical data used in this section may be divided into two groups. The first experiment will be run on the data used in the work of Jahanshahloo et al. [1]. After this experiment, the proposed method will be compared to interval TOPSIS using 11 randomly generated datasets. Note that the criteria weights are assumed to be 1, then $w_j=1$ for all j.

The first dataset is a four-criteria problem containing six cities as alternatives for constructing a date factory. The first two criteria are cost-oriented, while the others are benefit (*Table 4*).

Table 4. Datasets for comparing the introduced methods (interval TOPSIS and SVM).

City	Distance from Border	Cost of Construction	Finance	Products in the Region
City 1	1451	[2551, 3118]	[40, 50]	[153, 187]
City 2	843	[3742, 4573]	[63, 77]	[459, 561]
City 3	1125	[3312, 4049]	[48, 58]	[153, 187]
City 4	55	[5309, 6488]	[72, 88]	[347, 426]
City 5	356	[3709, 4534]	[59, 71]	[151, 189]
City 6	391	[4884, 5969]	[72, 88]	[388, 474]

Farajpour Khanaposhtani | Big. Data. Comp. Vis. 3(4) (2023) 137-145

The SVM model for solving this problem contains two constraints by using the upper ideal alternative and lower negative alternatives as follows:

$$\begin{aligned}&\min \frac{1}{2}\left(\omega_1^2+\omega_2^2+\omega_3^2+\omega_4^2\right)\\&+1\left(55\omega_1+2551\omega_2+88\omega_3+474\omega_4+b\right)\geq 1\\&-1\left(1451\omega_1+6488\omega_2+40\omega_3+151\omega_4+b\right)\geq 1.\end{aligned} \tag{11}$$

A separator hyperplane will be found by solving the above model, which may be used as a decision function for ranking the alternatives. The results of the comparisons are outlined in *Table 5*.

Table 5. Comparing SVM and interval TOPSIS.

Alternatives	Ranking	
	Interval TOPSIS [1]	Proposed Method (Interval SVM)
City 1	6	6
City 2	4	5
City 3	5	4
City 4	1	1
City 5	3	2
City 6	2	3

11 randomly generated datasets are used for more investigations, and the results are reported in *Table 6*. It is obvious that the rankings are very similar, indicating that the proposed method is a valid technique. Also, the new method, say SVM, outperforms the interval TOPSIS in terms of the consumed time.

Table 6. SVM versus interval TOPSIS- Random generated datasets.

Dataset	Size		Sorted Alternatives (Ascending)		Time (Seconds)	
	Alternatives	Criterion	Interval TOPSIS	SVM	Interval TOPSIS	SVM
1	5	5	1-5-2-3-4	1-3-2-5-4	0.002	0.007
2	5	20	3-4-1-5-2	3-4-1-2-5	0.002	0.0077
3	10	5	6-8-1-4-5-10-2-9-3-7	6-1-8-5-4-10-9-3-2-7	0.020	0.012
4	10	10	8-2-1-4-7-9-3-10-6-5	2-8-1-9-4-7-10-3-6-5	0.020	0.012
5	20	25	---	---	0.021	0.015
6	30	40	---	---	0.025	0.019
7	50	20	---	---	0.027	0.019
8	100	15	---	---	0.038	0.021
9	500	20	---	---	0.578	0.029
10	1000	30	---	---	3.69	0.039
11	1000	200	---	---	4..4	0.273

According to *Table 6*, it is obvious that the problems can be analyzed in a very short time rather than the interval TOPSIS, especially in the case of large scales. Also, adding more alternatives or criteria to the dataset will increase the time gap between the two methods.

6 | Conclusions

This paper introduced a new SVM method for ranking the alternatives in a multi-criteria decision-making problem. It was shown that a decision function may be found by using this method, which ranks the alternatives using the concept of distance from the ideal alternative. Finally, the validity and time superiority of the proposed technique were confirmed by computational experiments on both real and random generated datasets.

Availability of Data and Materials

The datasets generated and/or analyzed during the current study are not publicly available. However, the data will be available from the corresponding author upon reasonable request.

Competing Interests

The authors declare that they have no competing interests.

Funding

This study was self-financed by the authors. That is, no form of assistance was received from anybody or institution.

Authors' Contributions

The author read and approved the final manuscript.

Acknowledgments

The authors appreciate the effort made by all opinion leaders in the study communities who, out of their busy schedules, made this research work successful by paving the way for it to be carried out without a hitch.

References

[1] Jahanshahloo, G. R., Lotfi, F. H., & Davoodi, A. R. (2009). Extension of TOPSIS for decision-making problems with interval data: Interval efficiency. *Mathematical and computer modelling*, *49*(5–6), 1137–1142.

[2] Rasinojehdehi, R., & Valami, H. B. (2023). A comprehensive neutrosophic model for evaluating the efficiency of airlines based on SBM model of network DEA. *Decision making: applications in management and engineering*, *6*(2), 880–906.

[3] Nozari, H., Najafi, E., Fallah, M., & Hosseinzadeh Lotfi, F. (2019). Quantitative analysis of key performance indicators of green supply chain in FMCG industries using non-linear fuzzy method. *Mathematics*, *7*(11), 1020. https://doi.org/10.3390/math7111020

[4] Najafi, S. E., Amiri, M., & Abdolah Zadeh, V. (2015). Ranking practicable technologies within the science and technology corridor of Isfahan, using MCDM techniques. *Journal of applied research on industrial engineering*, *2*(3), 139–153.

[5] Rasi Nojehdehi, R., Bagherzadeh Valami, H., & Najafi, S. E. (2023). Classifications of linking activities based on their inefficiencies in Network DEA. *International journal of research in industrial engineering*, *12*(2), 165–176.

[6] Seiti, H., Hafezalkotob, A., Najafi, S. E., & Khalaj, M. (2019). Developing a novel risk-based MCDM approach based on D numbers and fuzzy information axiom and its applications in preventive maintenance planning. *Applied soft computing*, *82*, 105559. https://doi.org/10.1016/j.asoc.2019.105559

[7] Zeleny, M. (1982). *Multiple criteria decision making*. McGraw Hill.

[8] Chen, S. J., & Hwang, C. L. (1992). Fuzzy multiple attribute decision making methods. In *Fuzzy multiple attribute decision making: methods and applications* (pp. 289–486). Springer.

[9] Chen, T.-Y., & Tsao, C.-Y. (2008). The interval-valued fuzzy TOPSIS method and experimental analysis. *Fuzzy sets and systems*, *159*(11), 1410–1428.

[10] Wang, Y. M., & Elhag, T. M. S. (2006). Fuzzy TOPSIS method based on alpha level sets with an application to bridge risk assessment. *Expert systems with applications*, *31*(2), 309–319.

[11] Sayadi, M. K., Heydari, M., & Shahanaghi, K. (2009). Extension of VIKOR method for decision making problem with interval numbers. *Applied mathematical modelling*, *33*(5), 2257–2262.

[12] Xue-jun, T., & Jia, C. (2012). A dynamic interval decision-making method based on GRA. *Physics procedia*, *24*, 2017–2025.

[13] Khefacha, I., & Belkacem, L. (2015). Modeling entrepreneurial decision-making process using concepts from fuzzy set theory. *Journal of global entrepreneurship research*, *5*, 1–21.

[14] Hwang, C. L., Yoon, K., Hwang, C. L., & Yoon, K. (1981). Methods for multiple attribute decision making. *Multiple attribute decision making: methods and applications a state-of-the-art survey*, *186*, 58–191. https://link.springer.com/chapter/10.1007/978-3-642-48318-9_3

[15] Olson, D. L. (2004). Comparison of weights in TOPSIS models. *Mathematical and computer modelling*, *40*(7–8), 721–727.

[16] Lai, Y.-J., Liu, T.-Y., & Hwang, C.-L. (1994). Topsis for MODM. *European journal of operational research*, *76*(3), 486–500.

[17] Abo-Sinna, M. A., & Amer, A. H. (2005). Extensions of TOPSIS for multi-objective large-scale nonlinear programming problems. *Applied mathematics and computation*, *162*(1), 243–256.

[18] Kuo, M. S., Tzeng, G. H., & Huang, W. C. (2007). Group decision-making based on concepts of ideal and anti-ideal points in a fuzzy environment. *Mathematical and computer modelling*, *45*(3–4), 324–339.

[19] Shih, H.-S., Shyur, H.-J., & Lee, E. S. (2007). An extension of TOPSIS for group decision making. *Mathematical and computer modelling*, *45*(7–8), 801–813.

[20] Yue, Z. (2013). Group decision making with multi-attribute interval data. *Information fusion*, *14*(4), 551–561.

[21] Behzadian, M., Otaghsara, S. K., Yazdani, M., & Ignatius, J. (2012). A state-of the-art survey of TOPSIS applications. *Expert systems with applications*, *39*(17), 13051–13069.

[22] Salih, M. M., Zaidan, B. B., Zaidan, A. A., & Ahmed, M. A. (2019). Survey on fuzzy TOPSIS state-of-the-art between 2007 and 2017. *Computers & operations research*, *104*, 207–227.

[23] Sengupta, A., & Pal, T. K. (2000). On comparing interval numbers. *European journal of operational research*, *127*(1), 28–43.

[24] Delgado, M., Vila, M. A., & Voxman, W. (1998). A fuzziness measure for fuzzy numbers: Applications. *Fuzzy sets and systems*, *94*(2), 205–216.

[25] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*, 273–297.