Best-worst Tradeoff method

Liang, Fuqi; Brunelli, Matteo; Rezaei, Jafar

DOI
10.1016/j.ins.2022.07.097

Publication date
2022

Document Version
Final published version

Published in
Information Sciences

Information Sciences 610 (2022) 957–976

Contents lists available at ScienceDirect

Information Sciences

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

Best-worst Tradeoff method

Fuqi Liang [a,b,*], Matteo Brunelli [c], Jafar Rezaei [b]

[a] School of Management, Zhejiang University, 310058 Hangzhou, China
[b] Faculty of Technology, Policy and Management, Delft University of Technology, 2628 BX Delft, the Netherlands
[c] Department of Industrial Engineering, University of Trento, Via Sommarive 9, 38123 Trento, Italy

ARTICLE INFO

Article history:
Received 14 July 2021
Received in revised form 16 May 2022
Accepted 17 July 2022
Available online 22 July 2022

Keywords:
Multi-attribute analysis
Best-Worst Method
Tradeoff procedure
Multi-Attribute Value Theory
Consistency

ABSTRACT

This study aims to develop a Multi-Attribute Decision-Making (MADM) method, the Best-Worst Tradeoff method, which draws on the underlying principles of two popular MADM methods (the Best-Worst Method (BWM) and the Tradeoff). The traditional Tradeoff procedure, which is based on the axiomatic foundation of multi-attribute value theory, considers the ranges of the attributes, but decision-makers/analysts find it hard to check the consistency of the paired comparisons when using this method. The traditional BWM, on the other hand, uses two opposite references (best and worst) in a single optimization, which not only frames the elicitation process in a more structured way, but helps decision-makers/analysts check the consistency. However, the BWM does not explicitly considers the attributes ranges in the pairwise comparisons. The method proposed in this study uses the "consider-the-opposite-strategy" and accounts for the range effect simultaneously. Specifically, the decision-maker considers the ranges of the attributes and provide two pairwise comparison vectors, then an optimization model is designed to determine the optimal weights of the attributes based on these two vectors. After that, consistency thresholds are constructed to check the consistency of the judgements. Finally, a case study is used to examine the feasibility of the proposed method.

1. Introduction

Nowadays, an increasing number of decisions are made in complex contexts in a host of different application domains. This ought to be supported by mathematically sound decision analysis methodologies. A number of these methodologies can be classified as multi-attribute decision-making (MADM), thanks to their capacity to handle problems where a multitude of, often conflicting, objectives arise [2]. The common thread of these methods is the representation of the final value of each alternative, as a function of the degrees to which the same alternative satisfies a number of attributes, where each attribute level approximates the level of achievement of one of the objectives. In this context, it is often important to quantify the contribution of different attributes by means of weights (scaling constants) to aggregate the performances of alternatives with respect to the attributes into single values.

Various methods have been proposed to elicit the weights of attributes [3,4] on the ground of the subjective preferences of experts. Some of the most popular methods are: the Analytic Hierarchy Process (AHP) [5], Simple Multi-Attribute Rating Technique using Swings (SMARTS) [6], Direct Rating method [7], Swing [8], the Best-Worst Method (BWM) [9,10], and the Tradeoff procedure [11]. In this study, we focus on the Tradeoff and BWM methods.

* Corresponding author at; School of Management, Zhejiang University, 310058 Hangzhou, China.
E-mail address: fuqi15@zju.edu.cn (F. Liang).

https://doi.org/10.1016/j.ins.2022.07.097
0020-0255/

One of the obvious shortcomings of some methods is that the weight elicitation phase is developed *a priori*, on the basis of the perceived importance of the attributes alone. Conversely, in decision analysis, the weights of the attributes (or scaling constants) should be sensitive to the range of each attribute, i.e., if alternatives are very close to each other with respect to a particular attribute, that attribute would play a small role in discriminating between alternatives [12,13]. Methods that cannot account for the range of attributes values may lead to errors in the estimation of weights [13–15]. According to previous studies [12,13], even if the range of the attribute is mentioned, DMs often do not adjust their judgements on the weights properly, which means that methods that do not consider the ranges, like simple ranking or direct rating methods, should only be used with great care [12,13]. Although BWM encourages DMs to consider the range of attributes in advance, in practice, this is not done systematically. In this sense, methods like Swing [8] and Tradeoff procedure [11] which require DMs to provide their preference based on the range of attributes could handle this problem better than BWM could. How to take the range of attributes into consideration in BWM in an explicit and systematic way to avoid distortion or biases, is an important issue that requires further investigation.

The idea of taking the range effect into account, like the Tradeoff procedure and Swing do, can be incorporated into the BWM to remedy the potential distortion and biases. Compared to the Tradeoff procedure, Swing cannot make the consistency check, which is a serious shortcoming. As such, the Tradeoff procedure could be a better choice compared with Swing when considering external validity (the evaluation of consistency of the given preferences) [3]. However, the consistency check in the Tradeoff procedure may also be problematic. Keeney and Raiffa [11] encouraged analysts to use the Tradeoff procedure to ask additional questions to increase the robustness of the results and identify possible inconsistencies. But how many and what additional questions should be asked remains unclear. In addition, analysts frequently fail to apply consistency checks when assessing value tradeoffs [16]. Although some researchers have tried to improve the Tradeoff procedure - one of the latest studies uses a flexible and interactive way to collect trade-off information from the DM [17] - the consistency estimation problem has not attracted enough attention.

By incorporating the philosophy of BWM, using two vectors of pairwise comparisons based on two opposite references (best and worst) within a single optimization model, our goal is to help mitigate the anchoring bias. This strategy, named "consider-the-opposite strategy", was initially developed by Bacon [18] and has since been used in many psychological studies. Developing a parsimonious MADM method that incorporates the "consider-the-opposite strategy that can check the consistency systematically as well as consider the range effect based on the axiomatic foundation of Multi-Attribute Value Theory (MAVT) is the main objective and contribution of this study. To achieve this, the underlying principles of the Tradeoff procedure [11] and the BWM [9,10] are adopted. More specifically, firstly, MAVT is used as a foundation to establish a tradeoff for the objectives based on the range of the attributes. Secondly, two vectors of pairwise comparisons, Best-to-Others (BO) and Others-to-Worst (OW), are constructed to avoid that the DM is only anchoring on a fixed value or reference point, the way it is done in SMARTS and Swing. Thirdly, based on the BO and OW obtained from the DM, an optimization model is proposed to derive the weights of attributes. In addition, a cardinal consistency index and an ordinal consistency index are proposed to estimate the extent to which the preferences of a DM deviate from the perfect consistency, and thresholds will be proposed to decide whether or not the deviation is acceptable.

The remainder of the paper is organized as follows: In Section 2, the basic knowledge of MAVT and the procedure of the classical Tradeoff procedure are reviewed, while, in Section 3, the procedure of the Best–Worst Tradeoff (BWT) method is illustrated. Section 4 focuses on the consistency-checking, which includes the proposed consistency ratios and the thresholds for the BWT. A case study is used to illustrate the proposed method in Section 5, while the features of BWM are discussed in Section 6 and some conclusions are presented in Section 7.

2. Preliminaries

Of the many methods and theories developed to support MADM processes, MAVT is one of the most widely used, as well as the one for which Keeney and Raiffa [11] proposed the well-known Tradeoff procedure. In the following sub-sections, MAVT and the classical Tradeoff procedure are discussed, after which the recently developed BWM method is introduced.

2.1. Multi-attribute value theory and the additive value function

A typical MADM problem consists of a non-empty finite set of m alternatives $A = \{A_1, A_2, \cdots, A_m\}$ and a set of objectives that represent the goals of the DM. The satisfaction of the objectives is assumed to depend on a finite set of n attributes (also known as criteria) $C = \{C_1, C_2, \cdots, C_n\}$. The set of possible levels that can be achieved by a generic alternative with respect to the j th attribute is denoted by X_j. Assuming that all the relevant attributes have been considered, each alternative can be associated with a *consequence* $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in X_1 \times X_2 \times \cdots \times X_n$, such that x_j indicates the level of the j th attribute achieved by the alternative.

Given the fact that, to our scope, an alternative can be fully described by its consequence, for the sake of simplicity, in the following section, we will consider consequences, $\mathbf{x}$, instead of alternatives. Moreover, according to MAVT, each attribute is rescaled and normalized into the interval $[0, 1]$ thanks to a function $v_j : X_j \rightarrow [0, 1]$. In fact, it is customary to assign values 0 and 1 to the least $(\underline{x}_j)$ and most $(\bar{x}_j)$ desirable attribute levels, respectively. Hence, $v_j(x_j) \in [0, 1]$ can be interpreted as the value of the consequence evaluated with respect to the j th attribute. At this point, taking into account that the value of a consequence is a function of its attribute values, the main problem in MAVT is to find a function $V : [0,1]^n \to [0,1]$, which can correctly aggregate the n values $v_j(x_j)$ and represent the preferences of a DM.

The search for a simple and easily interpretable form for the function V is simplified by some known results stating that, under mild assumptions (i.e., mutual preference independence and measurability), the function V is additive. Namely,

$$V(v(\mathbf{x})) = V(v_1(x_1), \cdots, v_n(x_n)) = \sum_{j=1}^{n} w_j v_j(x_j) \tag{1}$$

with $w_j \geq 0, \forall j$ and $w_1 + w_2 + \cdots + w_n = 1$.

At this point, what is left to do is find suitable values for the weights w_j. Among the various methods used to elicit weights, the Tradeoff procedure has a strong axiomatic foundation [19].

2.2. The classical Tradeoff procedure

The original Tradeoff procedure essentially consists of three parts. The goal of the first part is to obtain the preference relation of the attributes. Attributes are considered in pairs, and in each pair two hypothetical consequences are constructed and presented to DM for tradeoffs. These two consequences are only different with regard to the performance of the two attributes under consideration, and the performance of the other attributes is set to the worst level. In the first hypothetical consequence, the performance of the two attributes is set to their worst and best levels, respectively, and in the second consequence, they are set the other way around. The DM is asked to indicate which consequence is preferred. After having applied these pairwise comparisons to all the attributes, the relation of order on the set of weights can be obtained.

The second part is designed to obtain indifference relations. The DM still uses the hypothetical consequences created in the first part, but now, the DM is asked to manipulate the level of the more important attribute until indifference is reached. By iterating this procedure for a properly chosen set of $(n-1)$ comparisons between pairs of artificially constructed payoffs, indifference relations on these hypothetical paired consequences can be obtained.

The third part is intended to determine the weights. On the basis of the indifference relations obtained in the second part, together with the constraint $w_1 + w_2 + \cdots + w_n = 1$, the analyst can form a system of linear equations for which a unique set of weights exists and can be identified.

Despite the simplicity of the original Tradeoff procedure, Keeney and Raiffa [11] questioned its robustness and consistency, by saying that “it may be desirable to ask additional questions thereby getting an over-determined system of equations”. The same was argued by Eisenführ et al. [20], who wrote that “it is sensible not to limit ourselves to the determination of $(n-1)$ tradeoffs”. At this point, we face significant questions like: “how many additional tradeoffs should we assess?”, and “which pairs should we choose to compare?” As such, it makes sense to propose a more structured and justified methods to deal with redundancy of preferences and their inconsistencies, which is the aim of this study.

2.3. The original Best-Worst method

The original BWM uses ratios of the relative importance of attributes in pairs estimated by a DM, from the two opposite anchored vectors, A^{BO} and A^{OW}. The basic steps of the original BWM can be summarized as follows [9,10]:

Step 1. The set of attributes $\{C_1, C_2, \cdots, C_n\}$ is determined by the DM.

Step 2. The best (e.g. the most influential or important) and worst (e.g. the least influential or important) attributes are determined by the DM. The two attributes are shown by C_B and C_W, respectively.

Step 3. The preference of the best over all the other attributes is determined by the DM using a number from $\{1, 2, \cdots, 9\}$. The obtained Best-to-Others vector is: $A^{BO} = (a_{B1}, a_{B2}, \cdots, a_{Bn})$, where a_{Bj} represents the preference of the best attribute C_B over attribute C_j, $j = 1, 2, \cdots, n$.

Step 4. The preferences of all the attributes over the worst attribute are determined by the DM using a number from $\{1, 2, \cdots, 9\}$. The obtained Others-to-Worst vector is: $A^{OW} = (a_{1W}, a_{2W}, \cdots, a_{nW})$, where a_{jW} represents the preference of attribute C_j over the worst attribute C_W, $j = 1, 2, \cdots, n$.

Step 5. The weights $(w_1^*, w_2^*, \cdots, w_n^*)$ can be calculated in this step. The optimal weights for the attributes are determined by setting the conditions where, for each pair of w_B/w_j and w_j/w_W, $w_B/w_j = a_{Bj}$ and $w_j/w_W = a_{jW}$. To find a good approximation for all j, a solution in which the maximum absolute differences $\left|\frac{w_B}{w_j} - a_{Bj}\right|$ and $\left|\frac{w_j}{w_W} - a_{jW}\right|$ for all j are minimized is formulated in the following model:

$$\begin{aligned} &\text{minimize} \quad \max_j \left\{ \left|\frac{w_B}{w_j} - a_{Bj}\right|, \left|\frac{w_j}{w_W} - a_{jW}\right| \right\} \\ &\text{subject to} \quad w_1 + w_2 + \cdots + w_n = 1 \\ &\qquad\qquad\quad w_j > 0, \ j = 1, 2, \cdots, n \end{aligned} \tag{2}$$

In words, the goal of this optimization problem is to find the weight vector that minimizes the maximum discrepancy between the judgements and their theoretical valued obtained as ratios between their corresponding weights.

The classical Tradeoff procedure has the strong axiomatic foundation of MAVT, and explicitly takes the ranges of the attributes into consideration, but it is difficult to check the consistency of the given preferences when the DMs/analysts using this method. Contrarily, the inherent property of the original BWM enables DMs/analysts check the consistency, but it does not explicitly consider the ranges of the attributes in the pairwise comparisons. Therefore, a proposal that combines the two merits, consistency check and range sensitivity, is requested.

3. Best-Worst Tradeoff method (BWT)

Based on the concepts of the traditional Tradeoff procedure and of the Best-Worst Method, a structured method, called Best-Worst Tradeoff method (BWT), is proposed in this section to obtain the weights of attributes and help check the consistency via a structured framework.

Step 1. Determine alternatives and attributes.

We assume that m alternatives $A = \{A_1, A_2, \cdots, A_m\}$, and n attributes $C = \{C_1, C_2, \cdots, C_n\}$ have been determined by DM. Moreover, a different consequence vector $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ is assigned to each alternative.

Step 2. Determine value function for each attribute.

There are many alternative methods to elicit the attribute value functions v_j [2,11,21,22]. One of the most widely used methods is the mid-value splitting technique proposed by Keeney and Raiffa [11].

Step 3. Identify Best and Worst attributes.

Firstly, the DM needs to identify the "best" attribute C_B and the "worst" C_W. Hereafter, we will use the terms "best" and "worst", borrowed from the classical BWM, to indicate the attributes with the greatest and smallest weights, respectively. These two attributes will serve as the two yardsticks against which the other attributes will be compared, with the ultimate goal of avoiding the anchoring bias. According to the first part of the original Tradeoff procedure in Section 2.2, we need to create n hypothetical consequences $\mathbf{x}^j$ $(j = 1, 2, \cdots, n)$ to represent the best performance of the j th attribute and the simultaneous worst performance of the other attributes (see Fig. 1). The DM is asked to compare and rank the hypothetical consequences $\mathbf{x}^1, \cdots, \mathbf{x}^j, \cdots, \mathbf{x}^n$, so that C_B and C_W can be identified.

Step 4. Compare Best to others tradeoff.

Assume that we are interested in comparing the best attribute to all the other attributes. Let us consider the instance of the comparison of the best attribute C_B and the k th one C_k. For this scope we will need two auxiliary consequences: $\mathbf{x}^{B,k}$ and $\mathbf{x}^k$. The first consequence is defined such that all attributes, except attribute C_B, achieve the lowest levels ($v_j(\underline{x}_j) = 0$), and the level of the best attribute is left to be determined. Assuming that C_k is the "other" attribute to be compared to the best, then the second consequence $\mathbf{x}^k$ has all the components at the lowest level of satisfaction ($v_j(\underline{x}_j) = 0$), except for the k th, which, instead, has the highest ($v_j(\bar{x}_k) = 1$). Now, we wonder what degree of satisfaction of the best attribute in $\mathbf{x}^{B,k}$ would make the two consequences equally desirable, i.e., the value of $x_B^{B,k}$ for which,

$$\left(\underline{x}_1, \cdots, x_B^{B,k}, \cdots, \underline{x}_n\right) \sim (\underline{x}_1, \cdots, \bar{x}_k, \cdots, \underline{x}_n) \iff \mathbf{x}^{B,k} \sim \mathbf{x}^k \iff V\left(v\left(\mathbf{x}^{B,k}\right)\right) = V\left(v\left(\mathbf{x}^k\right)\right), \tag{3}$$

where '$\sim$' indicates a relation of indifference between the two consequences.

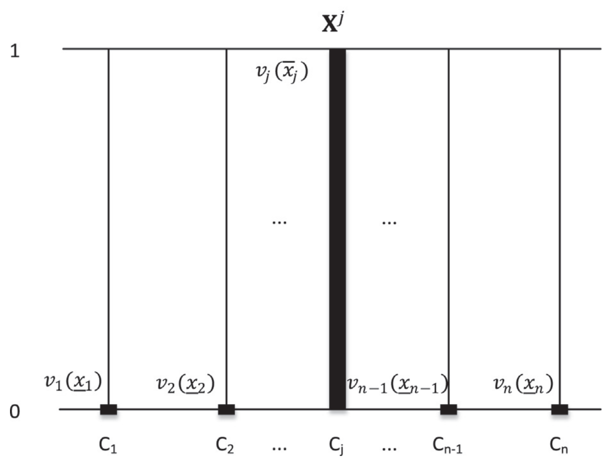

Fig. 1. Hypothetical consequence $\mathbf{x}^j$.

Formally, assuming that we can use the additive representation, this corresponds to.

$$V\left(v\left(\mathbf{x}^{B,k}\right)\right) = w_B v_B\left(x_B^{B,k}\right) + \sum_{\substack{j=1 \\ j\neq B}}^{n} w_j \underbrace{v_j\left(\underline{x}_j\right)}_{=0} = w_k \underbrace{v_k\left(\bar{x}_k\right)}_{=1} + \sum_{\substack{j=1 \\ j\neq k}}^{n} w_j \underbrace{v_j\left(\underline{x}_j\right)}_{=0} = V\left(v\left(\mathbf{x}^k\right)\right)$$

which collapses into:

$$w_B v_B\left(x_B^{B,k}\right) = w_k \tag{4}$$

At this point, asking for which value of $x_B^{B,k} \in X_B$, we obtain $\mathbf{x}^{B,k} \sim \mathbf{x}^k$ is equivalent to asking which $x_B^{B,k} \in X_B$ satisfies equation (4). The value $v_B\left(x_B^{B,k}\right)$, which we denote as a_{kB}, is thus the DM's estimation of the ratio w_k/w_B. Now, we use the vector $(a_{1B}, a_{2B}, \cdots, a_{nB})$ to collect the pairwise comparison values of $\mathbf{x}^{B,k}$ to $\mathbf{x}^k$ for all the k. Of course, its reciprocal $a_{Bk} = 1/a_{kB}$ corresponds to the DM's estimate of the ratio w_B/w_k. We use $A^{BO} = (a_{B1}, a_{B2}, \cdots, a_{Bn})$ to indicate the Best-to-Others vector. The value a_{Bk} has a double interpretation as both the following conditions are equivalent and should, in theory, hold:

$$w_k a_{Bk} = w_B \tag{5}$$

$$a_{Bk} = \frac{w_B}{w_k} \tag{6}$$

According to Equation (5), a_{Bk} is the scalar to which one needs to multiply w_k to get w_B, while Equation (6) stipulates that a_{Bk} is the ratio between the two weights w_B and w_k.

By considering the interpretation (5), if we consider all the $k\neq B$, we obtain the system of $(n-1)$ linear equations $w_B = a_{Bk}w_k, \forall k\neq B$. Similarly, we can obtain an equivalent system of equations (this time non-linear) if we consider the interpretation suggested by (6).

Example 1. *Suppose a DM, who is going to buy a car, considers four attributes (C_1 : price, C_2 : quality, C_3 : safety, C_4 : style) to evaluate the alternatives. We further assume that C_1 is identified by the DM as the best attribute. In order to make the tradeoffs for the attributes (C_1 to C_2; C_1 to C_3; C_1 to C_4), we need six hypothetical consequences:* $\mathbf{x}^{1,2} = \left(x_1^{1,2}, \underline{x}_2, \underline{x}_3, \underline{x}_4\right)$, $\mathbf{x}^{1,3} = \left(x_1^{1,3}, \underline{x}_2, \underline{x}_3, \underline{x}_4\right)$, $\mathbf{x}^{1,4} = \left(x_1^{1,4}, \underline{x}_2, \underline{x}_3, \underline{x}_4\right)$, $\mathbf{x}^2 = (\underline{x}_1, \bar{x}_2, \underline{x}_3, \underline{x}_4)$, $\mathbf{x}^3 = (\underline{x}_1, \underline{x}_2, \bar{x}_3, \underline{x}_4)$, $\mathbf{x}^4 = (\underline{x}_1, \underline{x}_2, \underline{x}_3, \bar{x}_4)$, *where $\underline{x}_j$ and $\bar{x}_j$ represent the least and most desirable attribute levels, respectively. Then, the decision maker is asked to find the values of $x_1^{1,2}$, $x_1^{1,3}$, and $x_1^{1,4}$ such that the following three indifference relations hold*:

$$\mathbf{x}^{1,2} \sim \mathbf{x}^2 \iff \left(x_1^{1,2}, \underline{x}_2, \underline{x}_3, \underline{x}_4\right) \sim (\underline{x}_1, \bar{x}_2, \underline{x}_3, \underline{x}_4)$$

$$\mathbf{x}^{1,3} \sim \mathbf{x}^3 \iff \left(x_1^{1,3}, \underline{x}_2, \underline{x}_3, \underline{x}_4\right) \sim (\underline{x}_1, \underline{x}_2, \bar{x}_3, \underline{x}_4)$$

$$\mathbf{x}^{1,4} \sim \mathbf{x}^4 \iff \left(x_1^{1,4}, \underline{x}_2, \underline{x}_3, \underline{x}_4\right) \sim (\underline{x}_1, \underline{x}_2, \underline{x}_3, \bar{x}_4)$$

These relations can be illustrated in Fig. 2 ($\mathbf{x}^{1,2} \sim \mathbf{x}^2$), Fig. 3 ($\mathbf{x}^{1,3} \sim \mathbf{x}^3$) and Fig. 4 ($\mathbf{x}^{1,4} \sim \mathbf{x}^4$).

After having determined the value functions and $x_1^{1,2}$, $x_1^{1,3}$, $x_1^{1,4}$, we can obtain the relations:

$w_1 v_1\left(x_1^{1,2}\right) = w_2$, $w_1 v_1\left(x_1^{1,3}\right) = w_3$, $w_1 v_1\left(x_1^{1,4}\right) = w_4$.

As we know that the first criterion is the best, with the convention, $a_{B2} = 1/v_1\left(x_1^{1,2}\right)$, $a_{B3} = 1/v_1\left(x_1^{1,3}\right)$, $a_{B4} = 1/v_1\left(x_1^{1,4}\right)$, one obtains:

$$w_1 = w_2 a_{B2} = w_3 a_{B3} = w_4 a_{B4}$$

Step 5. Compare Others to Worst tradeoff.

Similarly, the entire procedure can be repeated the other way round, this time to compare each attribute C_k ($C_k\neq C_W$) to C_W, the worst attribute. In this case, we continue to use two auxiliary consequences, $\mathbf{x}^W$, and $\mathbf{x}^{k,W}$, but with different components. Now, $\mathbf{x}^W$ has all components at the lowest level, i.e., $v_j(\underline{x}_j) = 0$, except for the worst attribute that reaches the highest level ($v_W(\bar{x}_W) = 1$). Consequence $\mathbf{x}^{k,W}$, instead, has all the components at the lowest level, i.e., $v_j(\underline{x}_j) = 0$, except for the k $^{\text{th}}$ attribute that is left undetermined. When we assume $\mathbf{x}^W \sim \mathbf{x}^{k,W}$, i.e.,:

$$V(v_1(\underline{x}_1), \cdots, v_W(\bar{x}_W), \cdots, v_n(\underline{x}_n)) = V(v_1(\underline{x}_1), \cdots, v_k(x_k), \cdots, v_n(\underline{x}_n)) \tag{7}$$

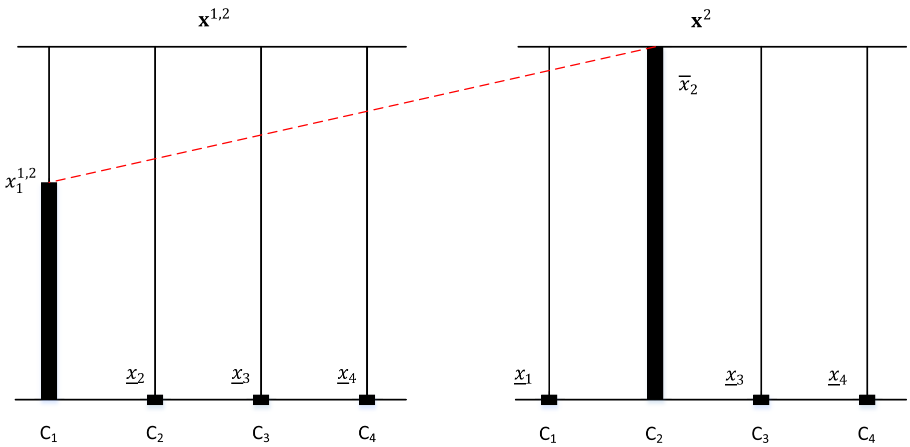

Fig. 2. The DM states the level of the attribute C_1 which makes the consequences $\mathbf{x}^{1,2}$ and $\mathbf{x}^2$ indifferent one to another.

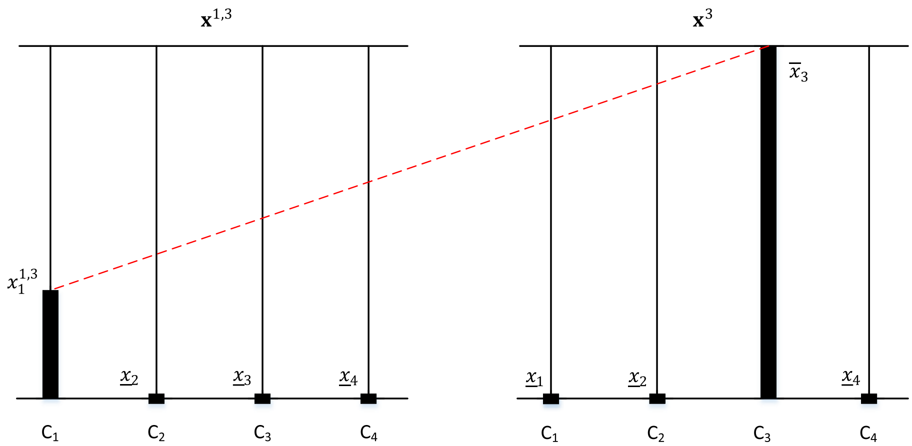

Fig. 3. The DM states the level of the attribute C_1 which makes the consequences $\mathbf{x}^{1,3}$ and $\mathbf{x}^3$ indifferent one to another.

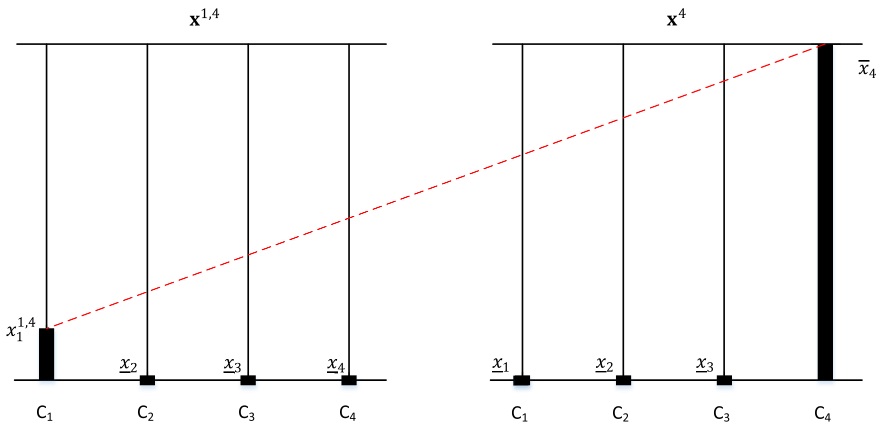

Fig. 4. The DM states the level of the attribute C_1 which makes the consequences $\mathbf{x}^{1,4}$ and $\mathbf{x}^4$ indifferent one to another.

and use the additive value function, we obtain:

$$w_W \underbrace{v_W(\bar{x}_W)}_{=1} + \sum_{\substack{j=1 \\ j\neq W}}^{n} w_j \underbrace{v_j(\underline{x}_j)}_{=0} = w_k v_k(x_k) + \sum_{\substack{j=1 \\ j\neq k}}^{n} w_j \underbrace{v_j(\underline{x}_j)}_{=0}$$

which can be simplified into:

$$w_W = w_k v_k(x_k) \tag{8}$$

By asking the DM what value of x_k makes the two consequences indifferent, we obtain $v_k(x_k)$, which we will call a_{Wk}. From this, thanks to reciprocity, i.e., $a_{kW} = 1/a_{Wk}$, all the values of the comparisons form the Others-to-Worst vector $A^{OW} = (a_{kW})$, where $k = 1, 2, \cdots, n$. We recover the following two interpretations for the value a_{kW},

$$w_W a_{kW} = w_k \tag{9}$$

$$a_{kW} = w_k / w_W \tag{10}$$

Implementing this procedure for all the paired comparison values a_{kW} of $\mathbf{x}^W$ to $\mathbf{x}^k$, we obtain the system of $(n-1)$ linear equations $w_k = a_{kW} w_W$, $\forall k \neq W$. Similarly, we can obtain an equivalent system of equations (this time non-linear) if we consider the interpretation suggested by (10).

Example 2. *We continue with Example 1, and suppose C_4 is identified by the DM as the worst attribute. To make the tradeoffs between the attributes (C_1 to C_4; C_2 to C_4; C_3 to C_4), we need to have 4 hypothetical consequences:* $\mathbf{x}^{1,4} = (x_1, \underline{x}_2, \underline{x}_3, \underline{x}_4)$, $\mathbf{x}^{2,4} = (\underline{x}_1, x_2, \underline{x}_3, \underline{x}_4)$, $\mathbf{x}^{3,4} = (\underline{x}_1, \underline{x}_2, x_3, \underline{x}_4)$, $\mathbf{x}^4 = (\underline{x}_1, \underline{x}_2, \underline{x}_3, \bar{x}_4)$, *where $\underline{x}_j$ and $\bar{x}_j$ represent the least and most desirable attribute levels, respectively. At this point the decision maker is asked to find the values of x_1, x_2, and x_3 such that the following indifference relations hold:*

$$\mathbf{x}^{1,4} \sim \mathbf{x}^4 \iff (x_1, \underline{x}_2, \underline{x}_3, \underline{x}_4) \sim (\underline{x}_1, \underline{x}_2, \underline{x}_3, \bar{x}_4)$$

$$\mathbf{x}^{2,4} \sim \mathbf{x}^4 \iff (\underline{x}_1, x_2, \underline{x}_3, \underline{x}_4) \sim (\underline{x}_1, \underline{x}_2, \underline{x}_3, \bar{x}_4)$$

$$\mathbf{x}^{3,4} \sim \mathbf{x}^4 \iff (\underline{x}_1, \underline{x}_2, x_3, \underline{x}_4) \sim (\underline{x}_1, \underline{x}_2, \underline{x}_3, \bar{x}_4)$$

These relations can be simplified as $\mathbf{x}^{1,4} \sim \mathbf{x}^{2,4} \sim \mathbf{x}^{3,4} \sim \mathbf{x}^4$, and represented in Fig. 5.

After having determined the values of x_1, x_2, x_3, we can obtain the relations:

$$w_1 v_1(x_1) = w_2 v_2(x_2) = w_3 v_3(x_3) = w_4 \underbrace{v_4(\bar{x}_4)}_{=1}$$

Remark. In our examples, the tradeoffs were made between fictitious alternatives where some attributes were set to their minimum levels. See, for instance, Fig. 5. This suits the results obtained by Vetschera et al. [23] according to which alternatives with extreme values may improve the consistency of results. However, this should not be interpreted as binding. Values of irrelevant attributes in tradeoffs can be set to values other than their minimums, as long as the same values are present in both consequences and can therefore be canceled out.

Step 6. Find the optimal weights.

After having obtained the pairwise comparison system, which we refer to as the set of judgements contained in vectors A^{BO} and A^{OW}, we can estimate the weights of the attributes. If we choose interpretations (6) and (10), we end up with the following system of linear equations in n variables:

$$\begin{cases} w_k a_{Bk} = w_B, \ \forall k \neq B \\ w_k = a_{kW} w_W, \ \forall k \neq W \\ w_1 + w_2 + \cdots + w_n = 1 \end{cases} \tag{11}$$

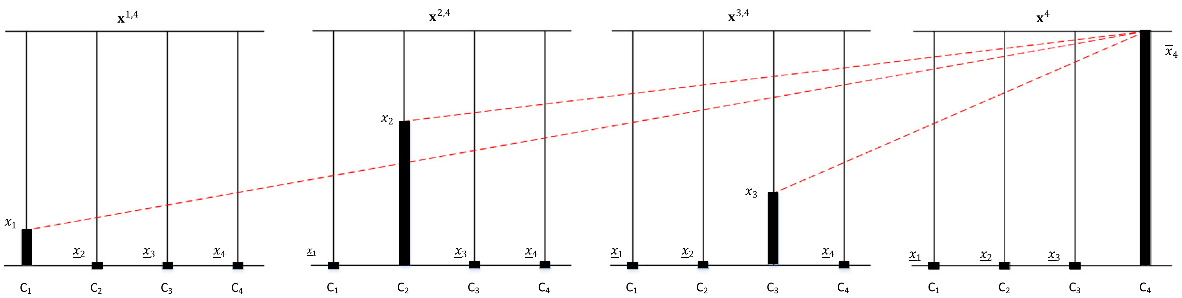

Fig. 5. The DM states the levels of the attributes C_1, C_2, C_3, such that $\mathbf{x}^{1,4} \sim \mathbf{x}^{2,4} \sim \mathbf{x}^{3,4} \sim \mathbf{x}^4$.

Unless the judgements elicited from the DM are fully rational, such an equation system does not have a solution. Since subjective judgements are seldom rational, it is necessary to use some methods to determine good estimates for the weights.

Following the formulation of the BWM [9,10] we want to identify the weight vector that minimizes the greatest absolute violation of the equations. Considering the non-negativity and normality conditions for the weights, this corresponds to solving the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & \max_j \{ |w_j a_{Bj} - w_B|, \ |w_W a_{jW} - w_j| \} \\ \text{subject to} \quad & w_1 + w_2 + \cdots + w_n = 1 \\ & w_j \geq 0, \ j = 1, 2, \cdots, n \end{aligned} \tag{12}$$

The optimization problem (12) can be equivalently rewritten as.

$$\begin{aligned} \text{minimize} \quad & \xi \\ \text{subject to} \quad & |w_j a_{Bj} - w_B| \leq \xi, \ \forall j \neq B \\ & |w_W a_{jW} - w_j| \leq \xi, \ \forall j \neq W \\ & w_1 + w_2 + \cdots + w_n = 1 \\ & w_j \geq 0, \ j = 1, 2, \cdots, n \end{aligned} \tag{13}$$

This last formulation can be easily linearized and its solution yields the optimal weights.

If, conversely, one wants to privilege the interpretations (6) and (10) of the judgements a_{Bk} and a_{kW}, the system of (non-linear) equations becomes.

$$\begin{cases} a_{Bk} = w_B / w_k, & \forall k \neq B \\ a_{kW} = w_k / w_W, & \forall k \neq W \\ w_1 + w_2 + \cdots + w_n = 1 \end{cases} \tag{14}$$

the optimization problem to minimize the maximum absolute discrepancy becomes,

$$\begin{aligned} \text{minimize} \quad & \max \{ |a_{Bj} - w_B / w_j|, |a_{jW} - w_j / w_W| \} \\ \text{subject to} \quad & w_1 + w_2 + \cdots + w_n = 1 \\ & w_j > 0, \ j = 1, 2, \cdots, n \end{aligned} \tag{15}$$

which is equivalent to.

$$\begin{aligned} \text{minimize} \quad & \xi \\ \text{subject to} \quad & |a_{Bj} - w_B / w_j| \leq \xi, \quad j \neq B, \\ & |a_{jW} - w_j / w_W| \leq \xi, \quad j \neq W \\ & w_1 + w_2 + \cdots + w_n = 1 \\ & w_j > 0, \ \ j = 1, 2, \cdots, n \end{aligned} \tag{16}$$

4. Consistency measurement

As also indicated by Keeney and Raiffa [11], it is important to keep inconsistencies at a “nominal” level. Keeney [16] mentioned inconsistency as one of the causes that can lead to errors in measuring the scaling constants, so it is important to quantify and localize the degree of inconsistency of sets of preferences.

There are two kinds of consistency, ordinal consistency and cardinal consistency, and one common desideratum in decision-making analysis is that the judgements of the DM be as ordinally consistent and cardinally consistent as possible. To measure how DMs deviate from these consistency conditions, we propose two indices: the ordinal consistency ratio and cardinal consistency ratio for the BWT, inspired by Liang et al. [24], Escobar et al. [25] and Cavallo et al. [26].

4.1. Ordinal consistency ratio

Definition 1. (*Ordinal consistency*)**:** *In the BWT, a pairwise comparison system is said to be ordinal-consistent if the order relations of the two paired comparison vectors (A^{BO} and A^{OW}) are the same. In formal terms* [27]:

$$(a_{Bk} - a_{Bj}) \times (a_{jW} - a_{kW}) > 0 \ \ or \ \ (a_{Bk} = a_{Bj} \text{ and } a_{jW} = a_{kW}), \ \forall k \text{ and } j \tag{17}$$

Checking the violation of ordinal consistency is very important because it has a vital impact on the ranking of the attributes. In order to measure to what extent DMs violate the ordinal consistency, we need to define an Ordinal Consistency Ratio.

Definition 2. (*Ordinal Consistency Ratio*)**:** *The Ordinal Consistency Ratio OR of a pairwise comparison system is defined as*:

$$OR = \max_j \; OR_j \tag{18}$$

where

$$OR_j = \frac{1}{n-1}\sum_{k=1}^{n} F\big((a_{Bk} - a_{Bj}) \times (a_{jW} - a_{kW})\big) \tag{19}$$

where $F(\gamma, \delta)$ is a step function, where $\gamma = a_{Bk} - a_{Bj}$, $\delta = a_{jW} - a_{kW}$, it is defined as:

$$F(\gamma,\delta) = \begin{cases} 1, & \text{if } \gamma \times \delta < 0, \\ 0.5, & \text{if } \gamma \times \delta = 0 \text{ and } (\gamma \neq 0 \text{ or } \delta \neq 0), \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

OR_j is called *local* ordinal consistency ratio,[1] indicating the degree of consistency with respect to the j^{th} attribute. With this ordinal consistency ratio ($OR_j \in [0,1]$), we can determine if the j^{th} attribute has different rank (and to what extent) in the two vectors.

OR is called *global* ordinal consistency ratio and reflects the ordinal consistency of the pairwise comparison system provided by the DM.

The rationale of OR_j formulation is that if attribute C_j overweighs attribute C_k, then the ordinal consistency should satisfy $a_{Bk} > a_{Bj}$ and $a_{jW} > a_{kW}$, i.e., $(a_{Bk} - a_{Bj}) \times (a_{jW} - a_{kW}) > 0$. If only one of $(a_{Bk} - a_{Bj})$ and $(a_{jW} - a_{kW})$ is equal to 0, we say that, in this situation, it violates weak ordinal relation, but if both are equal to 0, it is ordinal-consistent [25,26]. Overall, this approach is similar, but not identical, to Kendall's tau [27].

4.2. Cardinal consistency ratio

Definition 3. (*Cardinal consistency*)**:** *The preferences a_{Bj}, a_{jW} and a_{BW} elicited as in Section 3 are consistent if and only if*:

$$a_{Bj}a_{jW} = a_{BW}, \;\; \forall j. \tag{21}$$

To measure the deviation from the perfect cardinal consistency, Liang et al. [24] defined the following index:

Definition 4. (*Cardinal Consistency Ratio*)**:** *The Cardinal Consistency Ratio is formulated as follows*:

$$CR = \max_j \; CR_j \tag{22}$$

where,

$$CR_j = \begin{cases} \frac{|a_{Bj}a_{jW} - a_{BW}|}{a_{BW}a_{BW} - a_{BW}}, & a_{BW} > 1 \\ 0, & a_{BW} = 1 \end{cases} \tag{23}$$

CR is the *global* consistency ratio for all attributes and represents the overall consistency of the preferences; CR_j represents the *local* consistency level associated with attribute C_j, with which we can locate the most inconsistent attribute.

4.3. Thresholds

Identifying inconsistent judgements and knowing how much they deviate from a fully consistent status is not enough. Instead, we need to know under what threshold the inconsistency is acceptable, which is why, drawing from the study by Liang et al. [24], we propose a method to determine the thresholds of BWT based on the cardinal and ordinal consistency ratios.

[1] We opt for max instead of sum, because max could label a pairwise comparison system *in*consistent when there is at least one element that is not sufficiently consistent (for sufficiency, we use thresholds), which helps the analyst and the decision-maker locate the source of inconsistency for possible revision. The sum, on the other hand, looks at the whole system and not the individual pairwise comparisons. However, both max and sum can be used to aggregate the local measures of inconsistency [1].

If a DM is ordinally consistent, the rankings of the preferences obtained from the two judgement vectors (A^{BO} and A^{OW}) are the same, and only the intensities of preference may vary [28]. Therefore, in an ordinal sense, we can rely on the preferences provided by DM in this situation. Based on this idea, we can design a mechanism to find a suitable threshold.

Firstly, we use the Monte-Carlo method to estimate the probability distribution of CR s. We consider a number of attributes,n, ranging from 4 to 9 ($n = 2$ and $n = 3$ are excluded[2]). As $a_{BW} \in [1, \infty)$, which is a continuous set, in this study we consider only a subset of values, i.e., $a_{BW} \in \{2,3,4,5,6,7,8,9\}$. The choice of a discrete scale, with integers up to 9, is necessary to keep the Monte Carlo analysis computable, and is coherent with the standard approach used in the field of Design of Experiments (DoE) [29]. We will then carry out a full factorial analysis on the $6 \times 8 = 48$ combinations (6 is the number of attributes and 8 the number of different values of a_{BW}). For each combination, 10,000 pairs of *ordinal-consistent* vectors (A^{BO} and A^{OW}) will be generated randomly and will be labeled as *acceptable* comparisons. Also, we generate 10,000 pairs of *ordinal-inconsistent* vectors, which will be categorized as the *unacceptable* group.

In light of the distributions of the CR s of the two groups, there are significant overlaps, which means that there is no value of CR that can split the sets of acceptable and unacceptable preferences. In probabilistic terms, we expect that there is a threshold below which the CR s can be part of the acceptable group as much as possible, and above which the CR s can be part of the unacceptable group as much as possible (for example, see Fig. 6).

To achieve this, we can use the empirical cumulative distribution function, which can be defined as:

$$\widehat{F}_N(CR_i) = \frac{1}{N}\sum_{i=1}^{N} I\{CR_i \leq CR_T\} \tag{24}$$

where $I\{.\}$ is the indicator function:

$$I\{CR_i \leq CR_T\} = \begin{cases} 1, & if \quad CR_i \leq CR_T \\ 0, & otherwise \end{cases} \tag{25}$$

where N is the number of sampled vectors A^{BO} and A^{OW}, CR_i is the consistency ratio of the i th ($i \in \{1,2,\cdots,N\}$) pair of vectors, $CR_T \in [0,1]$ is the possible threshold.

We denote the cumulative distribution of CR s in ordinal-consistent vectors (the Acceptable group) as $\widehat{F}^A(CR_T)$, and the cumulative distribution of CR s from the Unacceptable group is denoted as $\widehat{F}^U(CR_T)$. We accept the $CR\prime$ s within the threshold (which is $\widehat{F}^U(CR_T)$), and reject the $CR\prime$ s beyond the threshold (which is $1 - \widehat{F}^A(CR_T)$). In order to make the proportion of ordinal-inconsistent $CR\prime$ s that we accept as small as possible, and also make the proportion of ordinal-consistent CR s that we reject as small as possible, we need to calculate the threshold, CR_T, so that $\widehat{F}^U(CR_T) = 1 - \widehat{F}^A(CR_T)$. Fig. 7 sketches the simulation algorithm used to calculate the thresholds.

Table 1 shows the thresholds for combinations of attributes range $n = \{4,\cdots,9\}$ and integer values of a_{BW} from 2 to 9.

It is worth to mention that, unlike the original BWM, where the scale of values of a_{BW} is a discrete set {1,2,$\cdots$,9}, the BWT relies on a continuous scale, so that, in principle, a_{BW} can take any value greater than 1. For this reason, the proposal by Liang et al. [24] must be readapted to this framework, but with two differences:

Firstly, in the calculation of the thresholds, all the real numbers in the interval $[1, a_{BW}]$ can be sampled, and not only the integers. The results of sampling from a continuous set substantially differs from the study by Liang et al. [24], where entries were sampled from the discrete set {1,2,$\cdots$,a_{BW}}.

Secondly, to acknowledge the continuous nature of a_{BW} we seek for a continuous function (a *response surface* in the DoE terminology) to approximate the values in Table 1, and help find the thresholds also for non-integer values of a_{BW}. Formally, we consider the variables x and y to represent n and a_{BW}, respectively, and use the following quadratic fit,

$$Q(x,y) = \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}^T \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}^T \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + c \tag{26}$$

where $x_0, y_0, a_{11}, a_{12}, a_{22}, b_1, b_2, c$ are the parameters. By using a least squares minimization approach, the optimal parameters[3] yield a good fit. See Fig. 8 for a graphical representation of the results.[4]

The adapted algorithm in this study considers a_{BW} as continuous space, which is closer to the reality. With these thresholds, we can now determine whether the consistency level (consistency ratio) of a DM is acceptable or not. Since a_{BW} could be a number with decimal digits, one may either use the threshold obtained through (26) or the threshold of the approximate integer in Table 1.

[2] Cases with $n = 2$ and $n = 3$ are excluded from the analysis, because in such cases, the ordinal-inconsistent situation does not appear.

[3] By minimizing the Euclidean norm, the optimal parameters are: $x_0 = 9.8584$, $y_0 = 7.8288$, $a_{11} = -0.0017$, $a_{12} = -0.0003$, $a_{22} = -0.0003$, $b_1 = 0.0058$, $b_2 = 0.0055$, $c = -0.4187$.

[4] It is worth mentioning that our decision to use a single quadratic function to identify thresholds for $n \in \{4,\cdots,9\}$ and $a_{BW} \in [2,9]$ was made for reasons of simplicity. In theory, we could have proposed different univariate quadratic fits for each single value of n.

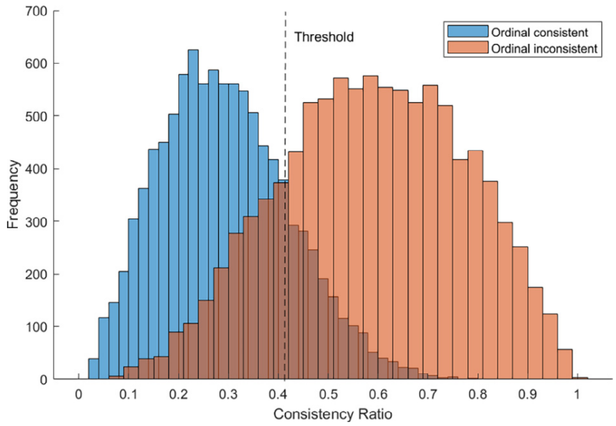

Fig. 6. The distribution of CR s in the two groups (in case $n = 9$, $a_{BW} = 9$).

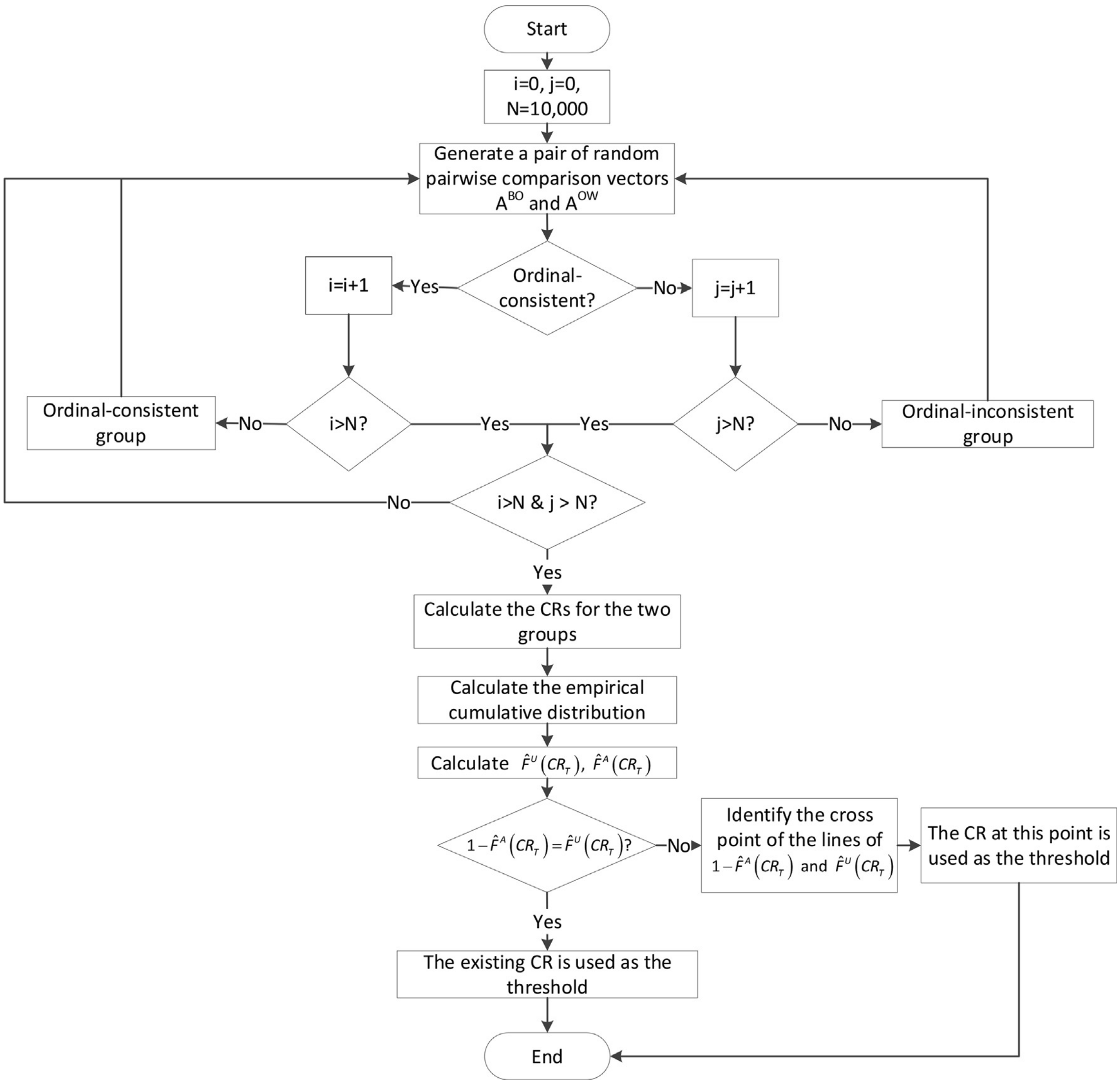

Fig. 7. The simulation algorithm used to calculate the thresholds (adapted from Liang et al. [24]). The code for the simulation algorithm can be seen in: https://github.com/fuqi15/BWT.

Table 1
The thresholds of the consistency ratios in different combinations.

a_{BW}	n					
	4	5	6	7	8	9
2	0.34	0.35	0.36	0.38	0.39	0.39
3	0.31	0.33	0.35	0.37	0.38	0.39
4	0.32	0.34	0.36	0.37	0.39	0.4
5	0.31	0.35	0.37	0.38	0.4	0.4
6	0.32	0.35	0.37	0.38	0.4	0.41
7	0.33	0.35	0.37	0.39	0.4	0.41
8	0.32	0.35	0.37	0.39	0.4	0.41
9	0.33	0.36	0.37	0.39	0.4	0.41

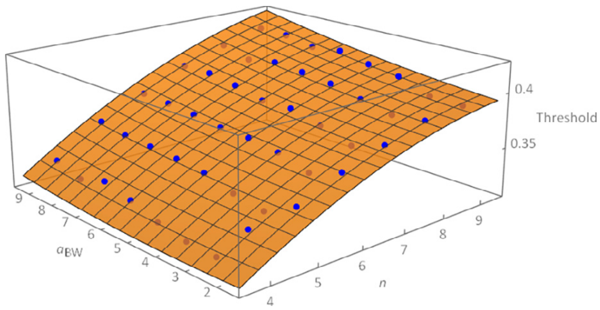

Fig. 8. The graphical representation of the thresholds.

4.4. Improving consistency

When DMs provide their preferences, we need to firstly check their consistency. If their cardinal consistency level is not acceptable, or if they violate the ordinal consistency, we suggest the DMs to revise their judgments.

Usually, the consistency improving process is guided by a moderator, who helps the DMs to revise their preferences. The following steps describe the procedure of the consistency improving process, which is also sketched in Fig. 9.

Step 1. Let $A^{BO(t)} = \left(a_{Bj}^{(t)}\right)$ and $A^{OW(t)} = \left(a_{jW}^{(t)}\right), j = 1, 2, \cdots, n$ be the original preferences provided by the DM. Let t indicate the iteration number.

Step 2. Compute the ordinal consistency ratio $\mathrm{OR}^{(t)}$ and the cardinal consistency ratio $CR^{(t)}$ by using equations (17)-(23). Based on a_{BW} (if a_{BW} is not an integer, round it up to the nearest whole number) and n, check the corresponding consistency threshold CR_T in Table 1.

Step 3. Check the cardinal consistency. If $CR^{(t)} > CR_T$, the consistency level is not acceptable, go to Step 4 to revise the preferences. If $CR^{(t)} \leq CR_T$, go to Step 5 to check the ordinal consistency.

Step 4. Revise preferences and improve the consistency. The inconsistent preferences should be adjusted to cardinal-consistent, which should satisfy:

$$\frac{\left|a_{Bj}^{(t)} a_{jW}^{(t)} - a_{BW}\right|}{a_{BW} a_{BW} - a_{BW}} \leq CR_T \tag{27}$$

Therefore, the DM should revise $a_{Bj}^{(t)}$ or $a_{jW}^{(t)}$ or both in the admissible ranges:

$$(a_{BW} - CR_T(a_{BW} a_{BW} - a_{BW}))/a_{jW}^{(t)} \leq a_{Bj}^{(t)} \leq (a_{BW} + CR_T(a_{BW} a_{BW} - a_{BW}))/a_{jW}^{(t)} \tag{28}$$

$$(a_{BW} - CR_T(a_{BW} a_{BW} - a_{BW}))/a_{Bj}^{(t)} \leq a_{jW}^{(t)} \leq (a_{BW} + CR_T(a_{BW} a_{BW} - a_{BW}))/a_{Bj}^{(t)} \tag{29}$$

and $a_{Bj}^{(t)}, a_{jW}^{(t)} \in [1, a_{BW}]$.

If the DM wants to revise his/her preferences in an acceptable range and also improve his/her consistency level,[5] then the ranges of $a_{Bj}^{(t)}$ or $a_{jW}^{(t)}$ for adjustment are:

[5] When the consistency level is not acceptable, the DM should revise his/her preferences. But if his/her preferences are ordinal-consistent and cardinal-consistent (acceptable consistency), then adjustments are not required, because the objective of elicitation is not achieving the optimal consistency level. However, if the DM tries to improve his/her consistency level, then the admissible ranges provide a way to reach that goal.

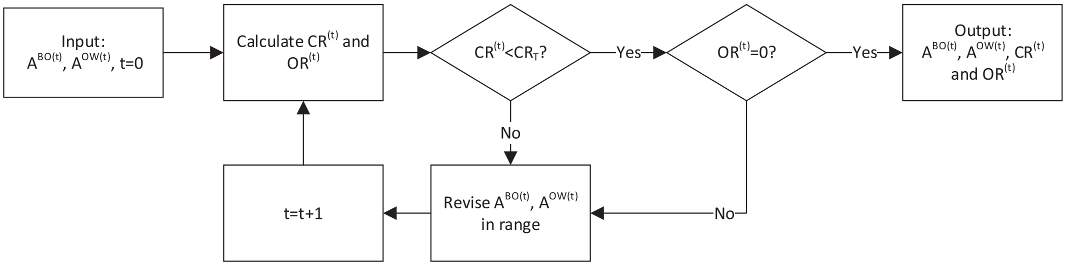

Fig. 9. Consistency improving process.

$$\min\left(a_{Bj}^{(t)}, a_{Bj}^{(t)} - \left(a_{Bj}^{(t)} a_{jW}^{(t)} - a_{BW}\right)/a_{jW}^{(t)}\right) \le a_{Bj}^{(t)} \le \max\left(a_{Bj}^{(t)}, a_{Bj}^{(t)} - \left(a_{Bj}^{(t)} a_{jW}^{(t)} - a_{BW}\right)/a_{jW}^{(t)}\right) \quad (30)$$

$$\min\left(a_{jW}^{(t)}, a_{jW}^{(t)} - \left(a_{Bj}^{(t)} a_{jW}^{(t)} - a_{BW}\right)/a_{Bj}^{(t)}\right) \le a_{jW}^{(t)} \le \max\left(a_{jW}^{(t)}, a_{jW}^{(t)} - \left(a_{Bj}^{(t)} a_{jW}^{(t)} - a_{BW}\right)/a_{Bj}^{(t)}\right) \quad (31)$$

Hint: If $a_{Bj}^{(t)} \times a_{jW}^{(t)} < a_{BW}$, then the DM needs to think about increasing $a_{Bj}^{(t)}$ or $a_{jW}^{(t)}$ or both, otherwise, decreasing the pairwise comparisons is recommended. In any case, the idea is to approach the condition $a_{Bj}^{(t)} \times a_{jW}^{(t)} \approx a_{BW}$.

After revision, set $t = t + 1$, and go back to Step 2.

Step 5. Check the ordinal consistency. If $OR^{(t)} = 0$, it means the given preferences are ordinal-consistent, go to next step. If $OR^{(t)} > 0$, it means the given preferences are ordinal-inconsistent, the DM should revise his/her judgments to reach full ordinal consistency, the local ordinal consistency ratio obtained by Equation (19) can help the DM to identify the most inconsistent attribute, then go back to Step 4.

Step 6. Output $A^{BO(t)}$, $A^{OW(t)}$, $OR^{(t)}$ and $CR^{(t)}$.

5. Case study

In this section, we illustrate the proposed BWT model by applying it to a port evaluation problem. The port performance measurement research helps ports anticipate and respond to possible future changes in port choice by shippers, freight forwarders and carriers. One of the studies on this problem was conducted by Rezaei et al. [30], from which we use the port leg-related services and facilities data (attributes, alternatives and evaluation scores) recorded by that study to present the BWT procedure.

To evaluate this study, we contacted a competent expert: a program manager at Digital Port Solutions of Vopak (in the Netherlands). She is an expert in the port choice problem and is familiar with MADM methods. The interview was conducted online in three phases. In phase one, we introduced the problem to the expert; in phase two we identified the preferences of the expert with regard to the BWT procedure; in phase three we collected the modified values after checking the consistency. After the interview, we calculated the weights of attributes to produce the ranking of the alternatives. Below, we describe the nine steps of the BWT we used in this case study.

Step 1. Determine alternatives and attributes.

Firstly, to evaluate the ports, the study by Rezaei et al. [30] included six attributes: terminal handling charges, International Ship and Port Facility Security Code (ISPS), customs service (rated on a 1-to-7 scale), port reputation (rated on a 1-to-7 scale), satisfaction with terminal operations (rated on a 1-to-7 scale), and number of container terminals. Terminal handling charges and ISPS are *cost attributes* and all the others are *benefit attributes*. Seven ports were examined in the study. The scores of the alternative ports with respect to the attributes are presented in Table 2.

Table 2
The recorded scores for the seven ports [30].

Ports	Attributes					
	Terminal handling charges (€/TEU)	ISPS (€/unit)	Customs service	Port reputation	Satisfaction with terminal operations	Number of container terminals
Piraeus	106	11	4.2	3.8	3.4	3
Koper	145	11	5.12	5.24	5	1
Genoa	179	13	4.2	4.4	4.4	2
Antwerp	179	12	5.44	5	5.11	4
Rotterdam	202	13	5.5	5.93	5.29	6
Hamburg	223	16	5.56	6.06	5.41	4
Gdansk	103	14	4.6	5	4.4	2

Step 2. Determine value functions for attributes.

Next, we need to determine the value functions of each attribute as stated in Section 3, and in this study, we adopt the mid-value splitting technique presented in the study of Keeney and Raiffa [11], where the readers can find the detailed definitions. Here we simplify the assessment procedure as follows:

(1) For the j th attribute, we let $v(\underline{x}_j) = 0$ and $v(\bar{x}_j) = 1$.
(2) Determine the mid-value point (denoted as $x_{.5}$) of $[x_0, x_1]$, we let $v(x_{.5}) = 0.5$.
(3) Determine the mid-value point, $x_{.75}$, of $[x_{.5}, x_1]$, and make $v(x_{.75}) = 0.75$.
(4) Determine the mid-value point, $x_{.25}$, of $[x_0, x_{.5}]$, and let $v(x_{.25}) = 0.25$.
(5) To check the consistency of the preferences, the DM needs to be certain that $x_{.5}$ is the midvalue point of $[x_{.25}, x_{.75}]$; otherwise modification is necessary to guarantee the consistency.
(6) Use the points (x_k, k) for $k = 0, 0.25, 0.5, 0.75, 1$, to estimate the value function v_j.

Subsequently, the mid-value points for the attributes of this study are obtained and shown in Table 3. These points can be plotted in Fig. 10 and interpolated by the v_1 - v_6 curves for C_1 to C_6, respectively.

Step 3. Identify the best and worst attributes.

Following Section 3, the expert considered C_1 as the best attribute and C_3 as the worst.

Step 4. Determine the best to others tradeoffs.

To tradeoff the best attribute C_1 to the other attributes, we generated the following hypothetical consequences based on the method described in Section 3 and asked the expert to provide the undetermined values $(x_1^{1,2}, x_1^{1,3}, x_1^{1,4}, x_1^{1,5}, x_1^{1,6})$ so that the following paired consequences are indifferent one to another,

$$\mathbf{x}^{1,2} \sim \mathbf{x}^2 \iff \left(x_1^{1,2}, 16, 4.2, 3.8, 3.4, 1\right) \sim (223, 11, 4.2, 3.8, 3.4, 1)$$

$$\mathbf{x}^{1,3} \sim \mathbf{x}^3 \iff \left(x_1^{1,3}, 16, 4.2, 3.8, 3.4, 1\right) \sim (223, 16, 5.56, 3.8, 3.4, 1)$$

$$\mathbf{x}^{1,4} \sim \mathbf{x}^4 \iff \left(x_1^{1,4}, 16, 4.2, 3.8, 3.4, 1\right) \sim (223, 16, 4.2, 6.06, 3.4, 1)$$

Table 3
The mid-value points for each attribute.

	C_1	C_2	C_3	C_4	C_5	C_6
$\underline{x}$	223	16	4.2	3.8	3.4	1
$x_{.25}$	194	14.8	4.6	4.5	4.2	2.25
$x_{.5}$	163	13.5	4.9	5	4.8	3.5
$x_{.75}$	127	12.25	5.3	5.7	5.2	4.75
$\bar{x}$	103	11	5.65	6.06	5.41	6

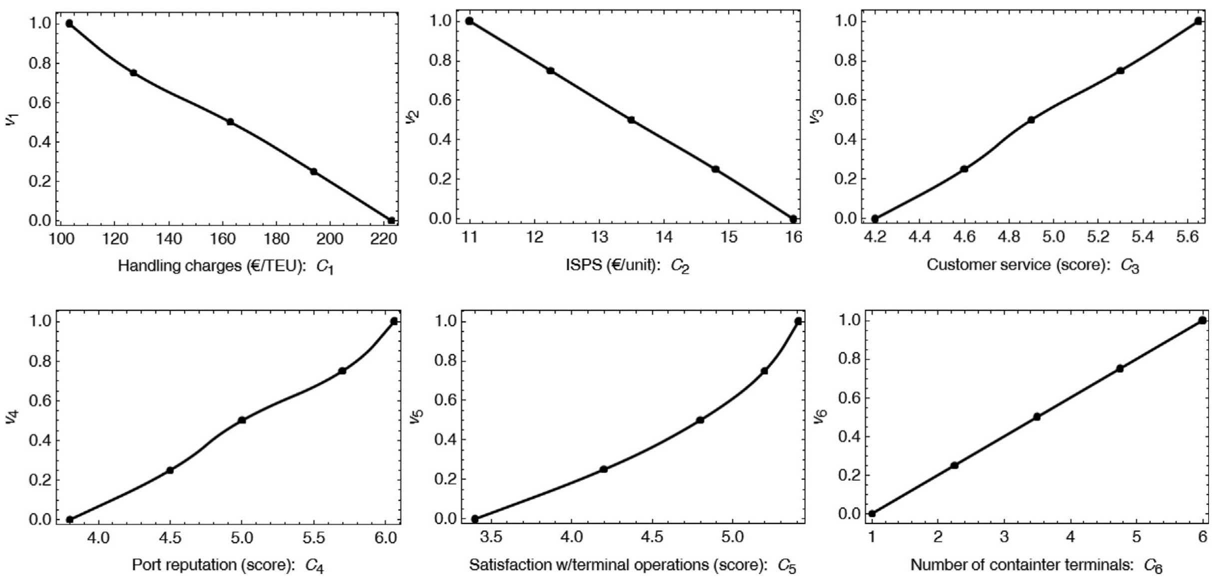

Fig. 10. The assessed value functions for attributes C_1 to C_6 as second order interpolations of the values in Table 3.

$$\mathbf{x}^{1,5} \sim \mathbf{x}^5 \iff \left(x_1^{1,5}, 16, 4.2, 3.8, 3.4, 1\right) \sim (223, 16, 4.2, 3.8, 5.41, 1)$$

$$\mathbf{x}^{1,6} \sim \mathbf{x}^6 \iff \left(x_1^{1,6}, 16, 4.2, 3.8, 3.4, 1\right) \sim (223, 16, 4.2, 3.8, 3.4, 6)$$

After assessment, the expert determined the following values:

$$\left(x_1^{1,1}, x_1^{1,2}, x_1^{1,3}, x_1^{1,4}, x_1^{1,5}, x_1^{1,6}\right) = (103, 200, 210, 180, 190, 185)$$

Step 5. Determine the others to the worst tradeoff.

In addition, the expert was asked to determine the values $(x_1, x_2, x_3, x_4, x_5, x_6)$ according to Section 3, so that the following indifference relations on the generated consequences can be satisfied:

$$\mathbf{x}^{1,3} \sim \mathbf{x}^3 \iff (x_1, 16, 4.2, 3.8, 3.4, 1) \sim (223, 16, 5.56, 3.8, 3.4, 1)$$

$$\mathbf{x}^{2,3} \sim \mathbf{x}^3 \iff (223, x_2, 4.2, 3.8, 3.4, 1) \sim (223, 16, 5.56, 3.8, 3.4, 1)$$

$$\mathbf{x}^{4,3} \sim \mathbf{x}^3 \iff (223, 16, 4.2, x_4, 3.4, 1) \sim (223, 16, 5.56, 3.8, 3.4, 1)$$

$$\mathbf{x}^{5,3} \sim \mathbf{x}^3 \iff (223, 16, 4.2, 3.8, x_5, 1) \sim (223, 16, 5.56, 3.8, 3.4, 1)$$

$$\mathbf{x}^{6,3} \sim \mathbf{x}^3 \iff (223, 16, 4.2, 3.8, 3.4, x_6) \sim (223, 16, 5.56, 3.8, 3.4, 1)$$

We received the following judgements from the expert:

$$(x_1, x_2, x_3, x_4, x_5, x_6) = (210, 15, 5.56, 5.2, 4, 3)$$

By using the value functions from Step 2, the revised vectors A^{BO} and A^{OW} are:

$$A^{BO} = (1, 0.2, 0.11, 0.37, 0.28, 0.32)$$

$$A^{OW} = (0.11, 0.21, 1, 0.54, 0.18, 0.4)$$

Step 6. Check the consistency of the preferences.

When we had the preferences A^{BO} and A^{OW}, we needed to check the consistency level and determine whether the preferences need to be revised. We first checked the ordinal consistency by using the method described in Section 4. Applying Equations (17)-(20), we found that $OR = 0.5$, which means that the judgements violated the ordinal consistency. To identify the locations of the violated values, we used the local ordinal consistency ratios, from Equation (19) , as shown in Table 4.

In Table 4, the value "1″ indicates that the judgements in two corresponding attributes are ordinal-inconsistent, and "0" means they are ordinal consistent. For example, the judgements in C_2 are not ordinal-consistent with C_4 and C_6. Based on Table 4 we could clearly identify which pairs of comparisons need to be modified.

Step 7. Modify the inconsistent preferences repeating Steps 4 and 5.

After we found and located inconsistency in the obtained preferences, we contacted the expert again to ask her to rethink about the judgements. After careful consideration, the expert acknowledged some inconsistencies, and then modified the preferences by applying the consistency improving process in Section 4.4. According to formulas (27)-(31) of this process, we obtain the admissible ranges of A^{BO} and A^{OW} in Table 5.

Based on the acceptable ranges, the expert revised some values of the Others-to-Worst tradeoffs, and left the Best-to-Others tradeoffs unchanged:

$$\left(x_1^{1,1}, x_1^{1,2}, x_1^{1,3}, x_1^{1,4}, x_1^{1,5}, x_1^{1,6}\right) = (103, 200, 210, 180, 190, 185)$$

$$(x_1, x_2, x_3, x_4, x_5, x_6) = (210, 14, 5.56, 4.2, 4, 1.8)$$

Table 4
The ordinal consistency check table.

	C_1	C_2	C_3	C_4	C_5	C_6	OR_j
C_1	0	0	0	0	0	0	0
C_2	0	0	0	1	0	1	0.4
C_3	0	0	0	0	0	0	0
C_4	0	1	0	0	1	1	0.6
C_5	0	0	0	1	0	1	0.4
C_6	0	1	0	1	1	0	0.6

Table 5
The admissible ranges for revision.

Attributes		C_1	C_2	C_3	C_4	C_5	C_6
Original		[103, 223]	[11, 16]	[4.2, 5.65]	[3.8, 6.06]	[3.4, 5.41]	[1, 6]
Acceptable ranges	A^{BO}	[103, 193.23]	[11, 15.34]	[4.4, 5.65]	[4.16, 6.06]	[3.96, 5.41]	[1.57, 6]
	A^{OW}	[103, 210]	[11, 15.31]	[4.61, 5.65]	[4.16, 6.06]	[3.80, 5.41]	[1.57, 6]
Improving ranges	A^{BO}	[103, 210]	[13.28, 15.05]	[5.65, 5.65]	[4.41, 4.7]	[4.31, 5.07]	[2.43, 2.62]
	A^{OW}	[103, 210]	[12.87, 15]	[5.65, 5.65]	[4.52, 5.2]	[4, 4.28]	[2.71, 3]

By using the value functions from Step 2, we obtained the new A^{BO} and A^{OW}:

$$A^{BO} = (1, 0.2, 0.11, 0.37, 0.28, 0.32)$$

$$A^{OW} = (0.11, 0.41, 1, 0.13, 0.18, 0.16)$$

Then we checked the consistency of the modified preferences again using the ordinal and cardinal consistency indices proposed in Section 4. Now the ordinal consistency ratio $OR = 0$, which means that the judgements are fully ordinal-consistent; the cardinal consistency ratio $CR = 0.18$, which is less than the threshold 0.37 when we refer to Table 1 (in this case we have 6 attributes and $a_{BW} = \frac{1}{a_{WB}} = \frac{1}{0.11} \approx 9$), which means the judgements can be accepted.

Step 8. Calculate the optimal weights of attributes.

The new set of values finally satisfies the ordinal consistency, while the cardinal consistency ratios are below the consistency threshold. Next, we can apply the BWT model in Section 3 to obtain the optimal weights for the attributes. In this study, we adopted the linear model (13), and obtained the optimal weights as follows.

$$w = (0.39, 0.1, 0.03, 0.18, 0.14, 0.16)$$

Step 9. Rank the alternatives.

Finally, we use the additive value function (1) to aggregate the weights and the assessment values of alternatives (after being normalized by the value functions presented in Fig. 10), and the aggregated values of alternatives are obtained as Table 6.

Therefore, according to the ranking, the port of Rotterdam is the most favorable option.

6. Discussion

There are several advantages in the BWT proposed in this study. Combined with the case study, we try to discuss it and compare it with other methods from the perspectives of anchoring bias, consistency check, computational complexity and the completeness of information.

6.1. Anchoring bias analysis

The anchoring bias is a cognitive bias which explains people's tendency towards the first piece of information they receive when they are evaluating something [31]. Such bias is persistent in several MADM methods, especially those with a single anchor like SMARTS, Swing, and Tradeoff (see, for instance, Buchanan and Corner [32], Montibeller and Von Winterfeldt [33], Rezaei [34]). A recent study in this area, conducted by Rezaei [34], shows that respondents tend to provide larger scores than their actual values to the other attributes when they are using a larger anchor (for example in the Swing method, where respondents start with identifying the most important attribute assigning it a score equal to 100), whereas in methods with a small anchor (like in the SMARTS method where respondents start with identifying the least important attribute assigning it a 10) the respondents assign scores to attributes lower than their actual ones. The preferences used to obtain weights in

Table 6
The normalized values, aggregated values and the final ranking.

Ports	Normalized value						Aggregated value	Ranking
	C_1	C_2	C_3	C_4	C_5	C_6		
Piraeus	0.97	1	0	0	0	0.4	0.54	6
Koper	0.61	1	0.65	0.55	0.59	0	0.74	3
Genoa	0.37	0.6	0	0.21	0.32	0.2	0.32	7
Antwerp	0.37	0.8	0.85	0.5	0.67	0.6	0.79	2
Rotterdam	0.18	0.6	0.89	0.9	0.85	1	0.87	1
Hamburg	0	0	0.93	1	1	0.6	0.73	4
Gdansk	1	0.41	0.25	0.5	0.32	0.2	0.68	5

the classical Tradeoff procedure could be equivalent to one of the two vectors that we used in BWT, Best-to-Others (A^{BO}) or Others-to-Worst(A^{BO}). For the vector A^{BO}, its anchor is the best attribute, which is similar to the Swing method; and for the vector A^{OW}, its anchor is the worst attribute, which is similar to the SMARTS method.

Taking the case study in Section 5 as an example, if we obtain the weights that are only based on the vector A^{BO} or A^{OW}, as we can see in Fig. 11, the weight of the best attribute ($C1$) obtained from A^{BO} is obviously larger than that from A^{OW}, and the weight of the worst attribute ($C3$) obtained from A^{BO} is also larger than that from A^{OW}. From this perspective, if we only consider one vector to calculate weights, as the classical Tradeoff procedure does, then we would encounter the anchoring bias.

One of the advantages of BWT is to remedy this anchoring bias [35]. By combining the two opposite reference attributes (best and worst), the potential anchoring bias is mitigated, as we can see from the BWT line, which is located between lines BO and OW in Fig. 11.

6.2. Consistency check

Experimental studies have found that the classical Tradeoff procedure has higher inconsistency rate than Swing and Ratio method, and 67 % of the subjects who applied Tradeoff elicitation procedure shown inconsistency [19,36]. Based on these studies, the FITradeoff method proposed by de Almeida et al. [17,37] considers that the DM is not able to specify the tradeoff values and this information cannot be obtained in a consistent way from the DM.

The elicitation of indifference relations may not be as easy as for Swing or Ratio methods, but according to the successful applications of the Tradeoff procedure and our experiments, not only the elicitation is possible, but often also not difficult. Since the classical Swing and Ratio method are not able to check consistency of preferences (neither ordinal nor cardinal), then, from the perspective of validation [3], it may be preferable to use the Tradeoff procedure. Although an extension of the FITradeoff method has considered the consistency checking and revision, it is for ordinal inconsistency and it still lacks detailed procedures for the revision of preferences [37]. Therefore, solving the consistency problem is necessary for the classical Tradeoff procedure.

The BWT method proposed in this study assumes that a DM can specify his/her preferences and these preferences can be adjusted to an acceptable level. Based on these assumptions, we have developed a systematic consistency check and improvement process. With this process, a DM can locate the inconsistencies and visualize the acceptable adjustment and improvement ranges. It makes the consistency revision process, which is absent in the classical BWM and Tradeoff procedure, possible and easy.

For example, in the case study of Section 5, the consistency checking and improving process helped the expert identify her inconsistencies by using the local ordinal inconsistency ratios in Table 4, and showing the admissible ranges for adjustment in Table 5. Although the original preferences were ordinal-inconsistent, and the cardinal consistency ratios were in the acceptable range, it was still suggested to revise the preferences to be fully ordinal consistent. The expert only revised some of her A^{OW} judgments within the acceptable ranges. After this revision, the ordinal and cardinal consistency were acceptable, with no need for further adjustments.

6.3. Computational complexity

In its basic form, the original Tradeoff procedure requires a minimum of $(n-1)$ pairs of comparisons and, with this number of comparisons, it does not allow to check the consistency. To that end, on the other extreme, it could consider all the possible combinations of comparisons, which would result in $n(n-1)$ pairs (bidirectional), or $n(n-1)/2$ pairs (unidirectional) [38]. The proposed BWT method requires $(2n-3)$ comparisons, which is linear with respect to the number of attributes, and, when a large number of attributes is used, the number of comparisons remains tractable. Fig. 12 represents the number of comparisons required by various methods. In this sense, the BWT appears to strike a fair balance between having

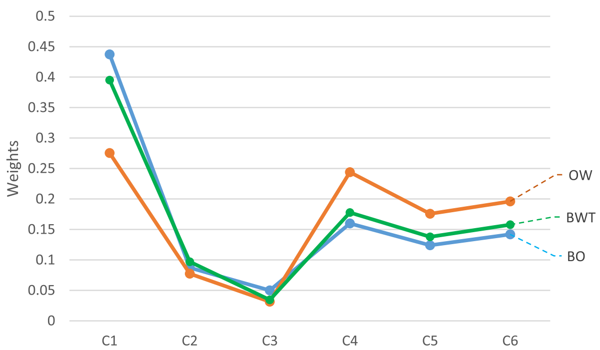

Fig. 11. The weights obtained separately by vectors A^{BO}, A^{OW} and BWT.

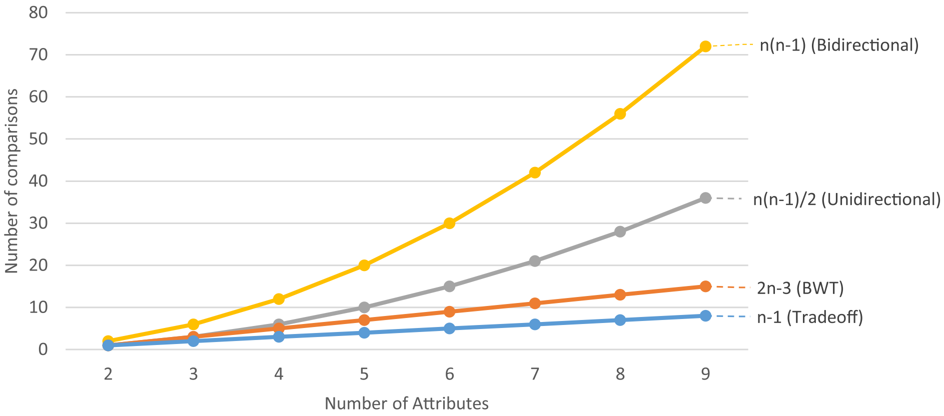

Fig. 12. Relation between number of attributes and number of comparisons required by questioning techniques.

an acceptable level of redundancy in the questioning process, necessary for the evaluation of inconsistency, and the cognitive burden required from the DM, and presents a good scalability for larger problems. In addition, although we know inconsistent preferences are practically unavoidable, it is still difficult to identify which (and how much) judgements contribute to that inconsistency. Therefore, it is important to have a module within in BWT which enables the DM to check the consistency in a systematic way.

6.4. Methods with complete information V.S. methods with incomplete information

This section discusses the BWT with respect to information availability. In real decision contexts, it may be difficult, for a DM, to provide indifference tradeoff information. Therefore, to ease the preference elicitation process, a number of researches proposed methods requiring partial and/or incomplete information as input [17,37,39,40]. Often, the incomplete information takes the form of weak order relations and inequalities instead of indifference relations and equalities. When we compare methods using incomplete information and methods using complete information, we shall consider three aspects: the availability of the complete information, the cost of eliciting preferences, and the consistency check of the comparisons.

The availability of the complete information. As the applications of classical Tradeoff procedure have proved that indifference relations can be elicited [41], so, similarly, complete information can also be obtained in BWT. Of course, some DMs cannot, or may not be willing to, provide complete or certain information, but there are also DMs who are familiar with the situation and can provide certain/complete information. If a DM can only provide incomplete information, then methods like FITradeoff [37] could be a good option. On the contrary, if complete information is available, using the BWT method has the advantage of helping mitigate DMs' anchoring bias and measure (and improve) the inconsistency of preferences.

The cost of eliciting preferences. One major argument in favor of using incomplete information is that eliciting complete information demands higher cognitive effort. As observed in the literature [19], using inequalities seems easier, with respect to the required cognitive effort, but it remains difficult to compare the cognitive effort of methods employing different techniques. Moreover, some methods have slightly different scopes. Consider, for example, the FITradeoff method [37], where the number of required comparisons is not fixed and depends on many factors, such as the values of the consequences and the order relation on the weights of the attributes. Last but not least, the FITradeoff aims at finding the optimal alternative with a reduced cognitive effort, while the BWT is a structured, and in this sense less flexible, used to elicit a unique weight vector. It is therefore difficult to compare methods on the ground of the demanded cognitive effort as (i) the cognitive effort associated with different types of questions is hard to assess, (ii) the number of questions is not always predetermined, and (iii) some methods are intrinsically different.

The consistency check of the comparisons. As Albert Einstein allegedly said, *everything should be made as simple as possible, but not simpler*. We try to make the elicitation procedure as simple as possible, but the minimum condition is that the consistency of these elicited preferences can be examined. The existing methods either lack a phase of consistency check, or they can only check ordinal consistency and seldom they are of any help in guiding the analyst and the DM through a consistency improvement process [37]. One of the advantages of BWT is that it can check both ordinal and cardinal consistency. Moreover, it can help DMs to improve their consistency level with acceptable ranges and improving ranges for references which can reduce the cognitive effort required from DMs.

DMs could choose to provide complete or partial/incomplete information, and use different methods to deal with these two types of information, and there is no right or wrong choosing one way or the other. To us, methods using incomplete information are complements to methods using complete information, especially when the cost of obtaining complete information is too high, if not even impossible.

7. Conclusion and future research

In this study, we developed a multi-attribute decision-making method called BWT that can be viewed as an attempt to combine the merits of the traditional BWM and the Tradeoff procedure, without losing the characteristics that have made the two methods popular. More specifically, with the BWT, we can elicit weights in a more structured way using the prescriptive MAVT approach, which considers the attribute range effect, and at the same time have a guided choice of the attributes to be compared with a check of the consistency of the preferences.

From the point of view of the original BWM, the BWT is better at eliciting preferences, obtaining weights, and complies with the theory of MAVT. From the point of view of MAVT, the BWT, represents a scheme for questioning (and testing the consistency of) DMs in the elicitation process. We want to emphasize that our proposal to include thresholds should not be interpreted as a strict acceptance/rejection rule, but as an effort to increase the intelligibility of the inconsistency index *CR*. Besides, BWT does not use ratio scale (from 1 to 9) for pairwise comparisons, which could avoid personal interpretations [42].

In addition, the BWT model may make another potential contribution. The traditional Tradeoff procedure may suffer from anchoring bias (or scale compatibility bias) and loss aversion bias, if the preferences obtained from DMs depend on a measuring stick (anchor) [31,43,44,45]. By introducing the consider-the-opposite-strategy, the use of two opposite reference attributes (best and worst) could reconcile the possible anchoring bias when eliciting judgements from the DMs [46].

While the method has been devised for situations where a decision-maker/analyst has full information about the alternatives including the ranges of attributes, in situations where a decision-maker has no full information about the alternatives, considering a nominal range [15] could help applying the method, however in those situations the findings should be interpreted more carefully.

We leave it to future studies to examine whether the anchoring bias affects the ultimate decisions resulting from using the Tradeoff procedure and how BWT helps remedy the bias (if it is able to). Besides, while conducting the surveys, the experts, sometimes, were hesitant with regard to providing the precise judgements when we asked the tradeoff questions, so future studies could examine how incomplete and partial information can be considered within the BWT method (and BWM as well). Moreover, it is helpful to develop a decision support system to facilitate DMs to operationalize BWT in the future.

CRediT authorship contribution statement

Fuqi Liang: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration, Funding acquisition. **Matteo Brunelli:** Conceptualization, Methodology, Software, Validation, Data curation, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Jafar Rezaei:** Conceptualization, Methodology, Validation, Formal analysis, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors gratefully acknowledge financial support for PhD study from the China Scholarship Council (No. 201708440305).

References

[1] M. Brunelli, Recent advances on inconsistency indices for pairwise comparisons—A commentary, Fundamenta Informaticae 144 (2016) 321–332.
[2] S. Greco, M. Ehrgott, J. Figueira, Multiple criteria decision analysis: state of the art surveys, Springer, New York, 2016.
[3] N.H. Zardari, K. Ahmed, S.M. Shirazi, Z.B. Yusop, Weighting methods and their effects on multi-criteria decision making model outcomes in water resources management, Springer International Publishing, New York, 2015.
[4] L.C. Dias, A. Morton, J. Quigley, Elicitation, Springer International Publishing, Switzerland, 2018.
[5] T.L. Saaty, A scaling method for priorities in hierarchical structures, J. Math. Psychol. 15 (1977) 234–281.
[6] W. Edwards, F.H. Barron, SMARTS and SMARTER: Improved simple methods for multiattribute utility measurement, Organ. Behav. Hum. Decis. Process. 60 (1994) 306–325.
[7] P.A. Bottomley, J.R. Doyle, A comparison of three weight elicitation methods: Good, better, and best, Omega 29 (2001) 553–560.
[8] D. Von Winterfeldt, W. Edwards, Decision analysis and behavioral research, Cambridge University Press, Cambridge, 1986.
[9] J. Rezaei, Best-worst multi-criteria decision-making method, Omega 53 (2015) 49–57.
[10] J. Rezaei, Best-worst multi-criteria decision-making method: Some properties and a linear model, Omega 64 (2016) 126–130.
[11] R.L. Keeney, H. Raiffa, Decision analysis with multiple conflicting objectives, John Wiley & Sons, New York, 1976.
[12] G.W. Fischer, Range sensitivity of attribute weights in multiattribute value models, Organ. Behav. Hum. Decis. Process. 62 (1995) 252–266.
[13] R. von Nitzsch, M. Weber, The effect of attribute ranges on weights in multiattribute utility measurements, Manage. Sci. 39 (1993) 937–943.
[14] G.W. Fischer, N. Damodaran, K.B. Laskey, D. Lincoln, Preferences for proxy attributes, Manage. Sci. 33 (1987) 198–214.

[15] T. Pajala, P. Korhonen, J. Wallenius, Judgments of importance revisited: What do they mean?, J Oper. Res. Soc. 70 (2019) 1140–1148.
[16] R.L. Keeney, Common mistakes in making value trade-offs, Oper. Res. 50 (2002) 935–945.
[17] A.T. de Almeida, J.A. de Almeida, A.P.C.S. Costa, A.T. de Almeida-Filho, A new method for elicitation of criteria weights in additive models: Flexible and interactive tradeoff, Eur. J. Oper. Res. 250 (2016) 179–191.
[18] F. Bacon, The new organon and related writings, Liberal Arts Press, New York, 1960.
[19] M. Weber, K. Borcherding, Behavioral influences on weight judgments in multiattribute decision making, Eur. J. Oper. Res. 67 (1993) 1–12.
[20] F. Eisenführ, M. Weber, T. Langer, Rational decision making, Springer, New York, 2010.
[21] J.S. Dyer, R.K. Sarin, Measurable multiattribute value functions, Oper. Res. 27 (1979) 810–822.
[22] P.C. Fishburn, Methods of estimating additive utilities, Manage. Sci. 13 (1967) 435–453.
[23] R. Vetschera, W. Weitzl, E. Wolfsteiner, Implausible alternatives in eliciting multi-attribute value functions, Eur. J. Oper. Res. 234 (2014) 221–230.
[24] F. Liang, M. Brunelli, J. Rezaei, Consistency issues in the best worst method: Measurements and thresholds, Omega 96 (2020) 102175.
[25] M.T. Escobar, J. Aguarón, J.M. Moreno-Jiménez, Some extensions of the precise consistency consensus matrix, Decis. Support Syst. 74 (2015) 67–77.
[26] B. Cavallo, L. D'Apuzzo, L. Basile, Weak consistency for ensuring priority vectors reliability, J. Multi-Criteria Decision Anal. 23 (2016) 126–138.
[27] M. Kendall, A new measure of rank correlation, Biometrika 30 (1938) 81–93.
[28] P. Amenta, A. Lucadamo, G. Marcarelli, On the transitivity and consistency approximated thresholds of some consistency indices for pairwise comparison matrices, Inf. Sci. 507 (2020) 274–287.
[29] R.H. Myers, D.C. Montgomery, C.M. Anderson-Cook, Response surface methodology: process and product optimization using designed experiments, John Wiley & Sons, 2016.
[30] J. Rezaei, L. van Wulfften Palthe, L. Tavasszy, B. Wiegmans, F. van der Laan, Port performance measurement in the context of port choice: an MCDA approach, Manag. Decis. 57 (2019) 396–417.
[31] A. Tversky, D. Kahneman, Judgment under uncertainty: Heuristics and biases, Science 185 (1974) 1124–1131.
[32] J.T. Buchanan, J. Corner, The effects of anchoring in interactive MCDM solution methods, Comput. Oper. Res. 24 (1997) 907–918.
[33] G. Montibeller, D. Von Winterfeldt, Cognitive and motivational biases in decision and risk analysis, Risk Anal. 35 (2015) 1230–1251.
[34] J. Rezaei, Anchoring bias in eliciting attribute weights and values in multi-attribute decision-making, J. Decision Syst. 30 (2021) 72–96.
[35] J. Rezaei, The balancing role of best and worst in Best-Worst Method. in The International Workshop on Best-Worst Method, Springer, Cham. 2021: 1-15.
[36] K. Borcherding, T. Eppel, D. Von Winterfeldt, Comparison of weighting judgments in multiattribute utility measurement, Manage. Sci. 37 (1991) 1603–1619.
[37] A.T. de Almeida, E.A. Frej, L.R.P. Roselli, Combining holistic and decomposition paradigms in preference modeling with the flexibility of FITradeoff, CEJOR 29 (2021) 7–47.
[38] P. Linares, S. Lumbreras, A. Santamaría, A. Veiga, How relevant is the lack of reciprocity in pairwise comparisons? An experiment with AHP, Ann. Oper. Res. 245 (2016) 227–244.
[39] A.A. Salo, R.P. Hämäläinen, Preference assessment by imprecise ratio statements, Oper. Res. 40 (1992) 1053–1061.
[40] C.W. Kirkwood, R.K. Sarin, Ranking with partial information: A method and an application, Oper. Res. 33 (1985) 38–48.
[41] R.L. Keeney, D. Von Winterfeldt, T. Eppel, Eliciting public values for complex policy decisions, Manage. Sci. 36 (1990) 1011–1030.
[42] V. Belton, P. Goodwin, Remarks on the application of the analytic hierarchy process to judgmental forecasting, Int. J. Forecast. 12 (1996) 155–161.
[43] H. Bleichrodt, J.L. Pinto, Loss aversion and scale compatibility in two-attribute trade-offs, J. Math. Psychol. 46 (2002) 315–337.
[44] T.J. Lahtinen, R.P. Hämäläinen, C. Jenytin, On preference elicitation processes which mitigate the accumulation of biases in multi-criteria decision analysis, Eur. J. Oper. Res. 282 (2020) 201–210.
[45] T.J. Lahtinen, R.P. Hämäläinen, Path dependence and biases in the even swaps decision analysis method, Eur. J. Oper. Res. 249 (2016) 890–898.
[46] J. Rezaei, A concentration ratio for nonlinear best worst method, Int. J. Inform. Technol. Decision Making 19 (2020) 891–907.