Proper and improper uses of MCDA methods in energy systems analysis

Cinelli, M.; Burgherr, P.; Kadzinski, M.; Slowinski, R.

Decision Support Systems 163 (2022) 113848

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Proper and improper uses of MCDA methods in energy systems analysis

Marco Cinelli [a,b,*], Peter Burgherr [c], Miłosz Kadziński [a], Roman Słowiński [a,d]

[a] Institute of Computing Science, Poznan University of Technology, 60-965 Poznań, Poland
[b] Leiden University College, Leiden University, The Hague 2595DG, the Netherlands
[c] Paul Scherrer Institut (PSI), Laboratory for Energy Systems Analysis, Villigen PSI 5232, Switzerland
[d] Systems Research Institute, Polish Academy of Sciences, 01-447 Warsaw, Poland

ARTICLE INFO

Keywords:
Energy systems analysis
Decision making
Multiple criteria decision analysis
Decision support system
Method recommendation

ABSTRACT

Over the past few decades, the strategies to perform energy systems analysis have evolved into multiple criteria-based frameworks. However, there still remains a lack of guidance on how to select the most suitable Multiple Criteria Decision Analysis (MCDA) method. These methods provide different decision recommendations for the Decision Makers, including ranking, sorting, choice, and clustering of the alternatives (e.g., technologies or scenarios) under evaluation. They deal with a variety of data typologies and preferences, and lead Decision Makers in shaping the energy systems of the future. Here, we evaluate the MCDA methods used in 56 case studies performing energy systems analysis at different scales. We find that close to 60% of these studies chose an MCDA method that was not the most adequate for the respective decision problem. In particular, this concerned the use of weighting methods (e.g., Analytical Hierarchy Process) in MCDA approaches not suited for this type of weights, sub-optimal selection of MCDA techniques for specific types of problem statements, and lack of handling rather evident interactions in preference models. Our analysis demonstrates that these deficiencies can be overcome by using a recently developed methodology and software that support Decision Makers and analysts in selecting the most suitable MCDA method for a given type of decision-making problem.

1. Introduction

As energy systems become increasingly complex, it is ever more critical to aid their assessment using structured decision support methods [7]. The energy systems analysis community has shown considerable advancements in the last decades in terms of approaches and strategies for assessing these systems. Mono-criterion and mono-disciplinary assessments such as risk analysis [22] have been extended with frameworks combining multiple criteria approaches with other disciplines, such as resilience [1], sustainability [29,77], energy security [4], supplier selection [13], and spatial analysis [36]. This shift has resulted in the need to include an increasing number and variety of stakeholders and their respective preferences [69], together with the need for transparency permitting tracking the development and performance of the systems under evaluation [60].

Given these requirements, decision support methods have been developed to integrate the multiple types of data and preferences to provide what is called a decision recommendation to the Decision Makers (DMs), which could be a ranking, sorting, clustering, or choice of the target systems [17]. In other words, analysts have to convey to the DMs a wealth of information in an easily digestible form. This led to the increasing use of Multiple Criteria Decision Analysis (MCDA) methods to provide the required decision recommendations [44] since these methods are specifically developed for delivering this type of decision support [3,28]. The selection of MCDA approaches that are adequate for a given Decision-Making Problem (DMP) is essential as they allow for its detailed analysis, in-depth exploration and interpretation of various possibilities, and structured interactions between the DMs and the analyst.

Even if a large variety of MCDA methods has been used to assess the comprehensive performance of energy systems [11,21,24,35], the fit between the types of DMPs and the suitability of the used MCDA methods has been rarely verified. A few studies that did it follow. For example, Polatidis et al. [58] considered several features: the structure and measurement scale of the criteria, the pairwise comparison thresholds, the compensation rate between criteria, and the type of

Abbreviations: MCDA, Multiple Criteria Decision Analysis; DMP, Decision-Making Problem; DSS, Decision Support System.

* Corresponding author at: Leiden University College, Leiden University, The Hague 2595DG, the Netherlands.
E-mail addresses: m.cinelli@luc.leidenuniv.nl (M. Cinelli), peter.burgherr@psi.ch (P. Burgherr), milosz.kadzinski@cs.put.poznan.pl (M. Kadziński), roman.slowinski@cs.put.poznan.pl (R. Słowiński).

https://doi.org/10.1016/j.dss.2022.113848
Received 1 January 2022; Received in revised form 10 August 2022; Accepted 10 August 2022
Available online 27 August 2022
0167-9236/

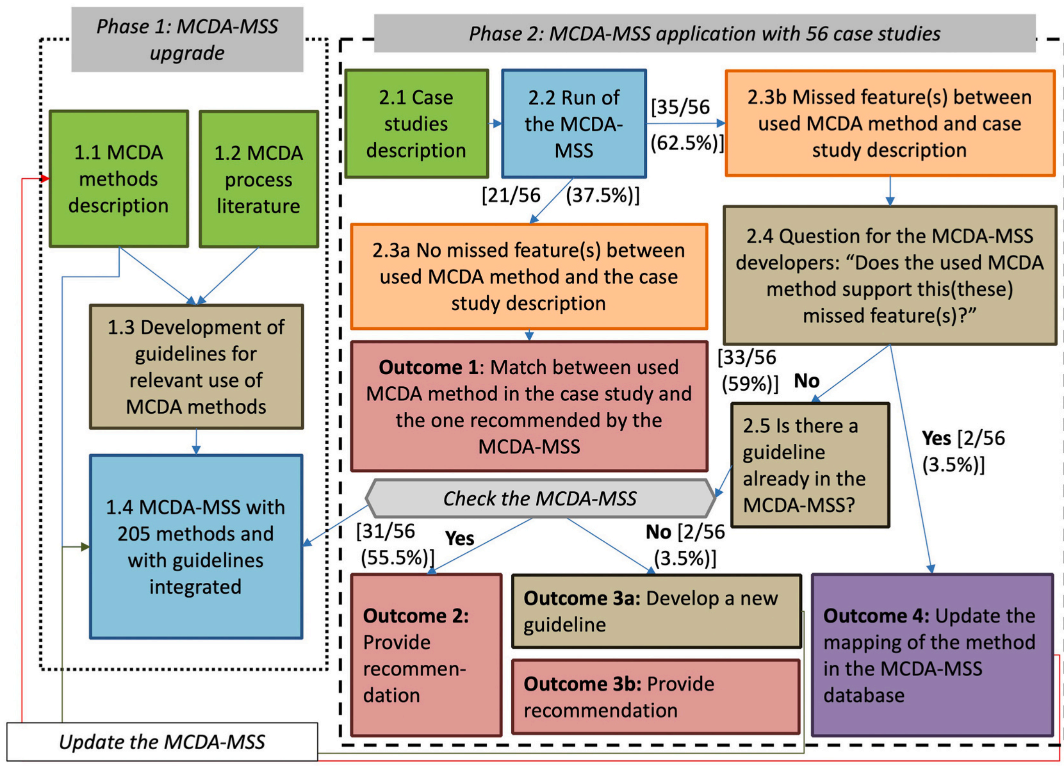

Fig. 1. The methodology used to upgrade and test the MCDA-MSS with the case studies (Green boxes: Knowledge conceptualization; Brown boxes: Knowledge development for the MCDA-MSS; Blue boxes: Software development and application; Orange boxes: Intermediate outcomes from the MCDA-MSS based on its application to the case studies; Light red boxes: Outcomes provided by the MCDA-MSS when there are either no missed features between the used MCDA method and the case study description or when guidelines from the software are missed; Purple box: Cases where the database of the MCDA-MSS has to be updated). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

exploitation of the preference model. Another study that justified the selection of MCDA methods for energy systems analysis is the one of Buchholz et al. [10], whose features included the type of problem statement, the structure of the criteria and their measurement scales, and the type and the exploitation of the preference model.

Even if there are a few published studies where the selection of the MCDA methods has been justified and studied based on a small set of features, the energy systems analysis community is still unaware of the proper and improper uses of MCDA methods in this very domain. The notable increase in the use of MCDA methods in this area demands that the chosen methods are tailored to the characteristics of the DMP under consideration. Ultimately, the employed approaches are expected to deliver results that follow an adopted way of reasoning, being consistent with the context of a decision process, characteristics of available data, possibilities for acquiring preferential information from the DMs, and working hypotheses on the aggregation of performances on multiple criteria. If the choice of a method neglects or contradicts some of these factors, we may end up with a dissatisfactory, suboptimal or even wrong recommendation.

In the area of energy systems analysis, there is no systematic assessment of the suitability of the MCDA methods that have been used. This paper starts tackling this research gap by considering a representative set of real-world case studies. By comparing the methods used by the authors of the case studies and those recommended by our newly developed methodology and software, we are capable of studying which past case studies used a proper MCDA method and which ones did not. In this regard, we provide three main contributions:

1. We tackle theoretical and technical issues in support of enhanced decision-making by upgrading the methodology and software that three of the authors of this paper developed to recommend MCDA methods, called the MCDA Methods Selection Software (MCDA-MSS) [18], accessible at http://mcdamss.com. The MCDA-MSS is freely accessible software that provides a structured and traceable path for the identification of the MCDA methods most suitable to a specific DMP, applicable to both new (prospective) and published

Table 1
Example of decision rules used in the MCDA-MSS and developed from the description of MCDA methods in the MCDA-MSS database.

Knowledge base (i.e., Rules)	MCDA methods	*"If the conjunction of conditions on some features is true*								*then*	*this MCDA method can be recommended"*
		Feature 1: Problem statement/typology				Feature 2: Criteria		Feature …			
		Ranking	Sorting	Clustering	Choice	Hierarchical criteria		…	…		
						No	Yes				
Rule 1	SMAA-TRI		✓			✓			✓		SMAA-TRI
Rule 2	DRSA for sorting		✓			✓		✓			DRSA for sorting
Rule 3	MAVT	✓					✓	✓			MAVT

Table 2
Literature reviews considered for the selection of the case studies for the MCDA-MSS test (* Reviews selected for the test).

Reference	Number of articles included in the reviews	Criteria for the evaluation of the literature reviews			
		Main types of alternatives	Included problem statements	Recurrent reported decision-making features	Most frequent MCDA methods
[46]*	51	Single and combinations of electricity generation technologies; Single and combinations of energy generation technologies; Multi-electricity & energy sources configurations; Geographic areas; Scenarios/policies; Strategic operational/business plans of energy companies; Types of heating systems; Designs of heating systems; Design of power plants (hydro); Energy crops; Cities; Desalination units; Countries	Choice; Ranking; Sorting	Criteria structure; Evaluation of criteria; Weighting; Aggregation algorithm	(Fuzzy) AHP/ANP, Weighted sums, (Fuzzy) VIKOR, (Fuzzy) TOPSIS, PROMETHEE, ELECTRE, MAVT, DEMATEL
[35]*	112	Single and combinations of electricity generation technologies; Single and combinations of energy generation technologies; Geographic areas; Design of power plants; Scenarios/policies; Countries; Strategic operational/business plans of energy companies	Choice; Ranking; Sorting	Evaluation of criteria; Weighting; Aggregation algorithm	AHP, ANP, TOPSIS, ELECTRE, PROMETHEE, VIKOR, TOPSIS, DEMATEL, Others
[7]	15	Energy storage systems	Choice; Ranking	Evaluation of criteria; Weighting; Aggregation algorithm	AHP, Weighted sums, MAVT, TOPSIS, PROMETHEE
[11]	79	Single and combinations of electricity generation technologies; Single and combinations of energy generation technologies; Geographical areas; Policies/scenarios	Choice; Ranking; Sorting; Description; Portfolio	Weighting	AHP, ANP, Weighted sums, TOPSIS, ELECTRE, VIKOR, PROMETHEE, MAUT, MAVT, MODM
[21]	12	Electricity generation technologies; Electricity generation scenarios	Ranking	Evaluation of criteria	MAVT, PROMETHEE, AHP, Weighted sum, TOPSIS
[24]	184	*(All alternatives focused on the assessment of renewable sources)* Energy technologies; Geographical areas; Scenarios for energy management; Barriers to energy project development	Choice; Ranking	Evaluation of criteria; Weighting; Aggregation algorithm	AHP, ANP, ELECTRE, PROMETHEE, MAVT
[41]	30	Scenarios/Policies for energy management; Single and combinations of energy generation technologies; Single and combinations of electricity generation technologies; Geographical areas; Energy storage; Energy recovery options	Choice; Ranking	Evaluation of criteria; Weighting	Weighted sums, ELECTRE, PROMETHEE, TOPSIS, MAUT, VIKOR, AHP
[43]	39	Electricity generation technologies; Energy supply strategies; Demand-side management programs; Energy resource scenarios;	Choice; Ranking; Sorting	Evaluation of criteria; Weighting; Aggregation algorithm	AHP, MAVT, MAUT, Goal programming, TOPSIS, PROMETHEE, ELECTRE,
[45]	54	Electricity generation technologies; Energy generation technologies; Geographical areas; Recycling strategies;	Choice; Ranking	Evaluation of criteria	(Fuzzy) AHP, ANP, VIKOR, (Fuzzy) TOPSIS, PROMETHEE
[49]	62	Single and combinations of energy generation technologies; Single and combinations of electricity generation technologies; Heating, ventilation, and air conditioning systems; Geographical areas; Scenarios/policies; Crops; Companies; Means of transport	Choice; Ranking	Evaluation of criteria; Aggregation algorithm	MAVT, MAUT, AHP, ELECTRE, PROMETHEE, Weighted sums, TOPSIS, VIKOR,
[70]	85	Geographical areas	Choice; Ranking	Evaluation of criteria; Normalization of criteria; Weighting; Aggregation algorithms	Weighted sums, TOPSIS, ELECTRE, VIKOR, Goal programming
[73]	55	Electricity generation technologies; Energy generation technologies; Policies	Choice; Ranking	Evaluation of criteria; Weighting; Aggregation algorithms	Weighted sums, AHP, PROMETHEE, ELECTRE, Fuzzy sets, Grey relational method
[74]	115	Electricity generation technologies; Energy generation technologies; Policies/scenarios; Geographical areas	Choice; Ranking	Aggregation algorithms	MAUT, AHP, ANP, ELECTRE, PROMETHEE

(retrospective) case studies. The upgrade consisted in formulating the guidelines to be respected when choosing an MCDA method.

2. We test the functionality of the MCDA-MSS by assessing whether the guidelines for MCDA methods selection embedded in the MCDA-MSS had been violated or not in a set of 56 case studies dealing with the analysis of energy systems at different scales.
3. We demonstrate how the expert knowledge in the MCDA-MSS can be updated based on its use with different case studies. The essential aspects of this update concern lessons learned from the irrelevant choice of MCDA approach, including the need to acquire knowledge about the structure of the DMP and requirements of the DMs, justify the process of selecting the method, and respect some rules facilitating the method's selection and allowing to avoid the mistakes made by others.

This paper is structured as follows. Section 2 presents the methodology used to develop the MCDA-MSS and apply it to the case studies. Section 3 provides the results of the test of the MCDA-MSS, focusing on the six guidelines that are embedded in the software and that were often not followed by the case studies. Section 4 discusses the main findings and Section 5 summarizes the main contributions of this research in the area of decision support.

Table 3
Example table of the application of the MCDA-MSS on a set of case studies.

Case studies	Used MCDA method in the case study	Missed features between used MCDA method and case study description by the authors of this paper	Recommended MCDA method (s) (i.e., complete match with case study)	MCDA methods match?	“I don’t know” features in case study mapping	Missed features from Guideline (s)	Reason(s) for lack of match between used MCDA method and case study description	Guideline(s) not followed	1st recommended MCDA method that complies with the guideline(s)	Ratio of covered features to required ones	Missed features	2nd recom mended MCDA method that complies with the guideline(s)
Case study 1	SMAA-TRI	None	SMAA-TRI	✓		Not applicable as there is a match between the MCDA method used in the case study and the one recommended by the MCDA-MSS						
Case study 2	MAVT	Weights = Relative importance coefficients	None	✘		Weights = Relative importance coefficients	Use of weights meaning importance coefficients in methods that require weights meaning trade-offs	1. Criteria weights are tailored to each MCDA method	PROMETHEE II	24/25	It works with pairwise comparisons between the alternatives and not with piecewise linear functions	…
Case study 3	ELECTRE III	Interactions between criteria = Yes	ELECTRE III-Inter actions	✘		Interactions between criteria = Yes	No consideration of interactions among criteria	6. The interdependencies between the criteria can refine the preference model	ELECTRE III-INT	22/22	None	None

2. Methodology

The methodological framework of this research is presented in Fig. 1. It includes two phases, the first one to upgrade the MCDA-MSS and the second one to test it with the case studies. It shows the iterative nature of the process, which can be used to identify the improper uses of the MCDA methods as well as iteratively improve our software. Fig. 1 shows that the application of the MCDA-MSS allowed us to (i) understand if the MCDA method used in the case study was a relevant one and within the set recommended by MCDA-MSS (outcome 1), (ii) provide recommendations that respect the existing guidelines (outcome 2), (iii) develop new guidelines and provide recommendations that respect them (outcome 3), and (iv) update the MCDA-MSS database (outcome 4).

2.1. Phase 1: MCDA-MSS upgrade

During phase 1, the focus has been on the upgrade of the software. This part was devoted to the study of 205 methods and their evaluation based on 156 features to describe their capabilities of dealing with different DMPs. The MCDA-MSS is structured in four main sections, and Cinelli et al. [18] provide a detailed description of each feature included in the software and the methods that are part of it. The first section is the problem typology, where the type and structure of the DMP are defined. The second one refers to the desired features of the preference model. The third section includes the type, modality, and frequency of the preferences of the DM. The last one is devoted to exploiting the preference relation induced by the preference model, which shapes the strategy used to derive and enrich the decision recommendation. A broad set of methods that include traditional and recent developments in the MCDA area compose the database. They are representative of the three main families of MCDA methods, namely scoring functions (109 methods), binary relations (81 methods), and decision rules (15 methods).

The MCDA-MSS adopts rule-based modeling to provide the recommendation of the MCDA methods to its users. This strategy belongs to the first group of Decision Support Systems (DSSs) presented in Cinelli et al. [17], which has transparency and simplicity as key advantages. This modeling uses rules in the form of *"if the conjunction of conditions on some features is true, then this MCDA method can be recommended"*. A simplified example is provided in Table 1, where the first rule states that if a decision analyst is searching for a sorting method without a hierarchical structure of the criteria, then SMAA-TRI and DRSA for sorting are suitable approaches.

During the upgrade of the MCDA-MSS, based on the available literature on the MCDA methods and process (steps 1.1 and 1.2 in Fig. 1, respectively), four guidelines that need to be respected when choosing an MCDA method were formulated (step 1.3 in Fig. 1) and integrated into the MCDA-MSS (step 1.4 in Fig. 1). These four guidelines were thus driven by the analysis of the 205 methods included in the MCDA-MSS, in addition to the original taxonomy of the MCDA process proposed in Cinelli et al. [17]. They are presented and discussed in Sections 3.1 to 3.4.

2.2. Phase 2: MCDA-MSS application with the case studies

During phase 2, the analysis of the case studies with the MCDA-MSS was implemented. This firstly consisted in describing each case study by means of the 156 decision-making features of the software (step 2.1 in Fig. 1). Each feature was scored with a 1 or 0 according to whether the case study qualified or not for its activation. The lack of information was accounted for as well.

2.2.1. Selected case studies

Several literature reviews were considered for the selection of the case studies. These were chosen from the search of Web of Science (WOS) and Scopus database with several keyword combinations,

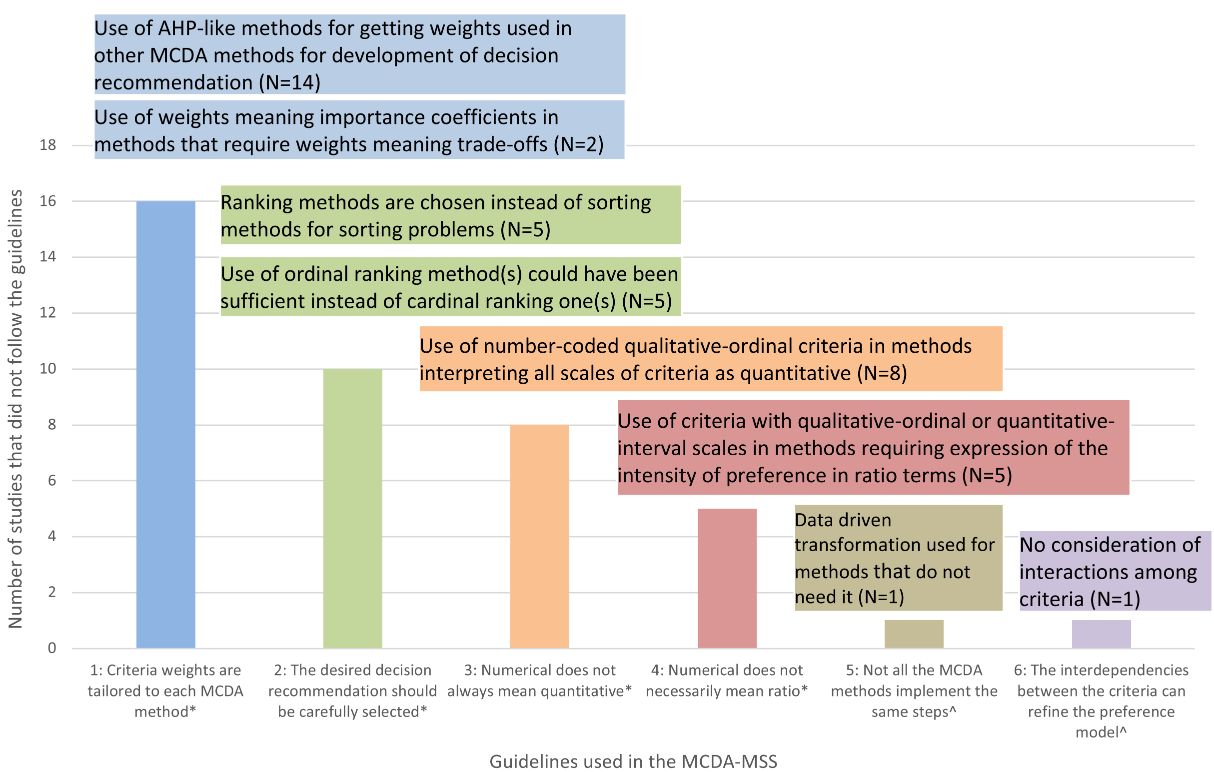

Fig. 2. Reasons for improper use of MCDA methods. Frequencies of lack of implementation for each guideline in the analyzed set of case studies with respect to the specific reasons for improper use of the MCDA method. *Guidelines defined during the upgrade of the MCDA-MSS; ^Guidelines developed during the test of the MCDA-MSS with the case studies.

including “MCDA”, “energy systems”, “analysis”, “decision making”, and “multiple criteria”.

Thirteen literature reviews were considered for the selection of the case studies. These are listed in Table 2, showing the criteria used to evaluate their suitability for the test of the MCDA-MSS. These included the main types of alternatives, problem statements, other recurrent decision-making features, and the most frequent MCDA methods. Four selection requirements were set to justify the inclusion of the reviews in the database of case studies. They include (i) the coverage of a broad range of alternatives (such as technologies, systems, designs, policies, scenarios, companies, or geographical units) to represent a broad scope of problems considered in energy systems analysis, (ii) the inclusion of several problem statements to reflect a variety of DMPs that are handled with MCDA, (iii) the consideration of several features so that the decision-making process conducted by the authors could be reconstructed and potentially revised, and (iv) the inclusion of a wide range of MCDA methods (not overlapping with methods from the other studies) so that the considered set is not dominated by a few most popular approaches that might serve as default choices for the less experienced analysts in a given domain. Two reviews were selected as they matched all the selection requirements, the one of Mardani et al. [46] and Kaya et al. [35]. When compared to the other reviews, the discriminatory factors were the coverage of a broad range of alternatives and several different MCDA methods. This was primarily evident in Mardani et al. [46], which was thus chosen to develop the bulk of the case studies, with 42 that were included in the database. Given that a wide majority of methods reported in the second chosen review are primarily the same as in the first one (i.e., AHP/ANP, TOPSIS, ELECTRE, PROMETHEE, VIKOR), the focus, in this case, has been placed on under-represented methods, resulting in additional 14 case studies selected from column “Other” in Table 4 in Kaya et al. [35]. In this research, a case study is defined as a DMP solved with an MCDA method. This means that one research paper can contain more than one case study.

In order to maintain the research within a manageable reach, the limit of approximately 50 case studies seemed reasonable, similar to other research studies, such as Wątróbski et al. [75]. It must be noted that some studies from these reviews were not included in our set for at least two main reasons. First, some studies reported very limited information regarding the features that are included in the MCDA-MSS and are needed to describe the DMP. Their description would have been highly subjective and unreliable. Second, some studies were out of the topic as they did not directly relate to energy system analysis, such as the one of Podgórski [57], which is actually tailored to operational safety assessment and only considered criteria weighting.

Table 4
A sample of case studies that did not follow the MCDA-MSS guideline 1.

Case studies	Ligus [42]	Villacreses et al. [71]
Used MCDA method in the case study	MAVT	VIKOR
Missed features between used MCDA method and case study description by the authors of this paper	1. Weights = Relative importance coefficients	1. Weights = AHP; 2. MCDA method inconsistency management = Inconsistency with respect to preference comparison of criteria
Recommended MCDA method(s) (i.e., complete match with case study)	None	None
MCDA methods match?	✘	✘
Missed features from Guideline(s)	1. Weights = Relative importance coefficients	1. Weights = AHP
Reason(s) for lack of match between used MCDA method and case study description	1. Use of weights meaning importance coefficients in methods that require weights meaning trade-offs	1. Use of AHP-like methods for weighting combined with other methods for the development of decision recommendation
Guideline(s) not followed	1. Criteria weights are tailored to each MCDA method	1. Criteria weights are tailored to each MCDA method
1st recommended MCDA method that complies with the Guideline(s)	QUALIFLEX	BWM
Ratio of covered features to required ones, e.g., 24/25	18/21	22/24
Missed features	1. Criteria structure = Hierarchical; 2. Input information used by the method = Qualitatively & Type = Performance-based; 3. Comparison of performances = Performances are compared by the DM with respect to the non-graded intensity of preference	1. Criteria profiles = Yes; 2. Exploitation of the preference model = Univocal recommendation without output variability analysis & single and deterministic model
2nd recommended MCDA method that complies with the Guideline(s)	…	…

Table 5
A sample of case studies that did not follow the MCDA-MSS guideline 2.

Case studies	Hobbs and Horn [30]	Baležentis and Streimikiene [5]
Used MCDA method in the case study	WAM with performances transformation	WASPAS
Missed features between used MCDA method and case study description by the authors of this paper	1. Problem statement = Sorting & Order of classes = Complete; 2. Problem statement = Sorting & scale leading the recommendation = Cardinal; 3. Problem statement = Sorting & Cardinality = without constraints	1. Problem statement = Ranking & Scale leading the recommendation = Ordinal
Recommended MCDA method(s) (i.e., complete match with case study)	1. TOPSIS-SORT-B	None
MCDA methods match?	✘	✘
Missed features from Guideline(s)	1. Problem statement = Sorting	1. Problem statement = Ranking & Scale leading the recommendation = Ordinal
Reason(s) for lack of match between used MCDA method and case study description	1. Method(s) not suited for this type of decision-making problem (in this case sorting) has (have) been selected	1. Use of ordinal ranking method(s) could have been sufficient instead of cardinal ranking one(s)
Guideline(s) not followed	2. The desired decision recommendation should be carefully selected	2. The desired decision recommendation should be carefully selected
1st recommended MCDA method that complies with the Guideline(s)	NA	EVAMIX
Ratio of covered features to required ones, e.g., 24/25	NA	21/22
Missed features	NA	1. Input information used by the method = Quantitatively & Type = Performance-based, linear
2nd recommended MCDA method that complies with the Guideline(s)	NA	…

Table 6
A sample of case studies that did not follow the MCDA-MSS guideline 3 and 4.

Case studies	Golić et al. [27]	Georgiou et al. [26]
Used MCDA method in the case study	VIKOR	AHP
Missed features between used MCDA method and case study description by the authors of this paper	1. Input information used by the method = Qualitatively & Type = Performance based	1. Input information used by the method = Qualitatively & Type = Pairwise comparisons based; 2. Input information used by the method = Quantitatively & Type = Pairwise comparisons based
Recommended MCDA method(s) (i.e., complete match with case study)	None	None
MCDA methods match?	✘	✘
Missed features from Guideline(s)	1. Input information used by the method = Qualitatively	1. Input information used by the method = Qualitatively; 2. Input information used by the method = Quantitatively
Reason(s) for lack of match between used MCDA method and case study description	1. Use of number-coded qualitative-ordinal criteria in methods interpreting all scales of criteria as quantitative	1. Use of criteria with qualitative-ordinal or quantitative-interval scales in methods requiring the expression of the intensity of preference in ratio terms (e.g., AHP)
Guideline(s) not followed	3. Numerical does not always mean quantitative	3. Numerical does not always mean quantitative; 4. Numerical does not necessarily mean ratio
1st recommended MCDA method that complies with the Guideline(s)	MAVT	EVAMIX
Ratio of covered features to required ones, e.g., 24/25	21/23	17/24
Missed features	1. Comparison of performances = Performances are transformed with a data-driven normalization approach and then compared; 2. Criteria profiles = Yes	1. Problem statement = Ranking & Scale leading the recommendation = Cardinal; 2. Criteria evaluation = Deterministic & Type = Relative performance comparison; 3. Comparison of performances = Performances are compared by the DM with respect to the graded intensity of preference; 4. MCDA method inconsistency management = Yes & Type = Cardinal inconsistency; 5. Weights = Relative importance with pairwise comparison ratio; 6. Indirect preferences for ranking or choice = Comparisons of reference alternatives with respect to the intensity of preference expressed on a ratio scale; 7. Exploitation of the preference model = Univocal recommendation without output variability & Type = Single representative and algorithmic
2nd recommended MCDA method that complies with the Guideline(s)	…	…

Table 7
A sample of case studies that did not follow the MCDA-MSS guideline 5 and 6.

Case studies	Jun et al. [34]	Papadopoulos and Karagiannidis [56]
Used MCDA method in the case study	ELECTRE II	ELECTRE III
Missed features between used MCDA method and case study description by the authors of this paper	1. Comparison of performances = Performances are transformed with a data-driven normalization approach and then compared	1. Problem statement = Ranking & Scale leading the recommendation = Cardinal; 2. Interactions between criteria = Yes
Recommended MCDA method(s) (i.e., complete match with case study)	None	None
MCDA methods match?	✘	✘
Missed features from Guideline(s)	1. Comparison of performances = Performances are transformed with a data-driven normalization approach and then compared	1. Interactions between criteria = Yes
Reason(s) for lack of match between used MCDA method and case study description	1. Data-driven transformation used for methods that do not need it (e.g., ELECTRE)	1. No consideration of interactions among criteria
Guideline(s) not followed	5. Not all the MCDA methods implement the same steps	6. The interdependencies between the criteria can refine the preference model
1st recommended MCDA method that complies with the Guideline(s)	ELECTRE II	ELECTRE III-INT
Ratio of covered features to required ones, e.g., 24/25	23/24	21/22
Missed features	1. Comparison of performances = Performances are transformed with a data-driven normalization approach and then compared	1. Problem statement = Ranking & Scale leading the recommendation = Cardinal
2nd recommended MCDA method that complies with the Guideline(s)	…	…

2.2.2. Test of the MCDA-MSS with the case studies

Each case study was described according to the 156 features in the MCDA-MSS, which can be found in Cinelli et al. [15], together with their reference, the type of alternatives, and the MCDA method used in each of them. The description of each case study was achieved by answering all questions in the MCDA-MSS (i.e., fill in all the features), using the information reported in the corresponding publication. For some features, information is reported explicitly, while for others, deductions and assumptions were made based on the method used by the authors. Lastly, there were cases where information about one or more features was not reported, so missing data was chosen for them. As a result of this process, a matrix of 56 rows and 156 columns was generated. The data used to develop the matrix consisted of 86.14% explicitly reported information in the case studies, 3.18% deductions by the authors of this paper, 7.92% assumptions by the authors of this paper, and 2.76% of missing data.

The input file with all the case studies described according to the 156 features was then run in the MCDA-MSS (step 2.2 in Fig. 1) to assess the correspondence between the description of the case studies with the 205 MCDA methods. The MCDA-MSS returned an output file, as shown in Table 3, distinguishing between the studies with a match or not of the used MCDA methods with those recommended by the software. In case of a match (step 2.3a and outcome 1 in Fig. 1), this means that there are no features missed between the MCDA method used by the authors of the case study and the description of the case study by the authors of this paper (case study 1 in Table 3). In case of no match (step 2.3b in Fig. 1), it implies that there is at least one feature missed between the MCDA method used by the authors of the publication and the case study description by the authors of this paper. At this point, a key question is raised for the developers of the MCDA-MSS, being "Does the used MCDA method support this (these) missed feature(s)?" (step 2.4 in Fig. 1). If the answer is "No", this leads to a further question for the analysts, being whether a guideline already exists or not in the MCDA-MSS (step 2.5 in Fig. 1). The database of the MCDA-MSS is thus analyzed, and in case the answer is "Yes", the reasons for lack of match between the used MCDA method and the case study description are listed. This is followed by the guideline(s) not respected in the case study, and then by the recommended MCDA methods with (case study 2 in Table 3) or without (case study 3 in Table 3) missed decision-making features (outcome 2 in Fig. 1). If the answer is "No", a new guideline is developed (outcome 3a in Fig. 1), and then the recommended MCDA methods with (case study 2 in Table 3) or without (case study 3 in Table 3) missed decision-making features are provided (outcome 3b in Fig. 1). In this research, two new guidelines (outcome 3a in Fig. 1) were developed based on the test with the selected case studies, and they are presented and discussed in Sections 3.5 and 3.6.

In case the answer is "Yes" to the question "Does the MCDA method support this(these) missed feature(s)?", it implies that the MCDA method chosen by the authors of the publication supports the missed feature(s), which leads to outcome 4 in Fig. 1. This implies that the developers of the MCDA-MSS can update its database, re-run the software and obtain outcome 1 in Fig. 1 for these case studies (i.e., the match between the used MCDA method in the case study and the one recommended by the updated MCDA-MSS).

3. Results

Four main findings can be derived from the test of the MCDA-MSS with the 56 literature case studies. First, 37.5% (i.e., $N = 21/56$) of the case studies used a method that was also recommended by the MCDA-MSS (outcome 1 in Fig. 1). One example is the case study of Alanne et al. [2], who used the PAIRS method [66] to score and rank ten micro-cogeneration and traditional heating systems for a Finnish single-family house based on environmental and economic criteria.

Second, four guidelines from the MCDA-related literature were not followed in more than half of the case studies (outcome 2 in Fig. 1). This means that 31 case studies out of 56 (~55%) used an MCDA method that is unsuitable for the specific DMP.

Third, the improper use of MCDA methods in two further case studies (~3.5% of the total) led to the development of two new guidelines (outcome 3a in Fig. 1).

Fourth, two case studies (~3.5%) led to an improvement of the MCDA-MSS database of MCDA methods (outcome 4 in Fig. 1). This update of the MCDA methods descriptions was implemented for the PAIRS [66] and MACBETH [6] methods. In the former, this implied adding the capability of accepting interval evaluation of performances and the option of data-driven normalization. In the latter, the update resulted in adding the capability of handling hierarchical as well as qualitative criteria.

The frequencies of lack of implementation of the guidelines in the analyzed case studies with respect to the specific reasons for the improper use of the MCDA methods are presented in Fig. 2. The figure distinguishes between the four guidelines defined during the development of the software (labeled with an asterisk "*") and the two guidelines developed and added during the test of the software with the case studies ("^"). It is noted that the overall number of improper uses is 41, which is higher than the one (i.e., 33) shown in Fig. 1, and this is because some case studies did violate more than one guideline.

The remainder of this section explains the rationale behind each guideline used in the MCDA-MSS, and discusses the case studies that did not follow a specific guideline. Tables 4 to 7 show at least one case study per guideline for which a mismatch was found, as well as the MCDA methods recommended by MCDA-MSS for these cases. All the results from the test of the MCDA-MSS with the 56 case studies are available in Cinelli et al. [16].

3.1. Guideline 1: Criteria weights are tailored to each MCDA method

The values of the weights of the criteria are essential to represent the preferences and priorities of the stakeholders involved in the decision-making process. The meaning of the weights assigned to criteria can be of different types, trade-offs on one side and importance coefficients on the other side.

Trade-off weights define substitution (trade-off) rates between the units of the different criteria in a weighted sum aggregation [31]. If w_{c_1} is the weight of one unit in criterion 1 (c_1) and w_{c_2} is the weight of one unit in criterion 2 (c_2), then one unit in c_1 is worth the same as $\frac{w_{c_1}}{w_{c_2}}$ units in c_2. For example, if $w_{c_1}= 0.3$ and $w_{c_2}= 0.6$, it means that one unit of c_1 is worth the same as $0.3/0.6 = 0.5$ units in c_2. In other words, their ratio means how much of one criterion compensates for a unit change of the other criterion. In this example, their ratio (i.e., $0.3/0.6 = 0.5$) means that two units of c_1 compensate for a unit change of c_2. Moreover, when a linear aggregation method is used (e.g., additive weighted mean), the substitution rates depend on the measurement scales [52]. Consequently, if the measurement scales change, the trade-offs need to be modified, hence the weights too.

Importance coefficients weights express the intrinsic importance of each criterion, meaning their voting power [23]. This type of weights is typically used in methods that compare two alternatives (at a time), say $a_1$and a_2, and conclude that a_1 is at least as good as (i.e., outranks) a_2 if a sufficient majority of the criteria supports this conclusion. This type of weights does not typically mean that being better in one criterion may compensate for being worse in another one. These weights are independent of any measurement scale employed since they represent the importance of the criterion per se [52]. As an example of this type of weights, assuming the cumulative weights of all criteria amount to a 100% majority (unanimity), the weights can be assigned as a share of this percentage to define the weight (or voting power) assigned to each criterion. Let us assume that c_1 has a weight of 45%, c_2 has a weight of 20%, c_3 has a weight of 35%, and a_1 outperforms a_2 on c_1 and c_2. The coalition of c_1 and c_2 has a joint weight of 65%, which would be sufficient to state that a_1 outranks a_2 if a simple majority is required but insufficient if a majority is set to 85% (for instance).

A further distinction in weighting is between weights obtained with relative preference comparisons on a ratio scale (e.g., AHP and its extensions) and those derived with absolute statements related to their measurement scales. Those obtained with the first approach (e.g., relative comparison) cannot be used in methods requiring substitution rates derived with the second approach.

Given all these differences in the meaning of the weights, they cannot be used interchangeably in MCDA methods. Their meaning depends on the type of aggregation formula used to derive the decision recommendation. This means that, according to the chosen MCDA methods, the weights must be elicited with a relevant approach that provides meaningful weights to be used in them.

In the application of MCDA-MSS, we found two improper uses of weights, which led to trigger our first guideline *"Criteria weights are tailored to each MCDA method"*. The first is the use of weights meaning importance coefficients in methods that require weights meaning trade-offs (3.5% of the case studies, $N = 2/56$). Both studies with this improper use employed multiple attribute value theory (i.e., MAVT [37]) as the MCDA method and chose the weights with approaches requiring points allocation to the criteria [42]. The essence of point allocation-based weighting methods is that they require the decision makers to assign numbers to criteria according to their perceived significance, and they do result in importance coefficient weights [51,79].

In the case of Ligus [42], the MCDA-MSS recommends QUALIFLEX [54] as one of the methods that is as close as possible to this DMP, as summarized in Table 4. When compared to the one chosen by the authors of the case study, this method uses weights suitable for those developed by the authors of this publication (i.e., relative importance coefficients), and it requires the structure of the criteria to be flat, which can easily be adapted at the problem formulation stage. In addition, it uses the information on the criteria performances with pairwise comparisons between the alternatives and not by accounting for the individual performance on each criterion. Lastly, it compares the raw performances directly instead of requesting the DM to compare them with respect to the non-graded intensity of preference.

The second and most common improper use (25% of the case studies, $N = 14/56$) is the use of weights obtained from methods like the Analytical Hierarchy Process (AHP) [64] or Analytical Network Process (ANP) [65] in other MCDA methods also requiring weights but of a different type. The authors of these case studies use the AHP/ANP weights in a variety of MCDA methods, like ELECTRE and TOPSIS [68], VIKOR [67], PROMETHEE [26], COPRAS [78], and TOPSIS [40]. As the weights obtained with AHP and ANP have the specific meaning of relative preference expressed in ratio terms, they can be meaningfully used only in the methods they originate from.

Table 4 shows an example of this improper use with the study of Villacreses et al. [71], who employed VIKOR as the chosen method. The MCDA-MSS identifies this error and proposes as a relevant solution the Best Worst Method (BWM) [59]. However, there are two decision-making features that this method does not satisfy from the provided description of the case study. The first one is the use of criteria profiles, as the BWM does not need any. The second one is the way the preference model is exploited, as in this case, it is representative algorithmic and not single deterministic as for VIKOR (see Cinelli et al. [18] for a description of the different exploitations of the preference relation induced by the preference model).

3.2. Guideline 2: The desired decision recommendation should be carefully selected

The formulation of every DMP requires selecting the type of decision recommendation the stakeholders would like to receive. In MCDA, this is defined as selecting the type of problem statement and the most common ones in MCDA include ranking (i.e., place alternatives in a preference-oriented order), sorting (i.e., classify alternatives to preference-ordered classes), choice (i.e., choose an alternative or a subset of alternatives that satisfy specific preference requirements), and clustering (i.e., group alternatives according to similarity features or preference relations) [50,61].

The selection of the correct problem statement is pivotal to identifying the most suitable MCDA method, which is the driving reason for naming this guideline *"The desired decision recommendation should be carefully selected"*. As it can be seen in Fig. 2, it was found that close to 9% ($N = 5/56$) of the studies whose problem statement was actually a sorting one, chose methods for ranking problems, primarily using the provided score to assign the alternatives to decision classes. The case study of Hobbs and Horn [30] shows an example of this in Table 5. The authors chose a scoring and ranking method (i.e., weighted additive mean with performances transformation [33]) to define the level of recommendation to assign to a series of demand-side energy management plans. Given that the final desired decision recommendation is a set of decision classes (e.g., the plan is recommended, the plan is strongly recommended, the plan is not recommended), the methods more suitable for this type of DMP are sorting ones instead of scoring and ranking. The MCDA-MSS accounts for this requirement and recommends two methods that fully match the description of the DMP. These are TOPSIS-SORT-B and TOPSIS-SORT-C, both proposed by de Lima Silva and de Almeida Filho [20].

Further differentiation within each problem statement can also be driven by the type of outcome that rules the decision recommendation, which can either be a score or a set of binary relations. Scoring methods provide a cardinal recommendation, meaning that the distance between the alternatives has a meaning from a quantitative standpoint. For example, if the score of alternative a is 0.2 and the score of alternative b is 0.4, it means that there is a 0.2 difference in score between the two alternatives. Binary relations, on the contrary, result in an ordinal decision recommendation, with only the position of the alternatives having a meaning. With the same example above, the only conclusion that can be provided could be that alternative a performs worse than alternative b. Hence it will be in a lower rank position. Fig. 2 shows that MCDA-MSS found that for five of the case studies, an ordinal ranking method could have been sufficient instead of a cardinal ranking one. The case study of Baležentis and Streimikiene [5] is an example of the latter improper use. These authors do not show or discuss the score obtained with the WASPAS method [14] and, instead, only focus their discussion on the ranking of their alternatives, which are manufacturing processes. The MCDA-MSS accounts for this aspect and proposes ordinal ranking methods instead. As shown in Table 5, it proposes, for example, EVAMIX [72] as the most suitable method, with the consideration that it does not operate with linear performance-based modeling, as it is one based on pairwise comparisons between the alternatives.

3.3. Guideline 3: Numerical does not always mean quantitative

The guideline *"Numerical does not always mean quantitative"* builds upon the work of Roy [62], who discusses the use of numbers and their intrinsic meaning in the measurement scale of criteria. On the one hand, there are criteria characterized by physical properties that can be verified with measurements and experiments in an objective manner. These are quantitative criteria, and some examples are the amount of raw materials used (in, e.g., kg), the time required to perform a task (in, e.g., minutes), and the length of the transport (in, e.g., km). In these cases, the type of measurement scale is defined as quantitative, where the differences between the performances have a well-defined meaning. For example, 10 kg is twice as much as 5 kg and four times as much as 2.5 kg.

On the other hand, there are criteria that are developed by assigning values to sensory perceptions that cannot be measured objectively, and they are called qualitative criteria. These perceptions are given numerical values for the convenience of communication and information management. Examples are the risk perception or social acceptability of technologies, the job creation potential of policies, and the contribution to the social sustainability of company strategies. The values used to characterize the achievement (or not) of these criteria can be low, medium, and high, and they can be coded with, e.g., low = 1, medium = 2, high = 3. These increasing or decreasing values have a subjective meaning, and it does not mean that the difference between low and medium is the same as the one between medium and high (unless this is agreed with the DM). These criteria are not suited to methods that assume that all criteria have a quantitative meaning, such as those computing averages and/or distances [62], more generally those using usual arithmetic operations, except those that only consider the orders [47].

As an example, let us consider two criteria c_1 and c_2, where c_1 is quantitative (e.g., distance autonomy of an electric car) and c_2 is qualitative (e.g., the comfort of a car). Although both criteria have six echelon scales, the meaning of the numbers coding the echelons of these two criteria is different. In c_1, e.g., echelon 6 means 2 times more autonomy than 3, while in c_2, 6 means only that it is a higher comfort than any lower comfort echelon, without the possibility of comparing the differences of comfort between echelons and without knowing if comfort 6 is 2 times better than 3. In this situation, the performances of these two criteria cannot be aggregated by an algorithm that computes averages and/or distances between the performances. This is because the numbers coding echelons of c_2 do not have a quantitative meaning. Remarque that if the order of the 6 echelons of c_2 would be coded by even numbers, instead of consecutive numbers, then the sum of two performances, c_1+c_2, would be different, and the ranking of the cars could also change solely based on this coding choice, which is the result of a wrong interpretation of numbers coding the order in case of c_2.

Of the 56 case studies we analyzed (see also Fig. 2), over 14% ($N = 8/56$) did not follow this guideline, using number-coded qualitative criteria in methods interpreting all scales of criteria as quantitative. The case study of Golić et al. [27] is one of these, using VIKOR to score and rank four designs for solar water heating systems in Belgrade, Serbia. Due to the use of qualitative criteria, the MCDA-MSS signals the lack of respect for this guideline, and it proposes MAVT [38] instead as a suitable alternative, as shown in Table 6. This method does not need to use criteria profiles, and it requires the DM to compare the criteria performances with respect to the non-graded intensity of preference.

3.4. Guideline 4: Numerical does not necessarily mean ratio

Numbers are used to define the performances of alternatives. According to the criterion being evaluated, they can be of two types as presented above, qualitative and quantitative. The latter can yet be divided into two categories, being interval and ratio scales.

Interval scales are those where the zero-point has no meaning of an absolute zero as it is defined arbitrarily, and the values of such a criterion can be below the zero-point [55]. An example is a temperature expressed in Celsius degrees, where the values can be below zero. With interval scales, only distances between evaluations define the intensity of preference and can be compared. With this scale, one cannot interpret the ratio of evaluations as the intensity of preference. For example, on the Celsius temperature scale, 40-degree air is not two times warmer than 20-degree air.

Ratio scales have the same characteristics as interval ones, and in addition, the zero-point has the meaning of an absolute zero, so a ratio of evaluations may be interpreted as the intensity of preference [55]. An example is weight, where a 40 kg object is two times heavier than a 20 kg object. The zero-point, in this case, is the lowest available number. Other examples of ratio scales are age, height, and speed.

We found that 9% ($N = 5/56$) of the case studies used AHP and its extensions with qualitative and interval criteria, which these methods have not been developed for. Hence, we propose the guideline *"Numerical does not necessarily mean ratio"*. This means that analysts should be aware of the types of measurement scales of their criteria to receive a recommendation of a method that can meaningfully use the input information. To illustrate the reasoning behind this guideline, we provide an example where we consider one of the used criteria in a case study as the social acceptability of energy technologies, expressed in an ordinal Likert scale (1 to 5, with 1 as the worst, 3 as neutral, and 5 as the best value). In methods like AHP, the DM must define how many times more she prefers higher social acceptability values over lower ones. The answer is the supposed ratio of the scores. However, the differences in social acceptability do not have the meaning of the intensity of preference. These numbers (1 to 5) are just numerical codes of ordinal echelons. The difference between, e.g., 2 and 3 is not equal to the difference between 3 and 4, so these questions cannot be answered meaningfully. Consequently, a different method should have been used with such a case study, which only considers the order of the performances, and a difference in evaluations in such a scale has no meaning of the intensity of preference. An example of this type of improper use is presented in Table 6 for the paper of Georgiou et al. [26], who use several qualitative criteria to assess the energy performance of desalination units.

The MCDA-MSS signals that the study does not follow guidelines 3 and 4, and proposes one that does instead, being EVAMIX [72]. This method does, however, require several changes to the DMP, as EVAMIX does not provide a cardinal ranking, and the performances are not compared by the DM with respect to the graded intensity of preference as the raw performances are compared directly. What is more, the weights are relative importance coefficients defined per criterion and not based on pairwise comparison ratios, and no indirect preferences in the form of comparisons of reference alternatives with respect to the intensity of preference on a ratio scale are needed.

3.5. Guideline 5: Not all the MCDA methods implement the same steps

A tailored set of steps needs to be followed to let MCDA methods provide a decision recommendation. These methods are many, and not all of them require the same ones [13]. To help analysts using MCDA methods, several implementation checklists have been provided, including, e.g., the selection and structuring of the criteria [39], as well as their weighting and aggregation [48]. One common step for composite indicators, also defined as indices, is the normalization of the performances, which means the transformation of the evaluation of the performances from different scales to a common one. This can include normalizations that are dependent on the dataset, such as the min/max or the target [53], those imposed by the pre-defined scale with preferences from the decision maker (e.g., AHP [64]), and those without a predefined scale with preferences from the decision maker (e.g., MAVT [9]).

Even if normalization is widespread among MCDA methods, it is not required for all of them, and there are, in fact, many which exploit the raw information on the criteria to assess whether one alternative performs better (and possibly also by how much) in comparison to another one. This is the common procedure for outranking methods [25]. Independence from normalization was actually one of the main reasons that justified their development, since working with the raw performances can be seen as less "invasive" on the dataset. Consequently, using normalization in outranking methods is not a relevant practice as it goes against the nature of these methods, especially when qualitative criteria are part of the assessment set. The case study by Jun et al. [34] actually showed this improper use as the authors explicitly stated that criteria should be dimensionless to apply the chosen outranking algorithm (i.e., ELECTRE II), which does not actually apply to this method. Hence, this improper use led to the guideline *"Not all the MCDA methods implement the same steps"*, and the recommendation from the MCDA-MSS, reported in Table 7, is to reapply the same method but without any normalization.

3.6. Guideline 6: The interdependencies between the criteria can refine the preference model

During the problem formulation, the collaboration between the analyst and DM should lead to a discussion of the relevance of the selected criteria and the possible presence of synergies and redundancies between them [12]. Synergies are treated with MCDA methods using positive interactions, while redundancies by means of negative interactions. Criteria that score the same for all the alternatives in the set can be kept in the criteria set only if at least one type of interaction is present. Otherwise, the criterion does not bring any valuable information and should be deleted, because its presence leads to violating the conditions of a consistent family of evaluation criteria.

The MCDA-MSS test found that the case study of Papadopoulos and Karagiannidis [56], as summarized in Table 7, chose a method that does not consider interactions, even if one criterion (i.e., black-out costs) has the same performance for all the alternatives. This triggered the last guideline *"The interdependencies between the criteria can refine the preference model"*. In this situation, the MCDA-MSS points out in Table 7 this sub-optimal selection of method and recommends ELECTRE III-INT [8] instead, an extension of ELECTRE III to deal with interacting criteria. The only remaining feature not matched by this method is the provision of a cardinal ranking since its outcome is "just" an ordinal one.

4. Discussion

The MCDA-MSS embeds in its structure six guidelines for the relevant choice of MCDA methods for DMPs of a different type. Since these guidelines are integrated into the MCDA-MSS, the user does not actually "see" them while answering its questions, as (s)he is guided with the knowledge embedded in the software, while respecting these guidelines automatically.

In this research, the MCDA-MSS has been used retrospectively, testing whether its guidelines had been violated or not in published case studies using MCDA methods in energy systems analysis. The research we propose in this paper shows an interesting link with what Robyn Dawes did in the late 1970s', when he studied the proper and improper strategies to develop (linear) psychological models [19]. In our research, we still studied the proper and improper uses of some methods that have been made to develop models, but instead of focusing on those that propose weights to describe correlations between variables, we focused on those that have been shaped for energy systems analysis using the MCDA methodology.

A large share of erroneous weighting practices was discovered. Thus, future training of decision analysts should concern a broad range of weighting methods. In this way, they could account for the different meanings (in particular, importance coefficients and trade-offs between criteria) and suitability of weights. Then, they would be able to justify "going the extra mile" when choosing a weighting approach that is potentially more demanding in terms of required implementation effort but aligns with the requirements of the specific method and hence provides meaningful decision recommendations.

As far as the problem statement is concerned, we found that 18% of the case studies suboptimally selected it. This indicates the need to spend some additional effort at the problem structuring stage to decide what type of recommendation would be the best fit for every specific research study. We suppose that many improper uses were implied by the users' unconsciousness of the existence of approaches adequate for different types of problems, using various scales leading the decision recommendation, or tolerating additional requirements that can be imposed on the delivered results.

Many ($N = 13/56$, just over 23%) improper uses of MCDA methods have been found due to the lack of consideration of the intrinsic meaning in the measurement scales of criteria. Different measurement scales exist, with qualitative and quantitative being the broadest categories, and interval and ratio the two types within the quantitative category. MCDA methods are suited for different types of measurement scales, so choosing the relevant method according to the type of DMP also requires considering the scales of the criteria. If there is even only one criterion whose scale is ordinal, then a method that is tailored to (also) exploiting the order of the performances should be chosen [55]. In case studies where all the criteria have a ratio scale, the methods that express an intensity of preference in ratio terms are relevant ones. Methods that do require ratio scales should not be used with criteria that are either qualitative and/or interval. The main example, in this case, is the AHP method and its extensions. They operate by comparing alternatives and criteria on a ratio scale to define how many times one alternative is better than another and how many times one criterion is preferred over another one [32,76].

This research points out the need to stress that MCDA methods have their own implementation checklists, which can vary notably from one method to another. For this reason, decision analysts are encouraged to carefully study these checklists to avoid any mixes that might add layers of unnecessary complexity to the case study. For example, normalization of criteria is needed for some scoring-based MCDA methods, while it is irrelevant for those based on binary relations or decision rules. Using normalization in methods from the latter families would not only result in additional work for the analyst, but could also bias the final decision recommendation.

Lastly, decision analysts should strive for consistent families of evaluation criteria. They should make sure that possible synergies and redundancies between the criteria are tested and included whenever relevant. At the same time, they should make sure that any criterion should provide added value to the assessment. If that is not the case, for example, because a criterion is not diversifying the performances of the alternatives, or it does not show a preference orientation, then it should be omitted from the set.

The MCDA-MSS provides the most comprehensive software to support complex decision-making with 205 MCDA methods and more than 150 selection features. Nevertheless, it only represents the first step to helping perform multiple criteria-based energy systems analyses. Therefore, its future use and broad acceptance will strongly depend on the input of scholars and practitioners in this domain that can enrich the MCDA-MSS database with new methods and case studies to shape the software in a dynamic and evolving fashion.

This first test of the MCDA-MSS has shown that two main research avenues could be explored in future software implementations. First, the application of the MCDA-MSS to more case studies in the energy systems analysis literature and/or to other research areas, possibly leading to further guidelines to be added to the software. This could result in the identification of other misuses in MCDA methods selection that were not observed in the considered studies. These could include some reported in, e.g., Roy and Słowiński [63] and Roy [62], which encompass, among others, (i) confusing indifference and incomparability relations, (ii) selecting a preference model characterized by an undesired compensation level, or (iii) neglecting evident imprecision, uncertainty, and indeterminacy when computing the decision recommendation.

Second, the extension of the software database with more methods and its application to the same case studies to verify if any method exists that completely matches the description of the case studies that did not have any complete match with the previous version of its database. In addition, a logical next step for published studies with improper uses of MCDA methods would be to re-calculate the results with the recommended method(s) to assess whether the main findings and conclusions would change. This would mean understanding if the MCDA-MSS does not just improve the whole methods selection process, but also leads to better decisions, both in terms of accuracy and suitable and correct methodology.

5. Conclusions

This paper presents the first systematic assessment of the suitability of using MCDA methods in a representative set of real-world case studies concerning energy systems analysis. While achieving this, it provides several contributions at different levels in the area of DSSs. These include:

1. Foundations-level: We embed an extensive taxonomy of the MCDA process and a set of guidelines for MCDA methods selection in a DSS infrastructure, enabling considerable improvements in decision-making;
2. Functionality- and interface-level: We streamline the implementation of the MCDA taxonomy and guidelines in a web-based DSS, available for free at http://mcdamss.com; we show the DSS's use in computer-supported research in a specific area and exhibit sound, non-trivial knowledge from such an application;
3. Implementation-level: We demonstrate the use of the DSS in more than 50 real-world case studies, together with its iterative improvement process;
4. Evaluation- and impact-level: We show that our DSS can have an actual impact in shaping decision-making processes by guaranteeing a proper use of MCDA methods. It can lead MCDA methods developers and practitioners (including consultants and analysts) who work with DMs in real-life case studies. Also, it can be used for educational purposes to convey knowledge to less experienced users

on the common mistakes and the critical guidelines to be followed in the method selection process.

Even if the software test is focused on the energy systems analysis, the described DSS, the formulated guidelines, and the arguments for updating the expert knowledge are also relevant for other domains in which the MCDA methods find use.

This research provides three main practical contributions to the practice of decision support. Firstly, it enables decision analysts to "look back" at what has been done and published in the scientific literature in this domain. It allowed us to discover (ir-)relevant uses of the MCDA methods, explain the rationale, propose more adequate methods, and motivate their selection by referring to the missed features. The conducted research confirmed that the authors of the case studies mostly focused on obtaining results from applying the MCDA methods rather than justifying the method selection. This has resulted in a large share (about 60%) of improper uses of MCDA methods, neglecting some essential aspects of the DMPs and potentially leading to biased, undesired, or sub-optimal decision-making. This suggests that decision analysts should allocate more time to learning about the structure of the DMP and collect information that can be used to learn the requirements of the DMs to select or develop the most suitable MCDA method. Such information concerns problem formulation, preference elicitation, information, and models, and decision recommendation construction. The MCDA-MSS can be used as the flagship tool to lead this effort, since it is freely accessible software that provides a structured and traceable path for the identification of the MCDA methods most suitable to a specific DMP, applicable to both new (prospective) and published (retrospective) case studies.

Secondly, the strategy proposed to update the MCDA-MSS during the test with the case studies shows an important virtuous feedback loop. In fact, it ensures validating if the MCDA methods support any of the features that the developers of the MCDA-MSS had not considered met by such methods. This is a valuable demonstration of co-constructive knowledge management of the MCDA methods. Overall, the analysis conducted with the 56 studies validates the practical usefulness of the MCDA-MSS, which is the broadest available database of MCDA methods and features relevant for their selection. Specifically, when describing all these studies, 76 features were used, whereas more than 70 different methods were recommended among the most suitable three. Such a great variety of characteristics and approaches used in the analysis and an even more extensive system's scope support the employment of MCDA-MSS for future real-world decision aiding.

Lastly, the conducted research provides considerations on the improper use of MCDA methods due to incorrect assumptions, unjustified interpretations, and false compromises. The conclusions of these considerations should be used in the training of decision analysts and subsequently followed by the practitioners when they analyze the characteristics of the DMPs and select the approaches that best fit their description.

Funding

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement No 743553, and from the Swiss National Science Foundation under grant agreement No IZSEZ0_193662. Miłosz Kadziński acknowledges financial support from the Polish National Science Center under the SONATA BIS project (grant no. DEC-2019/34/E/HS4/00045).

CRediT authorship contribution statement

Marco Cinelli: Conceptualization, Methodology, Data curation, Writing – original draft, Writing – review & editing. **Peter Burgherr:** Conceptualization, Writing – review & editing, Supervision. **Miłosz Kadziński:** Methodology, Writing – review & editing. **Roman Słowiński:** Writing – review & editing, Supervision.

Declaration of Competing Interest

None.

Data availability

See: https://doi.org/10.5281/zenodo.5808839, https://doi.org/10.5281/zenodo.5808861

Acknowledgments

The authors would like to thank Grzegorz Miebs for developing the software interface to test the implementation of the guidelines in the MCDA-MSS with the case studies.

References

[1] S. Ahmadi, Y. Saboohi, A. Vakili, Frameworks, quantitative indicators, characters, and modeling approaches to analysis of energy system resilience: A review, Renew. Sust. Energ. Rev. 144 (2021), 110988.
[2] K. Alanne, A. Salo, A. Saari, S.-I. Gustafsson, Multi-criteria evaluation of residential energy supply systems, Energy Build. 39 (2007) 1218–1226.
[3] P.A. Alvarez, A. Ishizaka, L. Martínez, Multiple-criteria decision-making sorting methods: a survey, Expert Syst. Appl. 183 (2021), 115368.
[4] B.W. Ang, W.L. Choong, T.S. Ng, Energy security: definitions, dimensions and indexes, Renew. Sust. Energ. Rev. 42 (2015) 1077–1093.
[5] T. Baležentis, D. Streimikiene, Multi-criteria ranking of energy generation scenarios with Monte Carlo simulation, Appl. Energy 185 (2017) 862–871.
[6] C.A. Bana E Costa, J.-C. Vansnick, The MACBETH approach: Basic ideas, software, and an application, in: N. Meskens, M. Roubens (Eds.), Advances in Decision Analysis, Springer, Netherlands, Dordrecht, 1999, pp. 131–157.
[7] M. Baumann, M. Weil, J.F. Peters, N. Chibeles-Martins, A.B. Moniz, A review of multi-criteria decision making approaches for evaluating energy storage systems for grid applications, Renew. Sust. Energ. Rev. 107 (2019) 516–534.
[8] M. Bottero, V. Ferretti, J.R. Figueira, S. Greco, B. Roy, Dealing with a multiple criteria environmental problem with interaction effects between criteria through an extension of the Electre III method, Eur. J. Oper. Res. 245 (2015) 837–850.
[9] M. Bottero, V. Ferretti, G. Mondini, Constructing multi-attribute value functions for sustainability assessment of urban projects, in: Computational Science and Its Applications – ICCSA 2014, Springer International Publishing, Cham, 2014, pp. 51–64.
[10] T. Buchholz, E. Rametsteiner, T.A. Volk, V.A. Luzadis, Multi criteria analysis for bioenergy systems assessments, Energy Policy 37 (2009) 484–495.
[11] S. Cajot, A. Mirakyan, A. Koch, F. Maréchal, Multicriteria decisions in urban energy system planning: A review, Front. Energy Res. 5 (2017), https://doi.org/10.3389/fenrg.2017.00010.
[12] G. Campanella, R.A. Ribeiro, A framework for dynamic multiple-criteria decision making, Decis. Support. Syst. 52 (2011) 52–60.
[13] J. Chai, E.W.T. Ngai, Decision-making techniques in supplier selection: recent accomplishments and what lies ahead, Expert Syst. Appl. 140 (2020), 112903.
[14] S. Chakraborty, E.K. Zavadskas, Applications of WASPAS method in manufacturing decision making, Informatica 25 (2014) 1–20.
[15] M. Cinelli, P. Burgherr, M. Kadziński, R. Słowiński, Case Studies Used in the Test of the MCDA-MSS for Energy Systems Analysis, Described According to its 156 Features, 2021, https://doi.org/10.5281/zenodo.5808839.
[16] M. Cinelli, P. Burgherr, M. Kadziński, R. Słowiński, Results from the Test of the MCDA-MSS with the 56 Case Studies from the Energy Systems Analysis Literature, 2021, https://doi.org/10.5281/zenodo.5808861.
[17] M. Cinelli, M. Kadziński, M. Gonzalez, R. Słowiński, How to support the application of multiple criteria decision analysis? Let us start with a comprehensive taxonomy, Omega 96 (2020), 102261.
[18] M. Cinelli, M. Kadziński, G. Miebs, M. Gonzalez, R. Słowiński, Recommending multiple criteria decision analysis methods with a new taxonomy-based decision support system, Eur. J. Oper. Res. 302 (2022) 633–651.
[19] R.M. Dawes, The robust beauty of improper linear models in decision making, Am. Psychol. 34 (1979) 571–582.
[20] D.F. de Lima Silva, A.T. de Almeida Filho, Sorting with TOPSIS through boundary and characteristic profiles, Comput. Ind. Eng. 141 (2020), 106328.
[21] J.G. de Paula Do Rosário, R. Salvador, M.V. Barros, C.M. Piekarski, L.M. da Luz, A. C. de Francisco, A Review on Multi-Criteria Decision Analysis in the Life Cycle Assessment of Electricity Generation Systems, International Business, Trade and Institutional Sustainability, 2020, pp. 575–590.
[22] S.L. Derby, R.L. Keeney, Risk analysis: understanding "How safe is safe enough?", Risk Anal. 1 (1981) 217–224.
[23] L.C. Dias, V. Mousseau, Eliciting multi-criteria preferences: ELECTRE models, in: L. C. Dias, A. Morton, J. Quigley (Eds.), Elicitation: The Science and Art of Structuring Judgement, Springer International Publishing, Cham, 2018, pp. 349–375.

[24] R.A. Estévez, V. Espinoza, R.D. Ponce Oliva, F. Vásquez-Lavín, S. Gelcich, Multi-criteria decision analysis for renewable energies: research trends, gaps and the challenge of improving participation, Sustainability 13 (2021) 3515.
[25] J.R. Figueira, V. Mousseau, B. Roy, ELECTRE methods, in: S. Greco, M. Ehrgott, J. R. Figueira (Eds.), Multiple Criteria Decision Analysis: State of the Art Surveys, Springer New York, New York, NY, 2016, pp. 155–185.
[26] D. Georgiou, E.S. Mohammed, S. Rozakis, Multi-criteria decision making on the energy supply configuration of autonomous desalination units, Renew. Energy 75 (2015) 459–467.
[27] K. Golić, V. Kosorić, A.K. Furundžić, General model of solar water heating system integration in residential building refurbishment—potential energy savings and environmental impact, Renew. Sust. Energ. Rev. 15 (2011) 1533–1544.
[28] S. Greco, M. Ehrgott, J. Figueira, Multiple Criteria Decision Analysis: State of the Art Surveys, Springer-Verlag, New York, 2016.
[29] S. Hirschberg, P. Burgherr, Sustainability assessment for energy technologies, in: Pages 1–22 Handbook of Clean Energy Systems, John Wiley & Sons, Ltd., 2015.
[30] B.F. Hobbs, G.T.F. Horn, Building public confidence in energy planning: a multimethod MCDM approach to demand-side planning at BC gas, Energy Policy 25 (1997) 357–375.
[31] J.P. Hubinont, Y. De Smet, Long-term multi-criteria improvement planning, Decis. Support. Syst. 149 (2021), 113606.
[32] A. Ishizaka, A. Labib, Review of the main developments in the analytic hierarchy process, Expert Syst. Appl. 38 (2011) 14336–14345.
[33] N. Itsubo, Weighting, in: M.Z. Hauschild, M.A.J. Huijbregts (Eds.), Life Cycle Impact Assessment, Springer, Netherlands, Dordrecht, 2015, pp. 301–330.
[34] D. Jun, F. Tian-tian, Y. Yi-sheng, M. Yu, Macro-site selection of wind/solar hybrid power station based on ELECTRE-II, Renew. Sust. Energ. Rev. 35 (2014) 194–204.
[35] İ. Kaya, M. Çolak, F. Terzi, Use of MCDM techniques for energy policy and decision-making problems: a review, Int. J. Energy Res. 42 (2018) 2344–2372.
[36] P.B. Keenan, P. Jankowski, Spatial decision support systems: three decades on, Decis. Support. Syst. 116 (2019) 64–76.
[37] L.R. Keeney, H. Raiffa, Decisions with Multiple Objectives: Preferences and Value Tradeoffs, Wiley, New York, 1976.
[38] L.R. Keeney, H. Raiffa, Decisions with Multiple Objectives: Preferences and Value Tradeoffs, Cambridge University Press, 1993.
[39] R.L. Keeney, R.S. Gregory, Selecting attributes to measure the achievement of objectives, Oper. Res. 53 (2005) 1–11.
[40] B. Kuleli Pak, Y.E. Albayrak, Y.C. Erensal, Renewable energy perspective for Turkey using sustainability indicators, Int. J. Comp. Intellig. Syst. 8 (2015) 187–197.
[41] A. Kumar, B. Sah, A.R. Singh, Y. Deng, X. He, P. Kumar, R.C. Bansal, A review of multi criteria decision making (MCDM) towards sustainable renewable energy development, Renew. Sust. Energ. Rev. 69 (2017) 596–609.
[42] M. Ligus, Evaluation of economic, social and environmental effects of low-emission energy technologies in Poland – multi-criteria analysis, Energy Procedia 136 (2017) 163–168.
[43] E. Loken, Use of multicriteria decision analysis methods for energy planning problems, Renew. Sust. Energ. Rev. 11 (2007) 1584–1595.
[44] C. Macharis, L. Turcksin, K. Lebeau, Multi actor multi criteria analysis (MAMCA) as a tool to support sustainable decisions: state of use, Decis. Support. Syst. 54 (2012) 610–620.
[45] A. Mardani, A. Jusoh, E. Zavadskas, F. Cavallaro, Z. Khalifah, Sustainable and renewable energy: an overview of the application of multiple criteria decision making techniques and approaches, Sustainability 7 (2015) 13947–13984.
[46] A. Mardani, E.K. Zavadskas, Z. Khalifah, N. Zakuan, A. Jusoh, K.M. Nor, M. Khoshnoudi, A review of multi-criteria decision-making applications to solve energy management problems: two decades from 1995 to 2015, Renew. Sust. Energ. Rev. 71 (2017) 216–256.
[47] J. Marichal, An axiomatic approach of the discrete Sugeno integral as a tool to aggregate interacting criteria in a qualitative framework, IEEE Trans. Fuzzy Syst. 9 (2001) 164–172.
[48] K. Marsh, P. Thokala, A. Mühlbacher, T. Lanitis, Incorporating preferences and priorities into MCDA: Selecting an appropriate scoring and weighting technique, in: K. Marsh, M. Goetghebeur, P. Thokala, R. Baltussen (Eds.), Multi-Criteria Decision Analysis to Support Healthcare Decisions, Springer International Publishing, Cham, 2017, pp. 47–66.
[49] M. Martín-Gamboa, L.C. Dias, P. Quinteiro, F. Freire, L. Arroja, A.C. Dias, Multi-criteria and life cycle assessment of wood-based bioenergy alternatives for residential heating: A sustainability analysis, Energies 12 (2019) 4391.
[50] P. Meyer, A.-L. Olteanu, Formalizing and solving the problem of clustering in MCDA, Eur. J. Oper. Res. 227 (2013) 494–502.
[51] M.M. Mousavi, J. Lin, The application of PROMETHEE multi-criteria decision aid in financial decision making: case of distress prediction models evaluation, Expert Syst. Appl. 159 (2020), 113438.
[52] G. Munda, The issue of consistency: Basic discrete multi-criteria "methods", in: Pages 85–109 Social Multi-Criteria Evaluation for a Sustainable Economy, Springer Berlin Heidelberg, Berlin, 2008.
[53] M. Nardo, M. Saisana, A. Saltelli, S. Tarantola, A. Hoffman, E. Giovannini, Handbook on Constructing Composite Indicators. Methodology and User Guide, OECD, Paris, 2008.
[54] J.H.P. Paelinck, Qualitative multiple criteria analysis, environmental protection and multiregional development, Pap. Reg. Sci. 36 (1976) 59–76.
[55] E. Pap, J.-L. Marichal, M. Grabisch, R. Mesiar, Properties for aggregation, in: E. Pap, J.-L. Marichal, M. Grabisch, R. Mesiar (Eds.), Aggregation Functions, Cambridge University Press, Cambridge, 2009, pp. 11–55.
[56] A. Papadopoulos, A. Karagiannidis, Application of the multi-criteria analysis method Electre III for the optimisation of decentralised energy systems, Omega 36 (2008) 766–776.
[57] D. Podgórski, Measuring operational performance of OSH management system – A demonstration of AHP-based selection of leading key performance indicators, Saf. Sci. 73 (2015) 146–166.
[58] H. Polatidis, D.A. Haralambopoulos, G. Munda, R. Vreeker, Selecting an appropriate multi-criteria decision analysis technique for renewable energy planning, Energy Sourc. Part B: Econ. Plan. Pol. 1 (2006) 181–193.
[59] J. Rezaei, Best-worst multi-criteria decision-making method, Omega 53 (2015) 49–57.
[60] P.D. Rigo, G. Rediske, C.B. Rosa, N.G. Gastaldo, L. Michels, A.L. Neuenfeldt Júnior, J.C.M. Siluk, Renewable energy problems: exploring the methods to support the decision-making process, Sustainability 12 (2020).
[61] B. Roy, Multicriteria Methodology for Decision Aiding, Kluwer Academic Publishers, 1996.
[62] B. Roy, Do not confuse, Int. J. Multicriteria Decis. Mak. 6 (2016) 112–117.
[63] B. Roy, R. Słowiński, Questions guiding the choice of a multicriteria decision aiding method, EURO J. Decis. Proc. 1 (2013) 69–97.
[64] T.L. Saaty, The Analytic Hierarchy Process, McGraw-Hill, New York, 1980.
[65] T.L. Saaty, The analytic hierarchy and analytic network processes for the measurement of intangible criteria and for decision-making, in: S. Greco, M. Ehrgott, J.R. Figueira (Eds.), Multiple Criteria Decision Analysis: State of the Art Surveys, Springer, New York, New York, NY, 2016, pp. 363–419.
[66] A.A. Salo, R.P. Hämäläinen, Preference assessment by imprecise ratio statements, Oper. Res. 40 (1992) 1053–1061.
[67] J.R. San Cristóbal, Multi-criteria decision-making in the selection of a renewable energy project in Spain: the Vikor method, Renew. Energy 36 (2011) 498–502.
[68] J.M. Sánchez-Lozano, M.S. García-Cascales, M.T. Lamata, Comparative TOPSIS-ELECTRE TRI methods for optimal sites for photovoltaic solar farms. Case study in Spain, J. Clean. Prod. 127 (2016) 387–398.
[69] W. Schenler, S. Hirschberg, P. Burgherr, M. Makowski, J. Granat, Final Report on Sustainability Assessment of Advanced Electricity Supply Options - NEEDS Deliverable D10.2 - RS2b, Accessed on 15 June 2021 at, https://www.psi.ch/en/ta/projects/new-energy-externalities-developments-for-sustainability-needs, 2009.
[70] M. Shao, Z. Han, J. Sun, C. Xiao, S. Zhang, Y. Zhao, A review of multi-criteria decision making applications for renewable energy site selection, Renew. Energy 157 (2020) 377–403.
[71] G. Villacreses, G. Gaona, J. Martínez-Gómez, D.J. Jijón, Wind farms suitability location using geographical information system (GIS), based on multi-criteria decision making (MCDM) methods: the case of continental Ecuador, Renew. Energy 109 (2017) 275–286.
[72] H. Voogd, Multicriteria evaluation with mixed qualitative and quantitative data, Environ. Plan. B: Plan. Design 9 (1982) 221–236.
[73] J.-J. Wang, Y.-Y. Jing, C.-F. Zhang, J.-H. Zhao, Review on multi-criteria decision analysis aid in sustainable energy decision-making, Renew. Sust. Energ. Rev. 13 (2009) 2263–2278.
[74] Q. Wang, K.L. Poh, A survey of integrated decision analysis in energy and environmental modeling, Energy 77 (2014) 691–702.
[75] J. Wątróbski, J. Jankowski, P. Ziemba, A. Karczmarczyk, M. Zioło, Generalised framework for multi-criteria method selection, Omega 86 (2019) 107–124.
[76] W.C. Wedley, Issues in aggregating AHP/ANP scales, in: Advances in Intelligent Decision Technologies, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010, pp. 29–42.
[77] C. Wu, Y. Lin, D. Barnes, An integrated decision-making approach for sustainable supplier selection in the chemical industry, Expert Syst. Appl. 184 (2021), 115553.
[78] A. Yazdani-Chamzini, M.M. Fouladgar, E.K. Zavadskas, S.H.H. Moini, Selecting the optimal renewable energy using multi criteria decision making, J. Bus. Econ. Manag. 14 (2013) 957–978.
[79] N.H. Zardari, K. Ahmed, S.M. Shirazi, Z.B. Yusop, Literature Review, in: N. H. Zardari, K. Ahmed, S.M. Shirazi, Z.B. Yusop (Eds.), Weighting Methods and their Effects on Multi-Criteria Decision Making Model Outcomes in Water Resources Management, Springer International Publishing, Cham, 2015, pp. 7–67.

Marco Cinelli is an Assistant Professor at Leiden University College (LUC), where he leads the interdisciplinary laboratory Decision Engineering for SustaInability and REsilience (DESIRE).

His research is focused on the development of Decision Support Systems (DSSs) with Multiple Criteria Decision Analysis (MCDA). The main frameworks he uses include sustainability, resilience and risk assessment, while the main application areas are energy systems and technologies assessment, materials and processes development, and chemical safety. Before joining LUC, between September 2018 and September 2021, Marco was a Marie Skłodowska-Curie Global Fellow at (i) the Laboratory for Energy Systems Analysis, Paul Scherrer Institute, Villigen PSI, Switzerland, (ii) the Center for Environmental Solutions and Emergency Response (CESER) (formerly National Risk Management Research Laboratory (NRMRL)) of the U.S. Environmental Protection Agency in Cincinnati, USA and (iii) the Institute of Computing Science at Poznan University of Technology, Poland. Marco received his Ph.D. in Engineering from the University of Warwick (UK) in 2016.

Miłosz Kadziński is an Associate Professor at Poznan University of Technology and a Polish Young Academy member. He holds the MCDM Doctoral Dissertation Award 2013. He has been distinguished with the scientific awards by EURO, INFORMS MCDM Section, Polish Academy of Sciences, and Ministry of Science. He has been acknowledged with thirteen Best Reviewer Awards by EJOR, Omega, and GDN. He acts as a co-president of the Decision Deck Consortium. His research interests include preference learning, robustness analysis, and strengthening the interfaces between MCDM and other sub-disciplines of algorithmic decision theory, such as interactive and evolutionary multiple objective optimization, data envelopment analysis, and machine learning. He has published more than 65 papers in premier international journals such as Omega, EJOR, DSS, INS, COR, KBS, ESWA, and ML.

Peter Burgherr joined the Paul Scherrer Institut (PSI) as a risk analyst in 2001. Since August 2008 he leads the interdisciplinary Technology Assessment group of the Laboratory for Energy Systems Analysis, which carries out integrated assessments of energy systems to support the complex decision-making processes towards a sustainable energy future.

His primary research interest is the comparative analysis of accident risks in the energy sector, and its relevance in the broader context of sustainability, resilience, energy security and critical infrastructure protection. He is also the primary responsible for PSI's database Energy-Related Severe Accident Database (ENSAD), which is the world's largest database on severe accidents in the energy sector. He initiated the development of the Energy Infrastructure Attack Database together with the Center for Security Studies at ETHZ, and subsequent collaborations with the Energy Policy Institute Chicago, Harris School of Public Policy, University Chicago. Furthermore, he is strongly involved in the sustainability assessment of current and future energy technologies, including their evaluation with Multi-Criteria Decision Aiding techniques.

He regularly presents his work at international conferences and meetings, publishing in peer-reviewed scientific journals, contributing to books, and various other publications. As part of his duties, he is also engaged in the supervision of students at different levels (PhD, MSc, BSc, interns), and lectured at the Swiss Federal Institute of Technology, University of Geneva, Zurich University of Applied Sciences, Lucerne University of Applied Sciences and Arts, NATO School Oberammergau, and KINGS (Busan, South Korea).

He holds a doctoral degree in environmental sciences from ETH Zurich, and MSc and BSc degrees in biology from ETH Zurich.

Roman Słowiński is a Professor and Founding Chair of the Laboratory of Intelligent Decision Support Systems at Poznań University of Technology, and a Professor in the Systems Research Institute of the Polish Academy of Sciences. As an ordinary member of the Polish Academy of Sciences he is its Vice President, elected for the term 2019-2022. He is a member of Academia Europaea and Fellow of IEEE, IRSS, INFORMS and IFIP. In his research, he combines Operational Research and Artificial Intelligence for Decision Aiding. Recipient of the EURO Gold Medal by the European Association of Operational Research Societies (1991), and Doctor *HC* of Polytechnic Faculty of Mons (Belgium, 2000), University Paris Dauphine (France, 2001), and Technical University of Crete (Greece, 2008). In 2005 he received the Annual Prize of the Foundation for Polish Science - the highest scientific honor awarded in Poland, and in 2020 - the Scientific Award of the Prime Minister of Poland. Since 1999, he is the principal editor of the *European Journal of Operational Research* (Elsevier), a premier journal in Operational Research.