Assessing Normalization Techniques for TOPSIS Method

Nazanin Vafaei(✉), Rita A. Ribeiro, and Luis M. Camarinha-Matos

CTS-UNINOVA and School of Science and Technology, NOVA University of Lisbon, 2829-516 Caparica, Portugal
{Nazanin.vafaei,rar,cam}@uninova.pt

Abstract. In recent years, data normalization is receiving considerable attention due to its essential role in decision problems. Especially, considering the new developments in Big data and Artificial Intelligent to handle heterogeneous data from sensors, normalization's role as a preprocessing step for complex decision problems is more distinguished. However, selecting the best normalization technique among several introduced techniques in the literature is still an open issue. In this study we focus on evaluating normalization techniques in Multi-Criteria Decision Making (MCDM) methods namely for Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) to recommend the most proper technique. A small numerical example, borrowed from literature, is used to show the applicability of the proposed assessment framework using several metrics for recommending the most suitable technique. This study helps decision makers to improve the accuracy of the final ranking of results in decision problems by selecting the best normalization technique for the related case study.

Keywords: Normalization · MCDM · TOPSIS · Decision making · Data fusion · Aggregation · Big data · Artificial Intelligence

1 Introduction

Nowadays, selecting the optimal solution/alternative based on multiple criteria for a given decision problem is a major task for decision makers due to the availability of criteria with different scales/ranges. Multi-Criteria Decision Making (MCDM) is used by most decision makers to deal with the decision problems. They carry out each MCDM method by defining a decision matrix that consists of a set of alternatives Ai (i = 1, …, m), criteria Cj (j = 1,…, n), the relative importance of the criteria (or weights) Wj, and rij, corresponding the rating of alternative i with respect to criteria j [1]. Most of the criteria are measured on different scales (e.g., weight, height, temperature, etc.) while they need to be defined in the same scale/range to enable decision makers to make a valid comparison and selection of the best solution. In other words, decision makers need some preprocessing to produce comparable and dimensionless data from heterogeneous input data sets that is called normalization process. Generally, MCDM methods have four main steps as [2]:

Published by Springer Nature Switzerland AG 2021
L. M. Camarinha-Matos et al. (Eds.): DoCEIS 2021, IFIP AICT 626, pp. 132–141, 2021.
https://doi.org/10.1007/978-3-030-78288-7_13

1) Determine the decision matrix
2) Normalize the criteria values in the decision matrix
3) Calculate weighted normalized values
4) Aggregation process (which differs for each MCDM method)
5) Sort alternatives in a decreasing ordering.

There are some research papers about the importance of normalization techniques in MCDM methods [3–9]. They discuss the fact that different normalization techniques may address different ranking of alternatives. So, selecting an improper normalization technique may cause deviation from the original solution. Therefore, normalization has an essential role in most MCDM methods and using the most proper technique will help decision makers to improve the accuracy of the final solution.

Furthermore, new developments in big data and Artificial Intelligence (AI) and the integration with MCDM methods provide novel aspects for handling collected data from heterogeneous sensors. These big data from sensors need a normalization process to produce dimensionless input data sets to be used in decision problems. So, normalization techniques have important role in these new points of view of science.

Several normalization techniques are introduced in the literature along with different MCDM methods [1]. In this study, we analyze the effects of different normalization techniques on the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) method (to get more details about TOPSIS please see [7]). Thus, the main research question related to this study is: *Which normalization technique is the most appropriate for usage with the TOPSIS method?*

This paper shows the applicability of the assessment framework proposed by [8] and compares the suitability of different normalization techniques using the TOPSIS method. In order to test the mentioned framework, a case study was borrowed from [10], which analyzed the effects of five normalization techniques as Max, Max-Min, Sum, Vector, and Fuzzification (Gaussian) with the TOPSIS method. The obtained results are compared with the results of initial work from [10].

2 Contribution to Applied Artificial Intelligence Systems

In recent years, with the advent of big data, organizations deal with complex decision problems consisting of large amounts of data. So, the integration of MCDM and Artificial Intelligence (AI) techniques is considered and reached success in handling real-world problems [11]. This integration enables decision makers to better structure complex decision problems in static and distributed environments. Doumpos and Grigoroudisby [11] mention that these help as "*handling of massive data sets, the modelling of ill-structured information, the construction of advanced decision models, and the development of efficient computational optimization algorithms for problem solving*".

Several AI techniques are introduced in the literature which are used with integration MCDM, such as Fuzzy Logic (FL), Genetic Algorithm (GA), Neural Network (NN), Heuristic or meta-heuristics, Knowledge-Based (KB), Expert Systems (ES), Tabu-Search (TS), Simulated-annealing (SA), Dampster Shafer (DS), and Self-Organizing-Map (SOM). However, FL is the most popular technique to be used with MCDM methods (e.g., Fuzzy-AHP, Fuzzy-ANP, and Fuzzy-TOPSIS, etc.) [12].

There are some researches on implementing AI techniques with MCDM methods such as Ho [13] that uses 8 meta-heuristics along with AHP, Pan [14] applies Fuzzy-AHP for bridge construction methods selection, Sheu [15] uses Fuzzy-AHP, Fuzzy-TOPSIS, and Fuzzy-MCDM for global logistic operational model, Kulturel-Konak et al. [16] apply TS for system redundancy allocation problem, Efendigil et al. [17] implement ANN and Fuzzy-AHP for third-party logistics providers selection, and Wu et al. [18] use Fuzzy-ANP for site selection problem.

3 Assessment Framework for Evaluation of Normalization Techniques

As mentioned above, normalization is a vital step of most MCDM decision problems that "transfers" all criteria to the same scale and provides an effective comparison between alternatives to select the optimal solution. Numerous normalization techniques have been introduced in the literature. For instance, Jahan and Edwards [1] listed 31 techniques and addressed the pros and cons of them and pointed the influence effects of normalization on the results of MCDM methods in material selection decision problems. The authors of [1] also discuss the different features that should be considered to evaluate a normalization technique namely ranking reversal, symmetry, handling negative values, and capability in removing scales. Furthermore, Vafaei et al. [8] added data type and topology to the initial features from [1] as an important aspect for considering in the evaluation process.

Among several introduced normalization techniques in the literature [1], some of them are more well-known and implemented for the specific mission or MCDM methods. For example, Vector normalization is often applied with TOPSIS method, Target base normalization technique is utilized for material selection decision problems, and Sum is used with AHP (Analytical Hierarchy Process) method [1]. However, there are no clear reasons in the literature for these specific usages, which motivated us to evaluate different normalization techniques in MCDM decision problems.

Some research papers investigated the evaluation of the effects of normalization techniques in MCDM decision problems using various metrics. For instance, Celen [3] analyzed the effect of four normalization techniques (Max, Max-Min, Sum, and Vector) in TOPSIS method using consistency conditions metrics and showed that Vector is the most appropriate normalization technique for the used case study [3]. Also, Charaborty and Yeh [4, 5] used Ranking Consistency Index (RCI) for evaluating the same normalization techniques (Max, Max-Min, Sum, and Vector) for TOPSIS [4] and SAW [5] methods. In another study, Lakshmi and Venkatesan [5] assessed five normalization techniques namely Max, Max-Min, Sum, Vector, Fuzzification (Gaussian membership function) with TOPSIS method using time and space complexity in MATLAB software and selected Sum normalization technique as the best for the case study. In addition, Mathew et al. [6] implemented Spearman correlation to evaluate 6 normalization techniques for the weighted aggregated sum product assessment (WASPAS) method and suggested Max-Min technique as the most suitable normalization for the case study. Recently, Vafaei et al. [7–9, 19, 20] developed an assessment framework to evaluate different normalization techniques for MCDM methods. The proposed assessment framework is consisted of several metrics that are presented in Fig. 1.

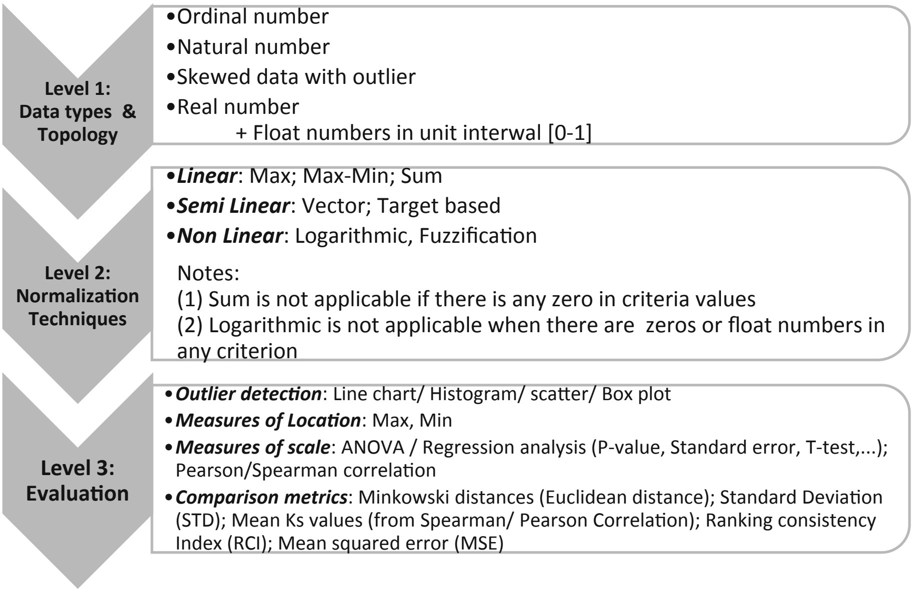

Fig. 1. Three level of the evaluation framework (adopted from [8]).

In this study, we show the applicability of the evolved framework that is proposed by [8]. Figure 1 shows that the first level of this framework is related to distinguishing the data type and topology of the input data sets. Proceeding to the second level enables decision makers to select the normalization techniques. Following, the third level that is the main part of the framework, includes several metrics to assess the used normalization techniques in the decision problems and recommend the most proper normalization techniques for the related case study (for more detail about the assessment framework please see [8]).

In the next section, we show the usage of this assessment framework using a case study that is borrowed from [10]. The related case study from [10] assessed five normalization techniques i.e. Max, Max-Min, Sum, Vector, Fuzzification (Gaussian membership function). So, in order to provide the possibility of benchmarking between our results and the initial results from the author we use the same normalization techniques. The formulas of the mentioned normalization techniques are presented in Table 1. It is noticeable that benefit formulas refer to the higher values the better and cost formulas represent the lower value is desirable.

Table 1. Well-known normalization techniques (adapted from [10]).

Normalization technique	Condition of use	Formula
Linear: Max	Benefit criteria	$n_{ij} = \frac{r_{ij}}{r_{max}}$
	Cost criteria	$n_{ij} = 1 - \frac{r_{ij}}{r_{max}}$
Linear: Max-Min	Benefit criteria	$n_{ij} = \frac{r_{ij} - r_{min}}{r_{max} - r_{min}}$
	Cost criteria	$n_{ij} = \frac{r_{max} - r_{ij}}{r_{max} - r_{min}}$
Linear: Sum	Benefit criteria	$n_{ij} = \frac{r_{ij}}{\sum_{i=1}^{m} r_{ij}}$
	Cost criteria	$n_{ij} = \frac{1/r_{ij}}{\sum_{i=1}^{m} 1/r_{ij}}$
Semi-Linear: Vector	Benefit criteria	$n_{ij} = \frac{r_{ij}}{\sqrt{\sum_{i=1}^{m} r_{ij}^2}}$
	Cost criteria	$n_{ij} = 1 - \frac{r_{ij}}{\sqrt{\sum_{i=1}^{m} r_{ij}^2}}$
Non- Linear: Fuzzification (membership function)	Benefit & cost criteria	E.g. trapezoidal: $f(x, a, b, c, d) = \begin{cases} 0 & x \le a \\ \frac{x-a}{b-a} & a \le x \le b \\ \frac{d-x}{d-c} & c \le x \le d \\ 0 & d \le x \end{cases}$

4 Comparison of Normalization Techniques with an Illustrative Example for TOPSIS Method

To test the applicability of the proposed assessment framework, we applied metrics from the framework to the small illustrative example borrowed from [10]. This case study analyzed and recommended normalization techniques for TOPSIS method in a car selection problem. It consists of 4 alternatives (A1, …, A4) regarding different car brands (Civic Coupe, Saturn Coupe, Ford Escort, and Mazda Miata) and 4 criteria (C1, …, C4) related to the car characteristics (style, reliability, fuel-eco, and cost). The decision matrix with the input data is depicted in Table 2.

Table 2. Decision matrix input data (borrowed from [10])

	C1	C2	C3	C4
A1	7	9	9	8
A2	8	7	8	7
A3	9	6	8	9
A4	6	7	8	6

The authors of [10] used five normalization techniques (Max, Max-Min, Sum, Vector, and Fuzzification (using Gaussian membership function)) to analyze the effects of using different normalization techniques on the TOPSIS method. They calculated the relative closeness and ranking of alternatives, with TOPSIS, using the selected normalization techniques for this case study (Table 3). For more details about the TOPSIS method please see [7].

Table 3. Relative closeness (RC) and Ranking of alternatives (R) the case study from [10]

	Vector		Max-Min		Sum		Max		Fuzzification (Gaussian)	
	RC	R	RC	R	RC	R	RC	R	RC	R
A1	0.74	1	0.88	1	0.38	1	0.26	4	0.75	1
A2	0.41	3	0.28	4	0.26	3	0.31	1	0.45	3
A3	0.17	4	0.31	3	0.009	4	0.29	3	0.02	4
A4	0.44	2	0.45	2	0.32	2	0.30	2	0.62	2

As Table 3 depicts, there is no consensus between the five selected normalization techniques for ranking alternatives. The work described in [10] calculated time complexity and space complexity for each normalization technique with the help of MATLAB and recommended the Sum normalization technique as the best one for the case study, using the TOPSIS method.

Now, we apply the proposed assessment framework and recommend the most appropriate normalization technique for the related case study and compare our results with the obtained results by the authors of [10].

In the first level of the assessment framework the topology of input data is considered. However, in this case study the decision matrix contains integer numbers between [6–9] without any outliers or decimal numbers, and zero. So, we skip the first level and proceed to the second level for selecting the normalization techniques to be assessed.

As mentioned before, the authors of [10] analyzed the effect of five normalization techniques (Max, Max-Min, Sum, Vector, and Fuzzification (using Gaussian membership function)) i.e. at least one normalization technique from three categories (Linear, Semi-linear, and Non-linear) that are mentioned in the second level of the assessment framework (Fig. 1) are selected.

The third level of the proposed assessment framework consists of several metrics from four classes (outlier detection; measures of location; measures of scale; and comparison methods). As mentioned above, the decision matrix includes integer numbers in the interval [6–9] without any outliers, decimal numbers and zero. Therefore, metrics for determining outliers are useless and we omitted these class of metrics for this case study. For the rest of the metrics, we focused on the comparison metrics namely Minkowski distances (Euclidean distance), Standard Deviation (STD), Mean Ks values (from Spearman/Pearson Correlation), Ranking consistency Index (RCI), and Mean squared error (MSE). For details about calculating Euclidean distance and STD please see [19] and for calculation of Mean Ks value and RCI please see [7].

For calculating MSE for aggregated data sets, it is obtained from the average of mean squared error for each normalization technique with other normalization techniques using the ranking of alternatives [21, 22]. It should be noticed that the favorable result for MSE is the lower value the better because it is desirable to have less error when comparing different normalization techniques [8]. For the rest of the metrics (Euclidean distance, STD, RCI, and Mean Ks) the higher values are better and more desirable [8]. The results of MSE are depicted in Table 4.

Table 4. MSE for the borrowed case study from [10] for TOPSIS method

	Vector	Max-Min	Sum	Max	Fuzzification (Gaussian)	MSE	Rank
Vector	–	0.5	0	3.5	0	1	2
Max-Min	0.5	–	0.5	0.5	0.5	0.5	1
Sum	0	0.5	–	3.5	0	1	2
Max	3.5	0.5	3.5	–	3.5	2.75	5
Fuzzification (Gaussian)	0	0.5	0	3.5	–	1	2

Implementing different metrics from the third level of the proposed assessment framework as Euclidean distance, STD, RCI, and Mean Ks plus the results of MSE are provided comparable results (Table 5).

Table 5. Results of different metrics of assessment framework using TOPSIS method for the borrowed case study from [10].

	Euclidean	STD	RCI	MSE	Mean Ks
Max	0.8094	0.2337	10.75	1	0.5211
Max-Min	0.9587	0.2768	9	0.5	0.3364
Sum	0.5648	0.1630	10.75	1	0.5690
Vector	0.0748	0.0216	6.25	2.75	−0.5452
Fuzzification (Gaussian)	1.1016	0.3180	10.75	1	0.5694

Table 6 shows the ordering of normalization techniques with respect to each used metrics. The different metrics produced different ranks for the chosen techniques and still recommending the most appropriate normalization technique is impossible just by looking at the results. Thus, plurality voting (PV) is used to sum up the obtained results of metrics [8]. PV counts the number of times that a normalization technique being the first ranking/ordering with respect to the applied metrics. So, the normalization technique that has the large number of times being the first ranked would be recommended as a more appropriate technique [8].

Table 6. Ordering of normalization techniques with respect to the metrics and using plurality voting

	Euclidean↑	STD↑	RCI↑	MSE↓	Mean Ks↑	PV
Max	3	3	1	2	3	1
Max-Min	2	2	4	1	4	1
Sum	4	4	1	2	2	1
Vector	5	5	5	5	5	0
Fuzzification (Gaussian)	1	1	1	2	1	4

From the obtained results (Table 6) using the proposed assessment framework and plurality voting, the Fuzzification (Gaussian) normalization technique is the best one due to having the large number of times being the first order considering the implemented metrics, while the approach by [10], recommended the Sum normalization technique.

Comparing the results of both approaches, we believe our framework provides more robust and reliable results than the ones obtained by [10], because we cover a wide range of metrics (STD, Euclidean distance, Mean Ks, RCI, and MSE). On the other hand, implementing PV enabled us to aggregate the obtained results from different used metrics. In this case study, the authors of [10] just calculated time and space complexity with MATLAB, and these results are highly dependent on the style of MATLAB users/programmers. For instance, someone can code Sum normalization technique in a manner that obtains time and space complexity twice higher than someone else. Therefore, our proposed framework ensures more accurate and reliable results to support decision makers.

5 Conclusions

This paper addressed the applicability and robustness of the assessment framework to evaluate different normalization techniques and recommend the most appropriate technique for the TOPSIS method. A case study borrowed from [10] was used to analyze the effects of five chosen normalization techniques namely Max, Max-Min, Sum, Vector, and Fuzzification (Gaussian) and recommended the Fuzzification (Gaussian) as the best technique to be used in the related case study. Moreover, the obtained results were compared with the initial results by the authors of [10] and showed that our results are more accurate and robust than the initial results due to the implementation of several metrics from the assessment framework.

There are several normalization techniques introduced in the literature [1] which are not evaluated yet. In future we plan to evaluate further normalization techniques for different MCDM methods (e.g. PROMETHEE, MOORA, COPRAS, and etc.) using the proposed framework. Also, the validation of the assessment framework has not been done with real-world case studies, which is planned as future work.

Acknowledgements. This work was funded in part by the Center of Technology and Systems (CTS) and the Portuguese Foundation for Science and Technology (FCT) through the Strategic Program UIDB/00066/2020.

References

1. Jahan, A., Edwards, K.L.: A state-of-the-art survey on the influence of normalization techniques in ranking: Improving the materials selection process in engineering design. Mater. Des. **65**(2015), 335–342 (2015). https://doi.org/10.1016/j.matdes.2014.09.022
2. Tzeng, G.-H., Huang, J.-J.: Multiple Attribute Desicion Making: Methods and Applications. Taylor & Francis Group, Boca Raton, FL. 33487–2742 (2011)
3. Celen, A.: comparative analysis of normalization procedures in topsis method: with an application to Turkish deposit banking market. Informatica **25**(2), 185–208 (2014)
4. Chakraborty, S., Yeh, C.-H.: A simulation comparison of normalization procedures for TOPSIS. In: 2009 International Conference on Computers and Industrial Engineering, pp. 1815–1820. IEEE, Troyes (2009). https://doi.org/10.1109/iccie.2009.5223811
5. Chakraborty, S., Yeh, C.-H.: A simulation based comparative study of normalization procedures in multiattribute decision making. In: AIKED'07 Proceedings of the 6th Conference on 6th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases, vol. 6, pp. 102–109, World Scientific and Engineering Academy and Society (WSEAS) Stevens Point, Wisconsin, USA (2007)
6. Mathew, M., Sahu, S., Upadhyay, A.K.: Effect Of Normalization Techniques In Robot Selection Using Weighted Aggregated Sum Product Assessment (2017)
7. Vafaei, N., Ribeiro, R.A., Camarinha-Matos, L.M.: Data normalisation techniques in decision making: case study with TOPSIS method. Int. J. Inf. Decis. Sci. **10**(1), 19 (2018). https://doi.org/10.1504/IJIDS.2018.090667
8. Vafaei, N., Ribeiro, R.A., Camarinha-Matos, L.M.: Selecting normalization techniques for the analytical hierarchy process. In: Technological Innovation for Life Improvement. DoCEIS 2020. IFIP Advances in Information and Communication Technology, pp. 43–52. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-45124-0_4
9. Vafaei, N., Ribeiro, R.A., Camarinha-Matos, L.M.: Selection of normalization technique for weighted average multi-criteria decision making. In: Camarinha-Matos, L.M., Adu-Kankam, K., Julashokri, M. (eds.) Technological Innovation for Resilient Systems. DoCEIS 2018. IFIP Advances in Information and Communication Technology, vol 521, pp. 43–52. Springer, Cham, (2018). https://doi.org/10.1007/978-3-319-78574-5_4
10. Lakshmi, T.M., Venkatesan, V.P.: A comparison of various normalization in techniques for order performance by similarity to ideal solution (TOPSIS). Int. J. Comput. Algorithm **3**(3) (2014)

11. Doumpos, M., Grigoroudis, E.: Multicriteria Decision Aid and Artificial Intelligence. John Wiley & Sons, Ltd, Chichester, UK (2013). https://doi.org/10.1002/9781118522516
12. Aliasi, M.A., Zaiton, S., Hashimi, M., Samsudin, S.: Multi crteria decision making and its applications: a literature review. J. Teknol. Mklm. **20**(2), 129–152 (2008)
13. Ho, W.: Integrated analytic hierarchy process and its applications – a literature review. Eur. J. Oper. Res. **186**(1), 211–228 (2008). https://doi.org/10.1016/j.ejor.2007.01.004
14. Pan, N.-F.: Fuzzy AHP approach for selecting the suitable bridge construction method. Autom. Constr. **17**(8), 958–965 (2008). https://doi.org/10.1016/j.autcon.2008.03.005
15. Sheu, J.-B.: A hybrid neuro-fuzzy analytical approach to mode choice of global logistics management. Eur. J. Oper. Res. **189**(3), 971–986 (2008). https://doi.org/10.1016/j.ejor.2006.06.082
16. Kulturel-Konak, S., Coit, D.W., Baheranwala, F.: Pruned Pareto-optimal sets for the system redundancy allocation problem based on multiple prioritized objectives. J. Heuristics **14**(4), 335–357 (2007). https://doi.org/10.1007/s10732-007-9041-3
17. Efendigil, T., Önüt, S., Kongar, E.: A holistic approach for selecting a third-party reverse logistics provider in the presence of vagueness. Comput. Ind. Eng. **54**(2), 269–287 (2008). https://doi.org/10.1016/j.cie.2007.07.009
18. Wu, C.-R., Lin, C.-T., Chen, H.-C.: Integrated environmental assessment of the location selection with fuzzy analytical network process. Qual. Quant. **43**(3), 351–380 (2009). https://doi.org/10.1007/s11135-007-9125-z
19. Vafaei, N., Ribeiro, R.A., Camarinha-Matos, L.M., Valera, L.R.: Normalization techniques for collaborative networks. Kybernetes **49**(4) (2019). https://doi.org/10.1108/k-09-2018-0476
20. Vafaei, N., Ribeiro, R.A., Camarinha-Matos, L.M.: Normalization techniques for multi-criteria decision making: analytical hierarchy process case study. In: Camarinha-Matos, L.M., Falcao, A., Vafaei, N., Najdi, S (eds.) DoCEIS 2016. IFIP Advances in Information and Communication Technology Technological Innovation for Cyber-Physical Systems, pp. 261–269. Springer, Cham, Lisbon, Portugal (2016). https://doi.org/10.1007/978-3-319-31165-4_26
21. Felinto de Farias Aires, R., Ferreira, L., Galdino de Araujo, A., Borenstein, D.: Student selection in a Brazilian university: using a multi-criteria method. J. Oper. Res. Soc. **69**(4), 528–540 (2018). https://doi.org/10.1057/s41274-017-0242-3
22. de F. Aires, R.F., Ferreira, L.: A new approach to avoid rank reversal cases in the TOPSIS method. Comput. Ind. Eng. **132**, 84–97 (2019). https://doi.org/10.1016/j.cie.2019.04.023