Belief-Based Best Worst Method

Liang, Fuqi; Brunelli, Matteo; Septian, Kevin; Rezaei, Jafar

DOI
10.1142/S0219622020500480

Publication date
2021

Document Version
Final published version

Published in
International Journal of Information Technology and Decision Making

International Journal of Information Technology & Decision Making
Vol. 20, No. 1 (2021) 287–320

DOI: 10.1142/S0219622020500480

Belief-Based Best Worst Method

Fuqi Liang[*,§], Matteo Brunelli[†,¶], Kevin Septian[‡,||] and Jafar Rezaei[*,**]

[*]*Faculty of Technology, Policy and Management*
Delft University of Technology, Delft, The Netherlands

[†]*Department of Industrial Engineering*
University of Trento, Trento, Italy

[‡]*Faculty of Civil Engineering & Geosciences*
Delft University of Technology
Delft, The Netherlands
[§]*f.liang-2@tudelft.nl*
[¶]*matteo.brunelli@unitn.it*
[||]*k.septian.92@gmail.com*
[**]*j.rezaei@tudelft.nl*

Published 3 February 2021

The Best-Worst Method (BWM) is a Multi-Criteria Decision Making (MCDM) method that has recently been introduced. The original BWM assumes that decision-makers are always certain about their judgments even if, in reality, decision-makers often express uncertain preferences. To deal with uncertainty, we introduce a belief structure in the BWM, a concept involving the preference degree adopted via Dempster-Shafer theory. A new approach is proposed to allow BWM to cope with this kind of information, where the level of belief in preferences being expressed is taken into account. In addition, an inconsistency measurement and an uncertainty measurement are proposed for the belief-based BWM, providing the foundation for a reliability degree of the decision-makers, after which the belief-based BWM is extended to include a group of decision-makers. Based on their reliability degrees and the weights of the criteria obtained from the various individuals, the overall criteria weights can be aggregated accordingly. Finally, a case study on the assessment of the infrastructure project criteria system in Indonesia is provided to demonstrate the applicability and feasibility of the proposed method.

Keywords: Best worst method; belief structure; inconsistency measurement; group decision-making.

1. Introduction

Multi-Criteria Decision-Making (MCDM) is an important area of operations research. It refers to finding an optimal result or ranking from a finite number of alternatives that are characterized in terms of multiple, usually conflicting, criteria.[1] There is a large and growing body of literature that has so far investigated MCDM methods.[2] One of the latest MCDM methods is the Best-Worst Method (BWM), proposed by Rezaei,[3] which uses pairwise comparisons to determine the weights of

[§]Corresponding author.

criteria. Thanks to its simplicity, flexibility and general applicability, since its inception, the BWM has been applied in a number of areas, including quality assessment,[4] supply chain management,[5,6] energy,[7] technology selection,[8] cloud service selection,[9] web service selection,[10] and hybrid vehicle engine selection.[11] In addition to its practical applications, many researchers have extended the BWM from a theoretical perspective as well. For example, since the original BWM can in some cases result in multi-optimality, Rezaei[12] proposed an interval weight analysis to deal with inconsistent comparisons with more than three criteria, as well as providing a linear BWM to generate a unique solution. Some researchers tried to combine subjective weights and objective weights together on the basis of BWM.[8,13] For a more exhaustive review, see the review study by Mi *et al.*,[14] and the bibliographical report.[a]

One of the critical issues in the BWM is the way it deals with uncertainty. Typically, there are three types of uncertainties, according to the summary by Klir and Wierman[15]: *fuzziness* (or vagueness), which results from the imprecise boundaries of fuzzy sets; *discord* (or strife), which expresses conflicts among the various sets of alternatives; and *nonspecificity* (or imprecision), which is connected to sizes (cardinalities) of relevant sets of alternatives. For example, a fuzzy set represents fuzziness, while a probability distribution represents only discord, and a classical set simply represents nonspecificity.[16] Although researchers have extended BWM to deal with uncertainty, most of them can only handle fuzziness.[11,13,17–20] A decision-maker (DM) who wants to provide his preferences with discord and nonspecificity cannot be handled properly in BWM. However, a belief structure defined in the Dempster–Shafer theory (D–S theory) framework[21] can handle both discord and nonspecificity.[16] Therefore, incorporating the belief system into the BWM will complement existing literature and make it possible to include these two types of uncertainty.

In D–S theory, subjective probabilities are replaced by "degrees of belief" within a belief structure, which can be used to express the extent to which a DM believes a specific proposition to be true.[22] Consider, for example, the comparison of the criteria price and quality in a sample involving cars, where a customer may state that he is 50% sure that the price is slightly more important than quality, 20% sure that the price is far more important than quality, and 30% sure that the price is extremely more important than quality. These "belief degrees" can be assigned to any subsets, making it possible to handle uncertainty and ignorance in a belief matrix. Such uncertainty and ignorance could be caused by imprecision in assessment, unfamiliarity with the problem at hand, a lack of data or the absence of certain stakeholders in a group decision.[23] Moreover, by using distribution assessment, the belief structure in question can capture precise data and as well as different types of uncertainties, such as probabilities and ambiguity in subjective judgments. As such, when modeling uncertainty by belief structure, D–S theory is more flexible and versatile than the traditional Bayes theory, where probabilities can only be assigned to individual hypotheses, instead of providing an explicit mechanism for dealing with ignorance.[24]

[a] From https://bestworstmethod.com/papers-and-slides/.

Belief structure was introduced to MCDM by Yang and Singh[25] in an Evidential Reasoning approach, since then, there have been a plethora of studies into the belief structure[24,26–29] and its extensions.[30,31] A recent study has tried to extend BWM to the belief structure,[32] however, since it uses pignistic probability function and weighted sum method to obtain an intermediate value, which is then used as input of the BWM, essentially speaking, it makes no change to the original BWM.

Next to uncertainty, complexity is another important issue that is considered in MCDM (including the BWM). In real-world decisions, it is difficult for a single DM to take all the relevant aspects of a decision-making problem into account. As a result, a group of DMs from different areas provides the advantages of synergy and information-sharing compared to the decisions that are made by a single individual. Thus, many of the decision-making processes that occur in the real world involve group settings designed to make the decision-making process more comprehensive and rational. Multi-Criteria Group Decision-Making (MCGDM), in which multiple DMs provide their evaluations regarding all the criteria of a decision-making problem, has been one of the most important and promising parts of modern decision-making theory.[33,34]

To date, several extensions of the BWM to group decision-making have been proposed.[11,17,35–41] However, in the existing group BWM approaches, the impact of the reliability of DMs is underestimated and rarely considered. The reliability of DMs in group BWM can be defined as their ability to provide a *certain* and *consistent* evaluation using pairwise comparisons. In the existing MCGDM research, the experts or DMs are usually assumed to be both rational and reliable. However, according to Simon,[42,43] our rationality is bounded due to our limited computational ability, selective memory and perception. As such, the judgements expressed by DMs in the BWM may be inconsistent and include some degree of uncertainty and imprecision.[37] Also, because the DMs reliability has a significant impact on the rationality and validity of the results,[44] neglecting it could lead to system accidents.[45] In other words, being able to measure that reliability effectively and apply it within the group aggregation process is significantly important to the group BWM.

The objective of this study is to incorporate information regarding the belief structure into BWM and enable the method to handle the opinions of a group of experts. Specifically, the belief structure preference is applied to pairwise comparisons and the original BWM is extended to handle that type of information. In order to solve the multiple optimal solutions problem of the nonlinear model, two models are used to obtain the boundary of the weights. Moreover, in order to check the reliability of DMs when they apply belief structure during the elicitation process, a reliability degree is defined based on the inconsistency and uncertainty levels of the DMs in question, and a group belief BWM framework is proposed. With the reliability degrees of DMs, the final weights of criteria can be determined by integrating the criteria and the weights obtained from each individual.

The remainder of this study is organized as follows. In Sec. 2, the original BWM is reviewed and the concept of belief structure is introduced. In Sec. 3, new BWM models are proposed to deal with belief structure preferences. In Sec. 4, a reliability measurement is proposed based on the inconsistency measurement (or consistency measurement) and the uncertainty measurement (or certainty measurement). The proposed method is then extended to include group decision-making problems, in Sec. 5. In Sec. 6, an application to the evaluation of the infrastructure project criteria system in Indonesia is provided to demonstrate the applicability and feasibility of the proposed method. Finally, some concluding remarks are presented in Sec. 7.

2. Preliminaries

In this section, we review the original BWM and discuss the basic terminology and definitions of the belief structure in D-S theory. The overall uncertainty measurement of belief structure, designed to measure random and nonspecific uncertainty, is also introduced.

2.1. *The original BWM*

As a weighting method based on pairwise comparisons, the BWM uses ratios of the relative importance of criteria in pairs, as estimated by a DM, from two evaluation vectors, the best criterion in relation to the other criteria, and the other criteria in relation to the worst criterion, whereby the weights of the criteria can be obtained by solving an optimization problem.[3] The basic steps of the original BWM can be summarized as follows:

Step 1. Determine the set of evaluation criteria $\{C_1, C_2, \ldots, C_n\}$.

Step 2. Determine the best (e.g. the most influential or the most important) and the worst (e.g. the least influential or the least important) criteria.

Step 3. Determine the preferences of the best criterion over all the other criteria, using a number between 1 to 9. The obtained Best-to-Others (BO) vector is: $A_{BO} = (a_{B1}, a_{B2}, \ldots, a_{Bn})$, where a_{Bj} represents the preference of the best criterion C_B over other criterion C_j, $j = 1, 2, \ldots, n$.

Step 4. Determine the preferences of all the criteria over the worst criterion. The obtained Others-to-Worst (OW) vector is: $A_{OW} = (a_{1W}, a_{2W}, \ldots, a_{nW})$, where a_{jW} represents the preference of other criterion C_j over the worst criterion C_W, $j = 1, 2, \ldots, n$.

Step 5. Determine the weights $(w_1^*, w_2^*, \ldots, w_n^*)$ by solving the following model:

$$\begin{gathered} \min\ \max_j \left\{ \left| \frac{w_B}{w_j} - a_{Bj} \right|, \left| \frac{w_j}{w_W} - a_{jW} \right| \right\} \\ \text{s.t.} \\ \sum_{j=1}^{n} w_j = 1, \\ w_j \geq 0, \qquad \text{for all} \quad j. \end{gathered} \tag{1}$$

Model (1) can be transferred into the following model:

$$
\begin{aligned}
&\min \xi \\
&\text{s.t.} \\
&\left|\frac{w_B}{w_j} - a_{Bj}\right| \le \xi, \qquad \text{for all} \quad j, \\
&\left|\frac{w_j}{w_W} - a_{jW}\right| \le \xi, \qquad \text{for all} \quad j, \\
&\sum_{j=1}^{n} w_j = 1 \\
&w_j \ge 0, \qquad \text{for all} \quad j.
\end{aligned} \tag{2}
$$

When the preferences are not fully consistent, the nonlinear model (2) usually generates multiple optimal solutions. Rezaei[12] proposed two models to derive interval weights, which include all the possible solutions, as well as a linear alternative designed to obtain a unique solution.

2.2. *Belief structure*

The basic concepts of belief structure are introduced in this part. The pignistic probability function and uncertainty measurement for belief structure are also discussed here, to be used at a later point.

2.2.1. *Basic terminology*

Suppose the DM is using a finite set of assessment grades $\Omega_{ij} = \{h_1, h_2, \ldots, h_k\}$ to express his preferences, which is commonly called the *frame of discernment* in the D–S theory. These grades are mutually exclusive and collectively exhaustive for all of the evaluations. The power set of Ω, which is the set of all the subsets of Ω, can be presented as

$$
\begin{aligned}
2^{\Omega} = \{H_l\} = \{H_1, H_2, \ldots, H_{2^K}\} = \{&\emptyset, \{h_1\}, \ldots, \{h_K\}, \{h_1, h_2\}, \ldots, \\
&\{h_1, h_K\}, \ldots, \{h_1, \ldots, h_{K-1}\}, \Omega\}, \qquad l = 1, \ldots, 2^K
\end{aligned}
$$

Definition 1 (Ref. 21). A *basic probability assignment* to all subsets H_l of 2^{Ω} is a function $m : 2^{\Omega} \to [0, 1]$, which satisfies $m(\emptyset) = 0$ and $\sum_{H_l \in 2^{\Omega}} m(H_l) = 1$.

The value $m(H_l)$ is assigned only to the set H_l and notto a smaller subset. Any subset H_l with $m(H_l) > 0$ is called a *focal element*. The set of all the focal elements is denoted with F. The pair $\langle F, m\rangle$ is called the body of evidence.

Based on the degree of belief, some other measures of confidence can be defined.

A *belief measure* is a function $\mathrm{Bel} : 2^{\Omega} \to [0, 1]$, which represents our confidence that the concerned element belongs to H or any of its subsets B and is defined by

$$
\mathrm{Bel}(H_l) = \sum_{B \subseteq H_l} m(B). \tag{3}
$$

A *plausibility measure* is a function $\text{Pls}: 2^{\Omega} \to [0,1]$, defined by

$$\text{Pls}(H_l) = \sum_{B \cap H_l \neq \emptyset} m(B), \tag{4}$$

$\text{Pls}(H_l)$ represents the extent to which we fail to disbelieve H_l. Thus, $\text{Bel}(H_l)$ and $\text{Pls}(H_l)$ can be interpreted as the lower and upper bound of probability to which H_l is supported.[46]

Definition 2 (Ref. 47). For an $m(H_l)$ on 2^{Ω}, its associated *pignistic probability function* $\beta_m : \Omega \to [0,1]$ is defined as

$$\beta_m = \sum_{H_l : h_k \in H_l} \frac{m(H_l)}{|H_l|}, \tag{5}$$

where $|H_l|$ is the cardinality of H_l.

The principle underlying the pignistic probability function is called the generalized insufficient reason principle, because the insufficient reason principle is used at the level of each focal element of the belief function. This pignistic probability can be interpreted as the degree of belief in each element of the frame of discernment Ω.

2.2.2. *Uncertainty measurement for belief structure*

A noteworthy uncertainty measure called the *Aggregated Uncertainty (AU) measure,*which was proposed by Harmanec and Klir[48] to quantify the total uncertainty of a belief function, is adopted in this paper to measure the uncertainty of the given preferences, because it can measure both discord and nonspecificity, and it satisfies all the basic requirements for a meaningful measure of aggregate uncertainty in evidence theory.

Definition 3 (Ref. 48). Let Ω be a finite frame of discernment, and Bel be a belief measure on Ω. The *Aggregated Uncertainty* AU associated with Bel is measured by

$$\text{AU(Bel)} = \max_{p_x \text{consistent with Bel}} \left[-\sum_{x \in \Omega} p_x \log_2 p_x \right], \tag{6}$$

where the maximum is taken over all distributions $\{p_x\}_{x \in \Omega}$ that are consistent with Bel, and $\{p_x\}_{x \in \Omega}$ should satisfy the following constraints:

$$\text{s.t.} \quad \begin{cases} p_x \in [0,1], & \forall x \in \Omega \\ \sum_{x \in \Omega} p_x = 1 & \\ \text{Bel}(A) \leq \sum_{x \in A} p_x \leq \text{Pls}(A), & \forall A \subseteq \Omega \end{cases}$$

As can be seen from the definition, AU is the maximum (upper) Shannon entropy of all probability distributions under the constraints according to the given basic belief assignments. This measure can capture both nonspecificity and discord, and it is a well-justified method to measure uncertainty within the D−S theory. It has been proven that AU satisfies a number of reasonable properties for uncertainty measures in evidential theory.[49]

3. The BWM with Belief Structure

In this section, a belief-based BWM is proposed to deal with uncertain information by using belief structures. Because the proposed method may generate multiple optimal solutions, we introduce a method to obtain the interval weights that can comprise all the possible solutions.

3.1. *Belief structure for pairwise comparison*

Suppose a finite set of assessments $\Omega = \{h_1, h_2, \dots, h_K\}$ is used by a DM to provide his pairwise comparison preferences, these assessments are assumed to be mutually exclusive.

In BWM, a set of 1–9 grades is usually defined to determine the preference of one criterion over another, to show their relative importance, serving as the frame of discernment

$$\Omega^{\text{IMPORTANCE}} = \{h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8, h_9\}.$$

Each element in this frame of discernment refers to a verbal judgment and a scale, as shown in Table 1.

After determining the frame of discernment, the DM can compare criteria C_i to C_j with subset $H_l(l = 1, \dots, 2^K)$ from 2^Ω to evaluate his preference and assign $m_{l,ij}$ (instead of calling this the *basic probability assignment*, we call it the *basic belief assignment* in BWM) to express his basic belief degree with regard to $H_l \subseteq \Omega$.

By using the pignistic probability function in Definition 2, the belief degree $\beta_{k,ij}$ (pignistic probability) associated with each grade $h_{k,ij}$ when comparing criteria C_i to C_j under the frame of discernment Ω can be obtained as follows:

$$\beta_{k,ij} = \sum_{H_l : h_{k,ij} \in H_l} \frac{m(H_l)}{|H_l|}. \tag{7}$$

Table 1. The linguistic terms and scales for the importance of pairwise comparisons.

Grades	Verbal description	Numerical values
h_1	Equally important	1
h_2	Equally to slightly more important	2
h_3	Slightly more important	3
h_4	Slightly to strongly more important	4
h_5	Strongly more important	5
h_6	Strongly to very strongly more important	6
h_7	Very strongly more important	7
h_8	Very strongly to extremely more important	8
h_9	Extremely more important	9

Then the pair of assessment of each grade $h_{k,ij}$ ($h_{k,ij} \in H_l$) and the *belief degree* $\beta_{k,ij}$ ($\langle h_{k,ij}, \beta_{k,ij} \rangle$) form the *body of assessment* (similar to the body of evidence in D–S theory), which can be profiled by a belief structure (denoted as S_{ij})):

$$S_{ij} = \{(h_{k,ij}, \beta_{k,ij}), k = 1, \dots, K\}. \tag{8}$$

Example 1. When a DM wants to buy a car and compares the relative importance of the criterion price over criterion style, suppose he decides to take $\Omega = \{h_1, h_2, h_3\} = \{$Equally important (h_1), equally to slightly more important (h_2), slightly more important $(h_3)\}$ as the frame of discernment, then he constructs his belief evaluations as

$$m : m\{\emptyset\} = 0, \qquad m\{h_1\} = 0, \qquad m\{h_2\} = 0, \qquad m\{h_3\} = 0.6, \qquad m\{h_1, h_2\} = 0,$$
$$m\{h_1, h_3\} = 0, \qquad m\{h_2, h_3\} = 0.1, \qquad m\{h_1, h_2, h_3\} = 0.3,$$

which means he is 60% sure that the price is *slightly more important than* style (grade h_3), 10% sure on grades h_2 and h_3, which leaves 30% belief for the remaining set, which represents his degree of ignorance.

According to Eq. (7), the belief degree (β_k) to each grade h_k can be computed as

$$\beta_1 = m\{h_1\} + \frac{m\{h_1, h_2\}}{2} + \frac{m\{h_1, h_3\}}{2} + \frac{m\{h_1, h_2, h_3\}}{3} = 0 + 0 + 0 + 0.1 = 0.1,$$
$$\beta_2 = m\{h_2\} + \frac{m\{h_1, h_2\}}{2} + \frac{m\{h_2, h_3\}}{2} + \frac{m\{h_1, h_2, h_3\}}{3} = 0 + 0 + 0.05 + 0.1 = 0.15,$$
$$\beta_3 = m\{h_3\} + \frac{m\{h_1, h_3\}}{2} + \frac{m\{h_2, h_3\}}{2} + \frac{m\{h_1, h_2, h_3\}}{3} = 0.6 + 0 + 0.05 + 0.1 = 0.75$$

Then the belief structure of comparing criterion price over criterion style can be constructed as

$$S_{\text{price,style}} = \{(h_1, \beta_1), (h_2, \beta_2), (h_3, \beta_3)\} = \{(h_1, 0.1), (h_2, 0.15), (h_3, 0.75)\}.$$

3.2. *The procedure of BWM with belief structure*

To incorporate the belief structure into the BWM, the model's procedure can be provided as follows:

Step 1. DM determines the set of evaluation criteria and the frame of discernment.

To evaluate an MCDM problem, the DM should identify the corresponding set of criteria to evaluate the performance of the alternatives involved. Here, we suppose there are n criteria $C = \{C_1, C_2, \dots, C_n\}$. A set of grades is identified by DMs to evaluate the pairwise comparisons, assuming that the frame of discernment consists of K grades: $\Omega = \{h_1, h_2, \dots, h_K\}$.

Step 2. DM selects the best (e.g. the most influential or the most important) and the worst (e.g. the least influential or the least important) criteria.

In this step, the DM is asked to identify the best and worst criteria, based on the criteria set. The best criterion is represented as C_B, the worst criterion as C_W.

Step 3. DM assigns the preference of the best criterion over all the other criteria, with basic belief assignments.

The DM needs to provide his preferences in comparing the best criterion C_B to the other criteria C_j under the set of identified assessment grade Ω. The entire subset H_l of 2^Ω will be complemented with the basic belief assignment $m_{l,Bj} \in [0,1]$. The subsets with $m_{l,Bj} > 0$ make up the body of assessment.

Step 4. DM assigns the preference of all the other criteria over the worst criterion, with basic belief assignments.

The DMs assigns basic belief scores ($m_{l,jW}$) to the entire subset H_l of 2^Ω when comparing the other criteria C_j to the worst criterion C_W. The body of assessment is made up by the subsets with $m_{l,jW} > 0$.

Step 5. Construct belief structures according to the pignistic probability function.

Determine the belief degree $\beta_{k,Bj}$ to each grade $h_{k,Bj}$ by using the pignistic probability function Eq. (7), after which the belief structure involved in comparing the best criterion to the others can be constructed as

$$S_{Bj} = \{(h_{k,Bj}, \beta_{k,Bj}), k = 1, \ldots, K\}.$$

The resulting Best-to-Others (BO) vector is: $S_B = (S_{B1}, S_{B2}, \ldots, S_{Bn})$, where S_{Bj} represents the preference of the best criterion C_B over the other criterion $C_j, j = 1, 2, \ldots, n$.

Similarly, the belief structure of comparing the others to the worst criterion can be constructed as

$$S_{jW} = \{(h_{k,jW}, \beta_{k,jW}), k = 1, \ldots, K\}.$$

The resulting Others-to-Worst (OW) vector is: $S_W = (S_{1W}, S_{2W}, \ldots, S_{nW})$, where S_{jW} represents the preference of other criterion C_j over the worst criterion C_W, $j = 1, 2, \ldots, n$.

Step 6. Determine the weights$(w_1^*, w_2^*, \ldots, w_n^*)$.

To determine the optimal weights with respect to a belief structure, we need to make each pair of w_B/w_j and w_j/w_W as close as possible to the grade $h_{k,Bj}^*$ ($h_{k,jW}^*$) with the maximum belief degree $\beta_{k,Bj}^*$ ($\beta_{k,jW}^*$) in the corresponding belief structure S_{Bj} (S_{jW}). The underlying idea is that the grade with the higher belief score should be valued more, and the grade with the lower belief score should be valued less. To operate this idea for all j, the maximum difference between $\frac{w_B}{w_j}$ and $h_{k,Bj}^*$ ($\frac{w_j}{w_W}$ and $h_{k,jW}^*$) for all j should be minimized, which means that the constrained optimization problem to determine the optimal weights is constructed as follows:

$$\begin{aligned}
&\min\max\left\{|\frac{w_B}{w_j} - h_{k,Bj}|\beta_{k,Bj}, |\frac{w_j}{w_W} - h_{k,jW}|\beta_{k,jW}\right\} \\
&\text{s.t.} \\
&\sum_{j=1}^{n} w_j = 1 \\
&w_j \geq 0, \qquad \text{for all} \quad j
\end{aligned} \tag{9}$$

Model (9) can be transferred into the following model:

$$
\begin{aligned}
&\min \xi \\
&\text{s.t.} \\
&\left|\frac{w_B}{w_j} - h_{k,Bj}\right|\beta_{k,Bj} \leq \xi, \qquad \text{for all} \quad j \quad \text{and} \quad k \\
&\left|\frac{w_j}{w_W} - h_{k,jW}\right|\beta_{k,jW} \leq \xi, \qquad \text{for all} \quad j \quad \text{and} \quad k \\
&\sum_{j=1}^{n} w_j = 1 \\
&w_j \geq 0, \qquad \text{for all} \quad j
\end{aligned}
\tag{10}
$$

By solving problem (10), the optimal weights $(w_1^*, w_2^*, \ldots, w_n^*)$ are obtained. The optimal value ξ^* obtained from this program indicates that the closer it is to 0, the more consistent the DM is.

Example 2. We use the same case that was studied by Rezaei[12] and suppose that the frame of discernment is $\Omega = \{h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8\}$ (Step 1). The second criterion, Price (C_2), is identified as the best criterion, and the fifth criterion Style (C_5) is identified as the worst criterion (Step 2). Next, the DM provides his basic belief assignments (only values for focal elements are listed) with regard to the best criterion compared to the others, and the other criteria compared to the worst, as seen in Tables 2 and 3 (Steps 3 and 4):

After applying the pignistic probability function (7), the basic belief assignments can be transformed into belief structures (Step 5):

$$
\begin{aligned}
S_{21} &= \{(h_2, 0.5), (h_3, 0.5)\}, \\
S_{22} &= \{(h_1, 1)\}, \\
S_{23} &= \{(h_1, 0.0125), (h_2, 0.0458), (h_3, 0.1058), (h_4, 0.2725), \\
&\quad\ (h_5, 0.2392), (h_6, 0.2392), (h_7, 0.0725), (h_8, 0.0125)\}, \\
S_{24} &= \{(h_2, 0.8), (h_3, 0.2)\}, \\
S_{25} &= \{(h_6, 0.0667), (h_7, 0.0667), (h_8, 0.8667)\}, \\
S_{15} &= \{(h_4, 1)\}, \\
S_{25} &= \{(h_1, 0.025), (h_2, 0.025), (h_3, 0.025), (h_4, 0.025), (h_5, 0.025), \\
&\quad\ (h_6, 0.025), (h_7, 0.325), (h_8, 0.525)\}, \\
S_{35} &= \{(h_2, 0.2333), (h_3, 0.5333), (h_4, 0.2333)\}, \\
S_{45} &= \{(h_3, 0.0667), (h_4, 0.8667), (h_5, 0.0667)\}, \\
S_{55} &= \{(h_1, 1)\}.
\end{aligned}
$$

Figure 1 visualizes the distribution of the belief degrees involving each individual grade. For example, the belief structure s_{21} has 0.5 belief degree on grade 2 and grade 3, respectively.

Table 2. Assessments of the Best criterion to the others.

Best to others	Best criterion: C_2
Quality (C_1)	$m_{21}\{h_2, h_3\} = 1$
Price (C_2)	$m_{22}\{h_1\} = 1$
Comfort (C_3)	$m_{23}\{h_3, h_4, h_5, h_6, h_7\} = 0.3, m_{23}\{h_4, h_5, h_6\} = 0.5,\ m_{23}\{h_2, h_3, h_4\} = 0.1, m_{23}\{\Omega\} = 0.1$
Safety (C_4)	$m_{24}\{h_2\} = 0.6, m_{24}\{h_2, h_3\} = 0.4$
Style (C_5)	$m_{25}\{h_8\} = 0.8, m_{25}\{h_6, h_7, h_8\} = 0.2$

Table 3. Assessments of the other criteria to the Worst.

Others to worst	Worst criterion: C_5
Quality (C_1)	$m_{15}\{h_4\} = 1$
Price (C_2)	$m_{25}\{h_7, h_8\} = 0.6, m_{25}\{h_8\} = 0.2, m_{25}\{\Omega\} = 0.2$
Comfort (C_3)	$m_{35}\{h_2, h_3, h_4\} = 0.7, m_{35}\{h_3\} = 0.3$
Safety (C_4)	$m_{45}\{h_4\} = 0.8, m_{45}\{h_3, h_4, h_5\} = 0.2$
Style (C_5)	$m_{55}\{h_1\} = 1$

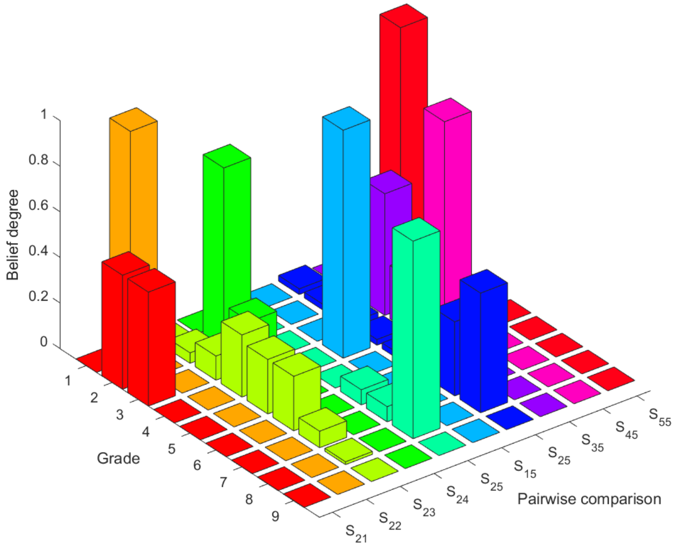

Fig. 1. The distribution of belief degrees in Example 2.

By solving the optimization problem, we can obtain one set of the optimal weights and ξ^* as follows (Step 6):

$$w_1^* = 0.1961,\ \ w_2^* = 0.4528,\ \ w_3^* = 0.1128,\ \ w_4^* = 0.1847,\ \ w_5^* = 0.0535,\ \text{ and }\ \xi^* = 0.4753.$$

The multiple optimal solutions issue is addressed in Sec. 3.3.

From the results, we can give another interpretation to the belief-based BWM. For instance, in the assessment C_2 over C_1, $m_{21}\{h_2, h_3\} = 1$, $S_{21} = \{\ (h_2, 0.5), (h_3, 0.5)\}$, the DM hesitates between h_2 and h_3, and the result $a^*_{21} = \frac{w^*_2}{w^*_1} = 2.5$ can capture this hesitation, since it lies in the middle. Also, for assessment C_4 over C_5, the basic belief assignment is $m_{45}\{h_4\} = 0.8$, $m_{45}\{h_3, h_4, h_5\} = 0.2$, so we expect the result can focus more on h_4 because the DM has expressed greater belief and certainty, instead of h_3 and h_5. The result of C_4 over C_5 is $a^*_{45} = \frac{w^*_4}{w^*_5} = 4$, which shows that it weighs more the strongest belief h_4.

The algorithm and analysis present the features of the belief-based BWM. The method not only allows a DM to provide his basic belief assignments in a more flexible way, it also balances the hesitation of the DM, taking all the preferences and beliefs into account and trying to come closer to the preferences with stronger beliefs and move further away from preferences associated with weaker beliefs.

If each belief structure provided by a DM is 100% sure on one single grade, that would mean the DM has no uncertainty at all, and this belief structure-based BWM in essence becomes the original BWM.

3.3. *Models to derive interval weights*

The nonlinear BWM can have multiple optimal solutions when the pairwise comparisons are not fully consistent. In order to handle that problem, we propose a method to obtain the minimum and maximum weights of each criterion. Two models are proposed to calculate the lower and upper bounds of the weights of criterion C_j based on the ξ^*, that is the optimal solution of models (9) and (10).

$$\begin{aligned} &\min \quad w_j \\ &\text{s.t.} \\ &\left|\frac{w_B}{w_j} - h_{k,Bj}\right| \beta_{k,Bj} \leq \xi^*, \qquad \text{for all} \quad j \quad \text{and} \quad k \\ &\left|\frac{w_j}{w_W} - h_{k,jW}\right| \beta_{k,jW} \leq \xi^*, \quad \text{for all} \quad j \quad \text{and} \quad k \\ &\sum_{j=1}^{n} w_j = 1 \\ &w_j \geq 0, \qquad \text{for all} \quad j \end{aligned} \tag{11}$$

$$\begin{aligned} &\max \quad w_j \\ &\text{s.t.} \\ &\left|\frac{w_B}{w_j} - h_{k,Bj}\right| \beta_{k,Bj} \leq \xi^*, \qquad \text{for all} \quad j \quad \text{and} \quad k \\ &\left|\frac{w_j}{w_W} - h_{k,jW}\right| \beta_{k,jW} \leq \xi^*, \quad \text{for all} \quad j \quad \text{and} \quad k \\ &\sum_{j=1}^{n} w_j = 1 \\ &w_j \geq 0, \qquad \text{for all} \quad j \end{aligned} \tag{12}$$

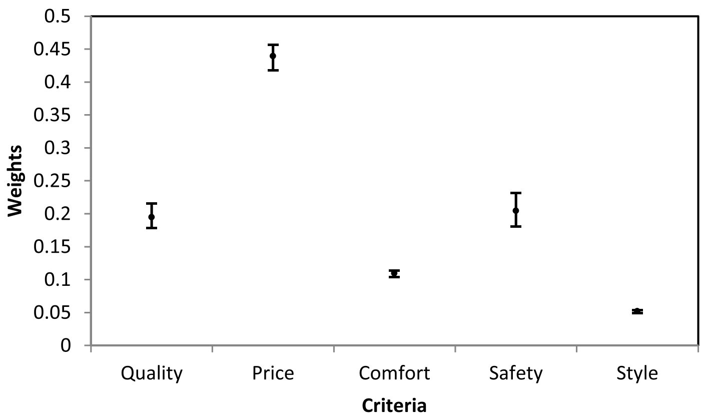

Fig. 2. The interval weights of belief BWM.

After solving these two models for all criteria, the optimal value of the objective function of (9) is taken as the minimum w_j^{*-} and, similarly, the optimal value of (10) is the maximum w_j^{*+}. Together, they identify intervals $[w_j^{*-}, w_j^{*+}]$. For the operations of interval weights and the method of ranking the criteria, the reader might refer to Ref. 12.

Example 3. From Example 2, we obtain $\xi^* = 0.4753$, which indicates that the system of pairwise comparisons is not fully consistent, and the nonlinear belief-based BWM model can generate multiple optimal solutions. To solve that problem, we use the interval weights to contain all the possible solutions. The optimal interval weights of belief-based BWM obtained, thanks to the optimization problems (11) and (12), are as follows:

$$w_1^* = [0.1784, 0.2156], \qquad w_2^* = [0.4179, 0.4563], \qquad w_3^* = [0.1041, 0.1137],$$
$$w_4^* = [0.1802, 0.2315], \qquad w_5^* = [0.0494, 0.0539].$$

The mean of all the optimal intervals can be used to indicate the middle position of these interval weights, the result being: $w_1^*(\text{mean}) = 0.1947$, $w_2^*(\text{mean}) = 0.4394$, $w_3^*(\text{mean}) = 0.1095$, $w_4^*(\text{mean}) = 0.2044$, $w_5^*(\text{mean}) = 0.0519$. The interval weights and their means are shown in Fig. 2.

4. The Reliability Measurement

After determining the weights, it is very important to check the reliability of the results. It has been a long debate on the measurement of the reliability or expertise of an expert/DM, especially when there is no external standard to verify.[50] Traditionally, the reliability of an expert is measured by the consensus with the other experts.[44,51] However, according to psychological investigations and empirical studies,[50,52] the agreement with other experts is neither necessary nor sufficient for expertise, rather, intra-individual consistency is a necessity.

Besides, the uncertainty degree of an expert is also highly related to his/her reliability. For example, if an expert provides a preference profile like {(1,0.5), (9,0.5)}, or like the highly nonspecific belief distributions {({1,2,3,4,5},0.5), ({5,6,7,8,9},0.5)}, the expert faces randomness and the nonspecificity problems.[15,53] Both cases could yield unreliable results, because the expert essentially has not provided sufficient information for a decision.

Therefore, in this section, we discuss a method designed to measure the reliability degree of an expert's judgments stemming from his inconsistency and uncertainty levels. To that end, an inconsistency measurement and an uncertainty measurement are proposed based on belief structure-based BWM.

4.1. *The inconsistency measurement for belief BWM*

The original BWM uses pairwise comparisons of criteria based on DMs' evaluations of the relative priorities of decision-making elements. As such, the pairwise comparisons are said to be perfectly (cardinal-) consistent if they satisfy the transitivity condition $a_{Bj} \times a_{jW} = a_{BW}$; otherwise, the DM is not fully consistent, which may imply some irrationality in the relative weight estimates.[54]

In belief-based BWM, to handle the information of belief structures, the utility-based approach[55] can be adopted to compute the value of belief structures. The expected utility of a belief structure S_{ij} is noted as u_{ij}, and can be computed as follows:

$$u_{ij} = \sum_{k=1}^{K} u(h_k)\beta_{k,ij}, \tag{13}$$

where $u(h_k) = k$. Then the value of a_{ij} in the original BWM can be replaced by the expected utility u_{ij}, thus the transitivity condition is transformed into

$$u_{Bj} \times u_{jW} = u_{BW}. \tag{14}$$

According to the definition of belief structure, suppose the DM identifies a set of evaluation grades $\Omega = \{h_1, h_2, \ldots, h_K\}$ which is applied to pairwise comparisons, then $S_{BW} = (h_K, 1)$ is the maximum belief structure that can generate the highest possible value to u_{BW}. If $u_{Bj} \times u_{jW} \neq u_{BW}$, the inconsistency will occur, whether $u_{Bj} \times u_{jW}$ is higher or lower than u_{BW}. When u_{Bj} and u_{jW} have the highest value, which is equal to u_{BW}, that will result in the largest inequality. According to $(w_B/w_j) \times (w_j/w_W) = w_B/w_W$, the following equation can be obtained:

$$(u_{Bj} - \xi) \times (u_{jW} - \xi) = u_{BW} + \xi. \tag{15}$$

For the maximum inconsistency of belief structure, $u_{Bj} = u_{jW} = u_{BW}$, Eq. (16) can be written as

$$(u_{\mathrm{BW}} - \xi) \times (u_{BW} - \xi) = u_{\mathrm{BW}} + \xi, \tag{16}$$

Table 4. Inconsistency index table.

K	1	2	3	4	5	6	7	8	9
Inconsistency index	0	0.44	1	1.63	2.30	3	3.73	4.47	5.23

and formulated as

$$\xi^2 - (1 + 2u_{BW})\xi + (u_{BW}^2 - u_{BW}) = 0. \tag{17}$$

Because $u_{BW} = u(K) = K$, $K \in \{1, 2, 3, \ldots\}$, Eq. (17) becomes

$$\xi^2 - (1 + 2K)\xi + (K^2 - K) = 0. \tag{18}$$

After solving Eq. (18) for different K, the maximum possible ξ can be obtained and used as the inconsistency index for belief-based BWM. The result of the inconsistency index is shown in Table 4. The inconsistency index obtained for belief-based BWM is the same as the original BWM, because the u_{BW} in the belief BWM is the same as a_{BW} in the original BWM.

We can now use the ξ^* obtained from the belief-based BWM models (10) to calculate the following Inconsistency Ratio (IR)[b]:

$$\mathrm{IR} = \frac{\xi^*}{\text{Inconsistency Index}}, \tag{19}$$

$\mathrm{IR} \in [0, 1]$, and the closer IR is to 0, the more consistent the judgments are. When $IR = 0$, the judgments of a DM are said to be *fully consistent*.

4.2. *The uncertainty measurement for belief BWM*

We stated earlier that the advantage of the belief-based BWM is the way it deals with uncertain preferences. However, it is important to quantify this very same uncertainty as it can be related to the reliability and the stability of the final results. The final goal of such an analysis would be to identify excessively uncertain preferences.

The measure of uncertainty for the belief-based BWM can be formulated as

$$\begin{aligned} \mathrm{AU}(\mathrm{Bel}) = &\max_{p_{k,ij}\text{consistent with Bel}} \left[-\sum_{k\in\Omega} p_{k,ij} \log_2 p_{k,ij} \right], \\ &\text{s.t.} \quad \begin{cases} p_{k,ij} \in [0,1], & \forall k \in \Omega \\ \sum_{k\in\Omega} p_{k,ij} = 1 & \\ \mathrm{Bel}(H_l) \le \sum_{k\in H_l} p_{k,ij} \le Pls(H_l), & \forall H_l \subseteq \Omega \end{cases} \end{aligned} \tag{20}$$

4.2.1. *The uncertainty measure algorithm for belief-based BWM*

To compute the AU function, an algorithm was proposed by Harmanec *et al.*,[56] which, in spite of being proved to be correct by Klir and Wierman,[15] is too complex in

[b] In the case $K = 1$, the preferences are always fully consistent, hence the IR is zero.

Table 5. The algorithm of uncertainty measurement.

Input: The set of focal elements F of belief function Bel and their corresponding basic belief assignments.
Output: AU(Bel), $\{p_k\}_{k\in\Omega}$ such that $\mathrm{AU(Bel)} = -\sum_{k\in\Omega} p_{k,Bj} \log_2 p_{k,Bj}$ and $\mathrm{AU(Bel)} = -\sum_{k\in\Omega} p_{k,jW} \log_2 p_{k,jW}$.
(1) Initialize $\mathrm{AU(Bel)} = 0$.
(2) Compute the belief measures for all elements of $U(F)$, which is the union of the focal elements from F.
(3) Find a set $H_l \in U(F)$, $(l = 1, \ldots, 2^K)$ such that $\mathrm{Bel}(H_l)/|H_l|$ is maximal. If there is more than one such set H_l, the one with the largest cardinality should be selected.
(4) For $k \in H_l$, put $p_{k,Bj} = \mathrm{Bel}(H_l)/|H_l|$ and $p_{k,jW} = \mathrm{Bel}(H_l)/|H_l|$; calculate $\mathrm{AU(Bel)} := \mathrm{AU(Bel)} - \mathrm{Bel}(H_l) \times \log_2 p_{k,Bj}$ and $\mathrm{AU(Bel)} := \mathrm{AU(Bel)} - \mathrm{Bel}(H_l) \times \log_2 p_{k,jW}$.
(5) Set $F' = \{H_f \backslash H_l | H_f \in F\} \backslash \{\emptyset\}$.
   (1) If $F' = \emptyset$, stop.
   (2) Otherwise, for each $S \in F'$, put
   $m(S) = \sum_{H_f \in F, H_f \backslash H_l = S} m(H)$ and set $F = F'$.
(6) If $|F| > 1$, return to step 2.
(7) If $|F| = 1$ and $F = \{S\}$, put $p_{k,Bj} = m(S)/|S|$ (or $p_{k,jW} = m(S)/|S|$) and $\mathrm{AU(Bel)} := \mathrm{AU(Bel)} - m(S) \times \log_2 p_{k,Bj}$ and $\mathrm{AU(Bel)} := \mathrm{AU(Bel)} - m(S) \times \log_2 p_{k,jW}$.

some cases, and it is why Liu *et al.*[57] proposed using another algorithm to reduce the computational complexity, which, unfortunately, was flawed, and it was subsequently corrected by Huynh and Nakamori[58] with an improved algorithm. This uncertainty measure for belief structure-based BWM uses Huynh and Nakamori's algorithm,[58] which is presented in its adapted form in Table 5.

4.2.2. *Global uncertainty*

The AU measure is used to quantify the total uncertainty of a given belief structure. To measure the global uncertainty of a DM, we need to take all the basic belief assignments into consideration. The DM's global uncertainty can be calculated as the average uncertainty of the given preferences

$$\overline{\mathrm{AU}} = \frac{1}{2n-3}\left(\max_{p_{k,Bj}\text{consistent with Bel}}\left[-\sum_{k\in\Omega} p_{k,Bj}\log_2 p_{k,Bj}\right] + \max_{p_{k,jW}\text{consistent with Bel}}\left[-\sum_{k\in\Omega} p_{k,jW}\log_2 p_{k,jW}\right]\right). \tag{21}$$

To compare the uncertainty degrees of different frames of discernment with different grades, we need to normalize the uncertainty degrees in the interval [0,1]. As the maximum value of $\overline{\mathrm{AU}}$ is $\log_2 K$, where K is the cardinality of discernment frame, the normalization of $\overline{\mathrm{AU}}$ can be formulated as follows:

$$\tilde{\mathrm{AU}} = \frac{\overline{\mathrm{AU}}}{\log_2 K}. \tag{22}$$

The range of $\widetilde{\mathrm{AU}}$ is $[0, 1]$, the closer $\widetilde{\mathrm{AU}}$ is to 0, the more certain the judgments are. When $\widetilde{\mathrm{AU}} = 0$, the judgments of a DM are said to be *fully certain*.

4.3. *The reliability degree*

The original BWM considers the reliability of a DM's assessments only through his inconsistency level, regardless of whether they use certain numbers or uncertain terms. However, highly uncertain judgments are unstable and lead to unreliable results. Therefore, in addition to looking at the inconsistency level, the uncertainty level also has to be taken into account to determine the reliability of a DM's judgments. In light of these considerations, we define the following reliability index.

Definition 4. The pairwise comparisons of a DM are said to be *fully reliable* if they are fully consistent and completely certain. The Reliability Degree (RD) of a DM's judgments can be formulated as

$$\mathrm{RD} = 1 - \frac{\sqrt{(IR)^2 + (\widetilde{AU})^2}}{\sqrt{2}}. \tag{23}$$

The RD ranges from 0 to 1, and when it is closer to 1, we say that the pairwise comparisons provided by this DM are more reliable because they are more consistent and more certain, as illustrated in Fig. 3. When $RD = 1$, the DM is considered to be fully reliable.

Unlike other formulations, e.g. $(\mathrm{IR} + \widetilde{\mathrm{AU}})/2$, our proposed formula for RD has a clear geometric interpretation: it is the *distance* from the point $(IR, \widetilde{\mathrm{AU}})$ to $(1, 1)$.

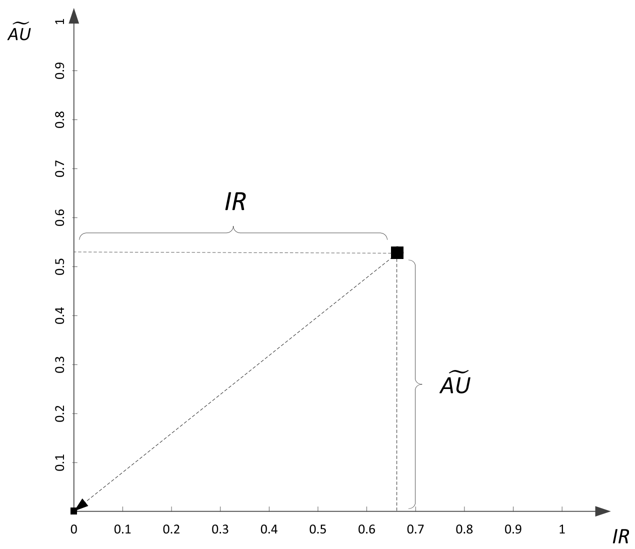

Fig. 3. The illustration of reliability.

In case of considering the expertise of DM as part of the reliability degree, we can use a generalized form: *Generalized reliability degree* $= \alpha\text{RD} + (1 - \alpha)$*expertise degree*, where *expertise degree* and $\alpha \in [0, 1]$.

5. Group BWM with Belief Structure

Due to the complexity of MCDM problems, it is common for several experts from different fields to form a group to assess the problems together. In addition, if the problems involve more than one stakeholder or multiple decision-makers, a group decision-making method is needed to aggregate individual preferences.

The existing aggregation methods for group-based BWM rarely take the reliability level of the DMs' judgments into account. As discussed in Sec. 4, we assume that the inconsistency level and uncertainty level contribute equally to a DM's reliability level. We propose an aggregation method for the group-and belief structure-based BWM, which uses the reliability degrees to determine suitable weights for the DMs.

We can extend the belief-based BWM proposed in Sec. 3 to multi-criteria group decision-making problems. We assume that the DMs express their preferences honestly, which means that the preferences reflect their inconsistency and uncertainty levels. The procedure is illustrated below and the flowchart of the steps involved is shown in Fig. 4.

Step 1. The group of DMs $D = \{D_1, D_2, \ldots, D_G\}$ negotiate and determine the set of evaluation criteria $\{C_1, C_2, \ldots, C_n\}$ and the frame of discernment, which contains K grades: $\Omega = \{h_1, h_2, \ldots, h_K\}$.

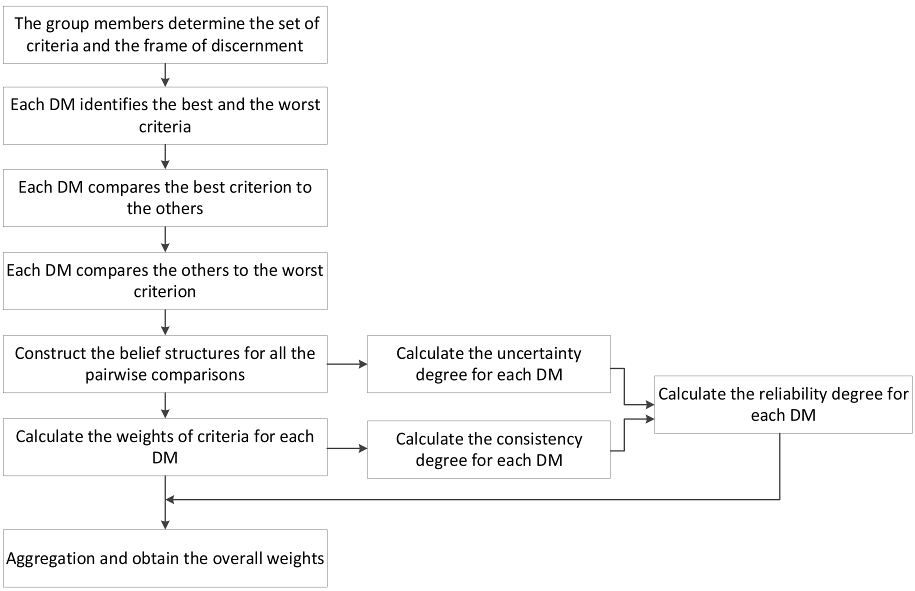

Fig. 4. The procedure of group BWM with belief structure.

Step 2. Each DM D_g determines his best and worst criteria (C_B^g and C_W^g, respectively).

Step 3. Each DM D_g assesses the best criterion over all the other criteria with basic belief assignments.

Step 4. Each DM D_g assesses all the other criteria over the worst criterion with basic belief assignments.

Step 5. Construct belief structures according to the pignistic probability function in Eq. (7).

Step 6. The nonlinear belief-based program (10) and the two decomposed models (11) and (12) in Sec. 3 are used to find the optimal criteria weights $w_j^g = \{w_1^g, w_2^g, \ldots, w_n^g\}$ for each DM D_g.

Step 7. From the preferences that have been provided, each DM D_g can obtain his IR^g and uncertainty degree $\tilde{\mathrm{AU}}^g$ by using the consistency measurement and uncertainty measurement discussed in Sec. 4.

Step 8. The weight of each DM is assumed to be a function of his reliability degree (s) RD^g obtained by Eq. (23). Under this assumption, we suggest deriving the weight of each DM D_g (λ^g) by means of

$$\lambda^g = \frac{\mathrm{RD}^g}{\sum_{g=1}^{G} \mathrm{RD}^g}. \tag{24}$$

Step 9. Aggregate all the criteria weights from each DM D_g into an overall weight $\tilde{w}_j$, which can be calculated by

$$\tilde{w}_j = \sum_{g=1}^{G} \lambda^g w_j^g. \tag{25}$$

6. Case Study: Application of the Proposed Method for Evaluating Infrastructure Project Criteria System in Indonesia

As one of the new emerging markets, Indonesia is striving to boost its economic development by making efforts to accelerate strategic projects which can be realized within a short period of time. Each of these projects and programs has its own objectives and responsibility, but due to the lack of coordination between various stakeholders in government and private sectors, there is potential to cause delay to the implementation.[c] Therefore, to deal with this problem, the Committee for Acceleration of Priority Infrastructure Delivery (KPPIP, shortly in Indonesian) was established. The mission of KPPIP is to screen and select the National Strategic Projects, and carry out monitoring activities for National Strategic Projects, as well as to conduct high-level debottlenecking strategies for Priority Projects.[59]

Before providing coordination in debottlenecking efforts for the 247 National Strategic Projects and programs, due to limited resources, KPPIP should shortlist 37

[c] The basic information of this case study is from KPPIP's website: https://kppip.go.id/en/.

Fig. 5. List of 37 KPPIP priority projects.[59]

projects as priority projects in line with the criteria established by KPPIP (Fig. 5). KPPIP will then monitor the shortlisted projects and ensure that they comply with quality standards and regulations. This case study will focus only on determining the importance level of the established criteria, not considering the monitoring and implementation part.

To support the decision-making process, KPPIP is equipped with a Project Management Office (PMO), which comprises of professional experts in their respective fields. These experts are responsible for providing recommendations to the implementation team in selecting priority projects. To evaluate various infrastructure projects, four sectors are formed in KPPIP, i.e. Energy and Electricity sector (EE), Road and Bridge sector (RB), Transportation sector (TT), Water and Sanitation sector (WS). The organizational structure of KPPIP can be seen in Fig. 6.

The National Strategic Projects are complex to evaluate. After discussion with the experts, 20 criteria were identified by KPPIP and classified into four categories as shown in Table 6. Almost all of these criteria were assigned equal weights initially (Executive Direction 0.08, Issuance of project permits and Number of authorities involved are 0.12, the others are all 0.04), which is unreasonable according to an interview we conducted with the leader of KPPIP. In addition to the arbitrariness, the assignment of weights to the criteria did not consider the variety of the four different sectors and the reliability degree of the experts in these sectors. Therefore, the weights of the criteria were suggested to be reevaluated by KPPIP with a more structured/analytic methodology.

In this study, we invited four experts from the four sectors (one from each sector - EE, RB, TT, WS) in KPPIP to reevaluate the importance of the given criteria by using the proposed method. They are asked to follow Steps 1 to 4 of the belief-based BWM in Sec. 5 and provided their assessments. Table 7 presents the pairwise comparison assessment for the main categories from the four sectors in KPPIP. Table 8–11 show the assessment for all the criteria in each category from the four

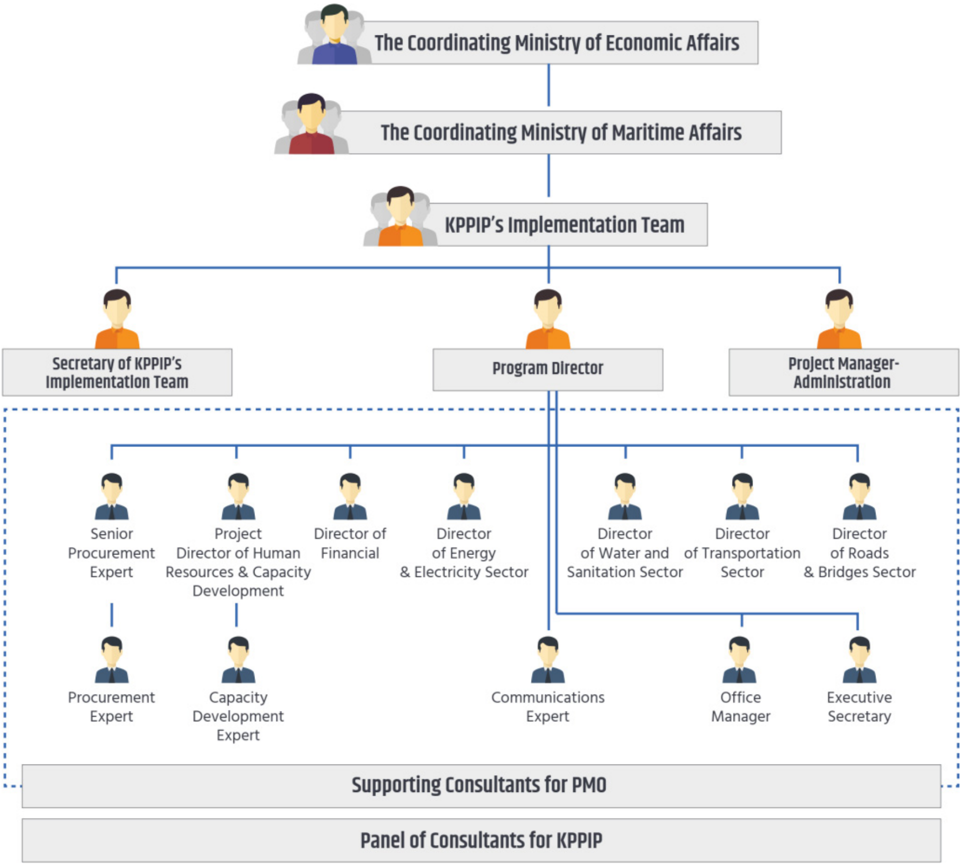

Fig. 6. The organizational structure of KPPIP.[59]

KPPIP sectors. In this case, the experts were suggested to evaluate the criteria by assigning each of the basic belief assignments to only one grade (for the sake of simplicity), and the unassigned degree represents ignorance. For example, the assessment {(2; 0,3), (3; 0,7)} and {(5; 0,5), (Ω; 0,5)} in the bottom right in Table 7, can be interpreted as: the Water and Sanitation sector compared the best category, which is Project Preparation (PP), to Policy (P) with 30% confidence that Project Preparation is equally to slightly more important (grade h_2) than Policy, and 70% confidence that Project Preparation is slightly more important (grade h_3) than Policy; This sector compared Policy to the worst category, which is Coordination (C), with 50% confidence thatPolicy is strongly more important (grade h_5) than Coordination, and the remaining 50% allocated to ignorance.

The nonlinear belief-based BWM is used in this case to determine the weights of the main categories for each sector, the interval weights for the main categories obtained from model (11) and (12) are shown in Table 12 (Steps 5 and 6). In this table, the IR_s and uncertainty degrees of each sector are obtained following the

Table 6. The criteria identified by KPPIP.

Category	Criteria	Definition
Project Preparation (PP)	PP1: Outline business case comprehensiveness	Preliminary thoughts, which contains the information, such as outcomes, benefits, and potential risks associated with the proposal.
	PP2: Economic benefits	It considers the benefit to the economy, environmental, and also social.
	PP3: Technical planning complexity	It considers environments that can bring the project into a complex development, such as land-use plan, environmental dispute, and relocation.
	PP4: Project development fund support	A programmatic approach to the funding of the cost for early tasks to encourage contracting agencies to use best practices.
	PP5: Infrastructure readiness/requirement surrounding the project	The government intends to accelerate the development in the country; it is implemented by integrating infrastructures that can carry out or perform more economic activities in the society surrounding the project.
Funding (F)	F1: Acquisition of interest from the investor(s)	Investors are one of the primary sources of funding for the project that is required to develop multiple projects
	F2: Determination of funding scheme	It shows the scheme of the funding which considers the strategic issue on the availability of the investors' interest.
	F3: Funding resources synchronization	It is a body provided by the central government when it is needed to assist the team in organizing the funding of the project.
	F4: Public sector organization structuration	Public Sector Organization is an entity that is formed to manage the policy and operating requirements that enable a government to achieve its goals of public governance.
	F5: Granting of credit risk	Credit risk is the possibility of a loss resulting from a debtor's failure to meet the obligations. It measures the availability of the assurance of the projects.
	F6: Granting of business feasibility support	The business feasibility support refers to the availability of elements, which support the continuity of the project development.

Table 6. (*Continued*)

Category	Criteria	Definition
Coordination (C)	C1: Stakeholder buy-in	Process of involving all the related stakeholders to reach consensus.
	C2: Land acquisition coordination	Most infrastructure projects mostly involve many areas to be cleared for the project and need complicated coordination.
	C3: Spatial plan synchronization	Most infrastructure projects involve many areas, and sometimes, it has a different land-use plan that can create some dispute.
	C4: Number of authorities involved	It refers to the complexity that happens due to the administrative and coordination time needed in the project.
	C5: Implementation of procurement between government and business entity	Some projects involve coordination between private parties or other stakeholders that do not have or little experience in the field or can be due to an innovative project.
	C6: Synchronization with other National Strategic Projects	This criterion intends to synchronize between two or more National Strategic Projects that relate to each other.
Policy (P)	P1: Executive direction	The president's vision of the country to distribute the wealth to the society, and have national-range impacts.
	P2: Publishing of supporting policies	The government tries to accelerate development by publishing some sectoral/general policies.
	P3: Issuance of project permits	A project permit is a critical milestone of the project, which lets the team continue/start/operate an activity in the project.

Table 7. The KPPIP main category assessment from the four sectors.

EE			RB			TT			WS		
Category	Best to others	Others to worst	Category	Best to others	Others to worst	Category	Best to others	Others to worst	Category	Best to others	Others to worst
PP^B	{(1; 1)}	{(7; 1)}	PP^B	{(1; 1)}	{(7; 0,7), (Ω; 0,3)}	PP^B	{(1; 1)}	{(2; 0,2), (3; 0,8)}	PP^B	{(1; 1)}	{(5; 0,8), (Ω; 0,2)}
F	{(1; 1)}	{(7; 1)}	F	{(3; 0,7), (Ω; 0,3)}	{(3; 0,8), (Ω; 0,2)}	F	{(1; 0,2), (2; 0,8)}	{(2; 0,8), (3; 0,2)}	F	{(2; 1)}	{(4; 0,7), (Ω; 0,3)}
C^W	{(7; 1)}	{(1; 1)}	C	{(5; 0,8), (Ω; 0,2)}	{(1; 0,6), (Ω; 0,4)}	C	{(1; 0,1), (2; 0,9)}	{(2; 0,9), (3; 0,1)}	C^W	{(5; 0,8), (Ω; 0,2)}	{(1; 1)}
P	{(1; 1)}	{(7; 1)}	P^W	{(7; 0,7), (Ω; 0,3)}	{(1; 1)}	P^W	{(2; 0,2), (3; 0,8)}	{(1; 1)}	P	{(2; 0,3), (3; 0,7)}	{(5; 0,5), (Ω; 0,5)}

B = Best category, W = Worst category.

Table 8. Assessment for all the criteria in the Project Preparation category from four sectors.

EE			RB			TT			WS		
Criteria	Best to others	Others to worst	Criteria	Best to others	Others to worst	Criteria	Best to others	Others to worst	Criteria	Best to others	Others to worst
PP1	{(1; 1)}	{(9; 1)}	PP1	{(2; 0,7), (Ω; 0,3)}	{(5; 0,8), (Ω; 0,2)}	$PP1^B$	{(1; 1)}	{(3; 0,1), (4; 0,9)}	PP1	{(2; 0,7), (Ω; 0,3)}	{(4; 0,8), (Ω; 0,2)}
PP2	{(1; 0,2), (2; 0,8)}	{(9; 1)}	PP2	{(3; 0,8), (Ω; 0,2)}	{(6; 0,8), (Ω; 0,2)}	PP2	{(1; 0,1), (2; 0,9)}	{(3; 0,8), (4; 0,2)}	$PP2^W$	{(6; 0,7), (Ω; 0,3)}	{(1; 1)}
PP3	{(1; 0,2), (2; 0,8)}	{(9; 1)}	$PP3^B$	{(1; 1)}	{(7; 0,7), (Ω; 0,3)}	PP3	{(2; 0,2), (3; 0,8)}	{(2; 0,8), (3; 0,2)}	PP3	{(2; 0,8), (Ω; 0,2)}	{(3; 0,8), (Ω; 0,2)}
$PP4^B$	{(1; 1)}	{(9; 1)}	PP4	{(2; 0,7), (Ω; 0,3)}	{(6; 0,7), (Ω; 0,3)}	PP4	{(3; 0,2), (4; 0,8)}	{(2; 0,9), (3; 0,1)}	$PP4^B$	{(1; 1)}	{(6; 0,7), (Ω; 0,3)}
$PP5^W$	{(9; 1)}	{(1; 1)}	$PP5^W$	{(7; 0,7), (Ω; 0,3)}	{(1; 1)}	$PP5^W$	{(4; 0,1), (5; 0,9)}	{(1; 1)}	PP5	{(4; 0,9), (Ω; 0,1)}	{(5; 0,9), (Ω; 0,1)}

B = Best criterion, W = Worst criterion.

Table 9. Assessment for all the criteria in the Funding category from four sectors.

EE			RB			TT			WS		
Criteria	Best to others	Others to worst	Criteria	Best to others	Others to worst	Criteria	Best to others	Others to worst	Criteria	Best to others	Others to worst
F1	{(9; 1)}	{(1; 1)}	$F1^{W}$	{(6; 0,8), (Ω; 0,2)}	{(1; 1)}	F1	{(4; 0,9), (5; 0,1)}	{(3; 0,1), (4; 0,9)}	F1	{(1; 1)}	{(5; 1)}
$F2^{B}$	{(1; 1)}	{(9; 1)}	$F2^{B}$	{(1; 1)}	{(6; 0,8), (Ω; 0,2)}	$F2^{B}$	{(1; 1)}	{(5; 0,9), (6; 0,1)}	F2	{(1; 1)}	{(5; 0,7), (Ω; 0,3)}
$F3^{W}$	{(9; 1)}	{(1; 1)}	F3	{(2; 0,6), (Ω; 0,4)}	{(3; 0,6), (Ω; 0,4)}	F3	{(3; 0,9), (4; 0,1)}	{(2; 0,1), (3; 0,9)}	$F3^{B}$	{(1; 1)}	{(6; 0,8), (Ω; 0,2)}
F4	{(9; 1)}	{(1; 1)}	F4	{(5; 0,7), (Ω; 0,3)}	{(2; 0,6), (Ω; 0,4)}	F4	{(4; 0,9), (5; 0,1)}	{(2; 0,1), (3; 0,9)}	F4	{(2; 0,7), (Ω; 0,3)}	{(5; 0,8), (Ω; 0,2)}
F5	{(9; 1)}	{(1; 1)}	F5	{(3; 0,6), (Ω; 0,4)}	{(3; 0,6), (Ω; 0,4)}	F5	{(4; 0,9), (5; 0,1)}	{(2; 0,1), (3; 0,9)}	$F5^{W}$	{(6; 0,8), (Ω; 0,2)}	{(1; 1)}
F6	{(9; 1)}	{(1; 1)}	F6	{(5; 0,7), (Ω; 0,3)}	{(5; 0,7), (Ω; 0,3)}	$F6^{W}$	{(5; 0,9), (6; 0,1)}	{(1; 1)}	F6	{(3; 0,8), (Ω; 0,2)}	{(2; 0,8), (Ω; 0,2)}

B = Best criterion, W = Worst criterion.

Table 10. Assessment for all the criteria in the Coordination category from four sectors.

EE			RB			TT			WS		
Criteria	Best to others	Others to worst	Criteria	Best to others	Others to worst	Criteria	Best to others	Others to worst	Criteria	Best to others	Others to worst
$C1^{B}$	{(1; 1)}	{(7; 1)}	$C1^{B}$	{(1; 1)}	{(7; 0,9), (Ω; 0,1)}	$C1^{B}$	{(1; 1)}	{(4; 0,7), (5; 0,3)}	$C1^{B}$	{(1; 1)}	{(6; 0,4), (Ω; 0,6)}
C2	{(1; 1)}	{(7; 1)}	C2	{(7; 0,9), (Ω; 0,1)}	{(3; 0,7), (Ω; 0,3)}	C2	{(2; 0,8), (3; 0,2)}	{(2; 0,2), (3; 0,8)}	C2	{(4; 0,6), (Ω; 0,4)}	{(4; 0,9), (Ω; 0,1)}
C3	{(1; 1)}	{(7; 1)}	C3	{(2; 0,8), (Ω; 0,2)}	{(5; 0,8), (Ω; 0,2)}	C3	{(2; 0,9), (3; 0,1)}	{(2; 0,1), (3; 0,9)}	$C3^{W}$	{(6; 0,4), (Ω; 0,6)}	{(1; 1)}
$C4^{W}$	{(7; 1)}	{(1; 1)}	C4	{(2; 0,8), (Ω; 0,2)}	{(6; 0,8), (Ω; 0,2)}	$C4^{W}$	{(4; 0,7), (5; 0,3)}	{(1; 1)}	C4	{(1; 0,8), (Ω; 0,2)}	{(4; 0,8), (Ω; 0,2)}
C5	{(1; 1)}	{(7; 1)}	C5	{(5; 0,7), (Ω; 0,3)}	{(2; 0,7), (Ω; 0,3)}	C5	{(3; 0,8), (4; 0,1), (Ω; 0,1)}	{(3; 0,1), (4; 0,8), (Ω; 0,1)}	C5	{(4; 0,7), (Ω; 0,3)}	{(5; 0,7), (Ω; 0,3)}
C6	{(1; 1)}	{(7; 1)}	$C6^{W}$	{(7; 0,9), (Ω; 0,1)}	{(1; 1)}	C6	{(2; 0,9), (3; 0,1)}	{(2; 0,1), (3; 0,9)}	C6	{(3; 0,8), (Ω; 0,2)}	{(3; 0,8), (Ω; 0,2)}

B = Best criterion, W = Worst criterion.

Table 11. Assessment for all the criteria in Policy category from four sectors.

EE			RB			TT			WS		
Criteria	Best to others	Others to worst	Criteria	Best to others	Others to worst	Criteria	Best to others	Others to worst	Criteria	Best to others	Others to worst
P1^{W}	{(8; 1)}	{(1; 1)}	P1	{(1; 0,8), (Ω; 0,2)}	{(2; 0,8), (Ω; 0,2)}	P1^{B}	{(1; 1)}	{(4; 0,2), (5; 0,8)}	P1^{W}	{(2; 1)}	{(1; 1)}
P2^{B}	{(1; 1)}	{(8; 1)}	P2^{W}	{(2; 0,8), (Ω; 0,2)}	{(1; 1)}	P2	{(3; 0,2), (4; 0,8)}	{(3; 0,8), (4; 0,2)}	P2	{(1; 0,9), (Ω; 0,1)}	{(2; 0,9), (Ω; 0,1)}
P3	{(1; 1)}	{(8; 1)}	P3^{B}	{(1; 1)}	{(2; 0,8), (Ω; 0,2)}	P3^{W}	{(4; 0,2), (5; 0,8)}	{(1; 1)}	P3^{B}	{(1; 1)}	{(2; 1)}

B = Best criterion, W = Worst criterion.

Table 12. Weights of the main categories.

	Category						
Sector	PP	F	C	P	IR	AU	λ
EE	0.3182	0.3182	0.0455	0.3182	0	0	0.3189
RB	[0.4714,0.6091]	[0.1720,0.3323]	[0.0786, 0.1554]	[0.0857, 0.1107]	0.2103	0.5341	0.1895
TT	[0.4040, 0.4075]	[0.2361, 0.2382]	[0.2251, 0.2318]	[0.1280, 0.1291]	0.0442	0.2011	0.2725
WS	[0.3862, 0.4549]	[0.2116, 0.3306]	[0.0638, 0.0751]	[0.2194, 0.2583]	0.1659	0.4104	0.2191
Aggregated	[0.3855, 0.4276]	[0.2448, 0.3018]	[0.1047, 0.1236]	[0.2007, 0.2143]	—-	—	—

procedure described in Sec. 4 (Step 7), and the respective reliability degrees (weights) are derived via Eq. (24) (Step 8). The last row of Table 12 contains the aggregated weights for each main category (Step 9).

Although the *IRs* and uncertainty degrees are relatively high, we did not ask the experts to revise their preferences in this study, because without a threshold for the belief-based BWM (which could be developed in the future), there is no way to determine whether or not the experts are sufficiently consistent and certain. As such, we accept all the experts' judgment, but with different weights for the experts based on their reliability degrees.

The IRs and AUs in Table 12 are pictured in Fig. 7, which shows how far the experts in the four sectors are from the perfect reliability status: the closer the coordinate is to the origin, the greater the reliability. As we can see, EE is the most reliable, so that his assessments carry a higher weight (0.3189) than the others.

Similarly, we can calculate the local weights for each criterion from Table 8–11 for each sector. By combining the aggregated weights of the main categories with the local weights of the criteria (see Ref. 12 for the interval operations), we can obtain the global weights for each criterion for each sector. Then, we follow Steps 5 to 9 again to obtain the overall weight for each criterion. The results can be seen in Table 13 and Fig. 8.

From the results we can see that the overall weights are rather different from the original weights used by KPPIP. Determination of funding scheme (F2) is considered to be one of the most important criteria with a maximum weight of 0.139. The importance of the criteria in the Coordination category is relatively low.

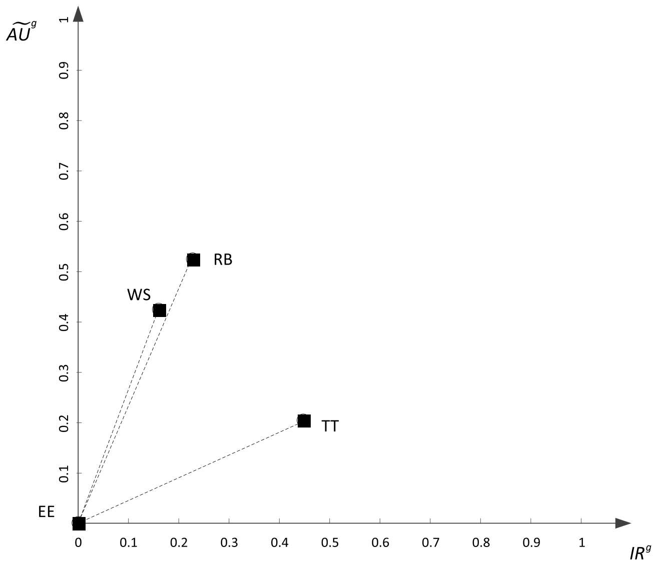

Fig. 7. Reliability of the four sectors.

Table 13. Overall weights of each criterion.

Category	Aggregated weights of categories	Criteria	EE	RB	TT	WS	Overall weights of criteria
PP	[0.3855, 0.4276]	PP1	[0.2336, 0.2614]	[0.1487, 0.2495]	[0.3746, 0.4337]	[0.1134, 0.2857]	[0.0889, 0.1337]
		PP2	[0.2251, 0.2336]	[0.1714, 0.2114]	[0.2047, 0.2850]	[0.0435, 0.0599]	[0.0654, 0.0884]
		PP3	[0.2251, 0.2336]	[0.2336, 0.3020]	[0.3724, 0.1130]	[0.1761, 0.1039]	[0.0707, 0.1064]
		PP4	[0.2614, 0.2712]	[0.1845, 0.2920]	[0.1130, 0.1308]	[0.3487, 0.4800]	[0.0857, 0.1198]
		PP5	[0.0270, 0.0280]	[0.0360, 0.0444]	[0.0813, 0.0941]	[0.1468, 0.2021]	[0.0268, 0.0372]
F	[0.2448, 0.3018]	F1	0.0714	[0.0391, 0.0621]	[0.1319, 0.1738]	[0.1941, 0.2874]	[0.0273, 0.0449]
		F2	0.6429	[0.3110, 0.4937]	[0.3351, 0.4416]	[0.1704, 0.2133]	[0.0957, 0.1390]
		F3	0.0714	[0.0812, 0.2768]	[0.0927, 0.2426]	[0.2367, 0.3396]	[0.0291, 0.0653]
		F4	0.0714	[0.0464, 0.1583]	[0.0844, 0.1627]	[0.1768, 0.2358]	[0.0235, 0.0453]
		F5	0.0714	[0.0675, 0.2724]	[0.0844, 0.1627]	[0.0244, 0.0422]	[0.0153, 0.0369]
		F6	0.0714	[0.1103, 0.1752]	[0.0519, 0.0684]	[0.0432, 0.0537]	[0.0162, 0.0254]
C	[0.1047, 0.1236]	C1	0.1944	[0.3072, 0.4094]	[0.2487, 0.3554]	[0.3193, 0.3999]	[0.0269, 0.0399]
		C2	0.1944	[0.0535, 0.0713]	[0.0967, 0.2405]	[0.1083, 0.1707]	[0.0127, 0.0216]
		C3	0.1944	[0.1497, 0.2785]	[0.1020, 0.2365]	[0.0378, 0.0474]	[0.0137, 0.0242]
		C4	0.0278	[0.1912, 0.3221]	[0.0467, 0.0668]	[0.1439, 0.2346]	[0.0095, 0.0176]
		C5	0.1944	[0.0489, 0.1127]	[0.1337, 0.1911]	[0.1304, 0.1634]	[0.0141, 0.0210]
		C6	0.1944	[0.0372, 0.0496]	[0.1020, 0.2365]	[0.0803, 0.1861]	[0.0119, 0.0212]
P	[0.2007, 0.2143]	P1	0.0588	[0.3611, 0.3626]	[0.6667, 0.6667]	[0.2003, 0.2004]	[0.0606, 0.0647]
		P2	0.4706	[0.1985, 0.2018]	[0.2222, 0.2222]	[0.3814, 0.3815]	[0.0667, 0.0714]
		P3	0.4706	[0.4371, 0.4389]	[0.1111, 0.1111]	[0.4182, 0.4182]	[0.0733, 0.0783]

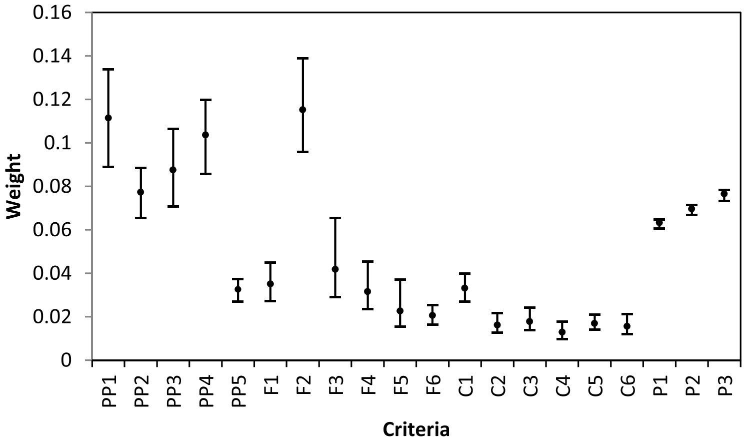

Fig. 8. The overall weights of the KPPIP criteria.

We also checked the weights obtained by the proposed belief-based BWM with the leader of KPPIP, and he confirmed that our findings are much more reasonable than the original ones used by KPPIP.

7. Conclusions

The aim of this study was to develop an extended BWM model to deal with belief structure-related information. Compared to the original BWM, the superiority of the proposed belief-based BWM method has to do with the fact that it can capture different types of uncertainties, including probabilities and vagueness in subjective judgments, and, as discussed in the introduction, that it is more flexible than the fuzzy BWM.

In the belief-based BWM, we first ask the DM to indicate his preferences in pairwise comparisons, with basic belief assignments, which are then transformed into the belief degrees (pignistic probabilities) associated with each grade. These degrees are then used to construct an optimization problem, to obtain the weights of the criteria. Since the nonlinear belief-based BWM was able to generate multiple solutions in cases where DMs are inconsistent, two models are developed to derive the interval weights of criteria.

In real-world contexts, it is likely that a group-based decision-making process is preferred over individual decisions, because of the complexity of the problems. The decision-making processes that take place in group settings tend to make the decisions more comprehensive and reasonable. However, the uncertainty contained in the estimations provided by the different DMs in such a group, and the inconsistency involved in the pairwise comparisons, can produce unreliable and unstable results, making it necessary to measure the uncertainty and inconsistency degree of the preferences being expressed, since these two degrees can reflect the reliability of a DM. To date, few studies have including the reliability of the judgments made by a

DM. To remedy that state of affairs, this study proposes a method to measure the reliability degree of a DM, based on his inconsistency and uncertainty levels. Based on the degree of inconsistency and uncertainty obtained from the preferences of the DMs, we can measure the relative reliability of DMs, which can then be used to assign weights to different DMs, based on which we can aggregate the weights of criteria from the belief BWM, and obtain the final weights of the criteria involved.

It is worth noting that instead of weighing the preferences of the DMs according to how much mutually supportive they are, we propose an approach to weighing the experts based on the quality of their preferences at an individual level. Although there is not a "gold standard" to aggregate preferences, our approach is supported by some empirical and psychological studies, e.g. see Ref. 50, 52, 60, which consider a number of factors contributing to the expertise of a DM. Among these factors, there are the experience, which is reflected in the precision of the judgments, and their internal coherence, i.e. the consistency. In our proposal, both these factors are taken into account.

The ideas underlying the belief-based BWM have been illustrated by numerical examples after each proposed model, and a real-world case study of infrastructure project criteria system assessment in Indonesia is demonstrating the applicability and feasibility of the models.

One of the aims of future research will be to increase our understanding of the inconsistency and uncertainty measures, so that we can determine the thresholds for acceptable levels uncertainty and inconsistency. In addition, it is important to take a closer look at the links between inconsistency and uncertainty, and to examine how they affect one another. In that regard, it might be also interesting to check the relation between the *concentration* of the weights provided by the nonlinear BWM[61] and their inconsistency and uncertainty. Furthermore, it is also possible to extend the belief-based BWM to linear model, but because it does not fit the framework of the group decision-making we leave it to another separate study. And finally, combining other MCDM methods (e.g. TOPSIS, VIKOR, ELECTRE) with belief-based BWM may provide another possible direction of further research.

Acknowledgments

An initial version of this paper was presented at the 25th International Conference on Multiple Criteria Decision Making, June 2019, Istanbul, and we thank the participants for their constructive comments. The authors gratefully acknowledge financial support for the Ph.D. project from the China Scholarship Council (No. 201708440305).

References

1. M. Zeleny, *Multiple Criteria Decision Making* (McGraw Hill, New York, 1982).
2. S. Greco, M. Ehrgott and J. Figueira, Multiple criteria decision analysis: State of the art surveys, in *International Series in Operations Research and Management Science*, 2nd edn. (Springer, New York, 2016).
3. J. Rezaei, Best-worst multi-criteria decision-making method, *Omega* **53** (2015) 49–57.

4. J. Rezaei *et al.*, Quality assessment of airline baggage handling systems using SERVQUAL and BWM. *Tourism Management* **66** (2018) 85–93.
5. H. Badri Ahmadi, S. Kusi-Sarpong and J. Rezaei, Assessing the social sustainability of supply chains using Best Worst Method, *Resources, Conservation and Recycling* **126** (2017) 99–106.
6. H. Gupta and M. K. Barua, Supplier selection among SMEs on the basis of their green innovation ability using BWM and fuzzy TOPSIS, *Journal of Cleaner Production* **152** (2017) 242–258.
7. P. Gupta, S. Anand and H. Gupta, Developing a roadmap to overcome barriers to energy efficiency in buildings using best worst method, *Sustainable Cities and Society* **31** (2017) 244–259.
8. J. Ren, Technology selection for ballast water treatment by multi-stakeholders: A multi-attribute decision analysis approach based on the combined weights and extension theory, *Chemosphere* **191** (2018) 747–760.
9. F. Nawaz *et al.*, An MCDM method for cloud service selection using a Markov chain and the best-worst method, *Knowledge-Based Systems* **159** (2018) 120–131.
10. W. Serrai *et al.*, Towards an efficient and a more accurate web service selection using MCDM methods, *Journal of Computational Science* **22** (2017) 253–267.
11. A. Hafezalkotob *et al.*, Interval MULTIMOORA Method Integrating Interval Borda Rule and Interval Best-Worst-Method-Based Weighting Model: Case Study on Hybrid Vehicle Engine Selection, *IEEE Transactions on Cybernetics* **50**(3) (2019) 1157–1169.
12. J. Rezaei, Best-worst multi-criteria decision-making method: Some properties and a linear model, *Omega* **64** (2016) 126–130.
13. R. Nie *et al.*, Risk evaluation by FMEA of supercritical water gasification system using multi-granular linguistic distribution assessment, *Knowledge-Based Systems* **162** (2018) 185–201.
14. X. Mi *et al.*, The state-of-the-art survey on integrations and applications of the best worst method in decision making: Why, what, what for and what's next? *Omega* **87** (2019) 205–225.
15. G. Klir and M. Wierman, *Uncertainty-Based Information: Elements of Generalized Information Theory*, Vol. 15 (Physica-Verlag, 1999).
16. A. Jousselme *et al.*, Measuring ambiguity in the evidence theory, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* **36**(5) (2006) 890–903.
17. Q. Mou, Z. Xu and H. Liao, An intuitionistic fuzzy multiplicative best-worst method for multi-criteria group decision making, *Information Sciences* **374** (2016) 224–239.
18. S. Guo and H. Zhao, Fuzzy best-worst multi-criteria decision-making method and its applications, *Knowledge-Based Systems* **121** (2017) 23–31.
19. H. Aboutorab *et al.*, ZBWM: The Z-number extension of Best Worst Method and its application for supplier development, *Expert Systems with Applications* **107** (2018) 115–125.
20. D. Pamučar, I. Petrović and G. Ćirović, Modification of the Best-Worst and MABAC methods: A novel approach based on interval-valued fuzzy-rough numbers, *Expert Systems with Applications* **91** (2018) 89–106.
21. G. Shafer, *A Mathematical Theory of Evidence*, Vol. 42 (Princeton University Press, 1976).
22. R. R. Yager and N. Alajlan, Evaluating belief structure satisfaction to uncertain target values, *IEEE Transactions on Cybernetics* **46**(4) (2015) 869–877.
23. I. N. Durbach and T. J. Stewart, Modeling uncertainty in multi-criteria decision analysis, *European Journal of Operational Research* **223**(1) (2012) 1–14.

24. D. Xu, J. Yang and Y. Wang, The evidential reasoning approach for multi-attribute decision analysis under interval uncertainty, *European Journal of Operational Research* **174**(3) (2006) 1914–1943.
25. J. Yang and M. G. Singh, An evidential reasoning approach for multiple-attribute decision making with uncertainty, *IEEE Transactions on Systems, Man, and Cybernetics* **24**(1) (1994) 1–18.
26. M. Beynon, B. Curry and P. Morgan, The Dempster-Shafer theory of evidence: An alternative approach to multicriteria decision modelling, *Omega* **28**(1) (2000) 37–50.
27. J. B. Yang *et al.*, The evidential reasoning approach for MADA under both probabilistic and fuzzy uncertainties, *European Journal of Operational Research* **171**(1) (2006) 309–343.
28. C. Y. Ng and K. B. Chuah, Evaluation of design alternatives' environmental performance using AHP and ER approaches, *IEEE Systems Journal* **8**(4) (2014) 1185–1192.
29. M. Zhou *et al.*, Evidential reasoning rule for MADM with both weights and reliabilities in group decision making, *Knowledge-Based Systems* **143** (2018) 142–161.
30. J. Yang and D. Xu, On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty, *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* **32**(3) (2002) 289–304.
31. M. Guo *et al.*, Evidential reasoning approach for multiattribute decision analysis under both fuzzy and Interval uncertainty, *IEEE Transactions on Fuzzy Systems* **17**(3) (2009) 683–697.
32. L. Fei, J. Lu and Y. Feng, An extended best-worst multi-criteria decision-making method by belief functions and its applications in hospital service evaluation, *Computers & Industrial Engineering* **142** (2020) 106355.
33. G. Li, G. Kou and Y. Peng, A group decision making model for integrating heterogeneous information, *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **48**(6) (2016) 982–992.
34. H. Zhang, G. Kou and Y. Peng, Soft consensus cost models for group decision making and economic interpretations, *European Journal of Operational Research* **277**(3) (2019) 964–980.
35. F. Jia and X. Y. Wang, BWM-TOPSIS multi-criteria group decision-making method based on rough number, *Kongzhi yu Juece/Control and Decision* **31**(10) (2016) 1915–1920.
36. X. You, T. Chen and Q. Yang, Approach to multi-criteria group decision-making problems based on the Best-Worst-Method and ELECTRE method, *Symmetry* **8**(9) (2016) 95.
37. A. Hafezalkotob and A. Hafezalkotob, A novel approach for combination of individual and group decisions based on fuzzy best-worst method, *Applied Soft Computing* **59** (2017) 316–325.
38. Q. Mou, Z. Xu and H. Liao, A graph based group decision making approach with intuitionistic fuzzy preference relations, *Computers & Industrial Engineering* **110** (2017) 138–150.
39. S. Safarzadeh, S. Khansefid and M. Rasti-Barzoki, A group multi-criteria decision-making based on best-worst method, *Computers & Industrial Engineering* **126** (2018) 111–121.
40. P. Hajek and W. Froelich, Integrating TOPSIS with interval-valued intuitionistic fuzzy cognitive maps for effective group decision making, *Information Sciences* **485** (2019) 394–412.
41. M. Mohammadi and J. Rezaei, Bayesian best-worst method: A probabilistic group decision making model, *Omega* **96** (2020) 102075.
42. H. A. Simon, A behavioral model of rational choice, *The Quarterly Journal of Economics* **69**(1) (1955) 99–118.

43. H. A. Simon, Rational choice and the structure of the environment, *Psychological Review* **63**(2) (1956) 129.
44. C. Fu, J. Yang and S. Yang, A group evidential reasoning approach based on expert reliability, *European Journal of Operational Research* **246**(3) (2015) 886–893.
45. A. Wang *et al.*, Quantitative evaluation of human-reliability based on fuzzy-clonal selection, *IEEE Transactions on Reliability* **60**(3) (2011) 517–527.
46. R. R. Yager, On the dempster-shafer framework and new combination rules, *Information Sciences* **41**(2) (1987) 93–137.
47. P. Smets and R. Kennes, The transferable belief model, *Artificial Intelligence* **66**(2) (1994) 191–234.
48. D. Harmanec and G. J. Klir, Measuring total uncertainty in Dempster-Shafer theory: A novel approach, *International Journal of General System* **22**(4) (1994) 405–419.
49. J. Abellán and A. Masegosa, Requirements for total uncertainty measures in Dempster–Shafer theory of evidence, *International Journal of General Systems* **37**(6) (2008) 733–747.
50. J. Shanteau *et al.*, *How Can You Tell If Someone Is an Expert? Performance-Based Assessment of Expertise*, J. Shanteau and S. L. Schneider, J. Shanteau and S.L. Schneider (eds.), Cambridge University Press: Cambridge, 2003, pp. 620–640.
51. Y. W. Du *et al.*, A reliability-based consensus model for multiattribute group decision-making with analytically evidential reasoning approach, *Mathematical Problems in Engineering* **2018** (2018) 1–14.
52. D. J. Weiss and J. Shanteau, The vice of consensus and the virtue of consistency, Cambridge University Press: New York, NY, US (2004), pp. 226–240.
53. N. R. Pal, J. C. Bezdek and R. Hemasinha, Uncertainty measures for evidential reasoning I: A review, *International Journal of Approximate Reasoning* **7**(3) (1992) 165–183.
54. G. Kou, D. Ergu and J. Shang, Enhancing data consistency in decision matrix: Adapting Hadamard model to mitigate judgment contradiction, *European Journal of Operational Research* **236**(1) (2014) 261–271.
55. J. Yang, Rule and utility based evidential reasoning approach for multiattribute decision analysis under uncertainties, *European Journal of Operational Research* **131**(1) (2001) 31–61.
56. D. Harmanec *et al.*, On the computation of uncertainty measure in Dempster-Shafer theory, *International Journal of General System* **25**(2) (1996) 153–163.
57. C. Liu *et al.*, Reducing algorithm complexity for computing an aggregate uncertainty measure, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* **37**(5) (2007) 669–679.
58. V. Huynh and Y. Nakamori, Notes on "Reducing algorithm complexity for computing an aggregate uncertainty measure", *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* **40**(1) (2010) 205–209.
59. KPPIP, *KPPIP Report: Period of January–June 2017* (KPPIP, 2017).
60. J. Shanteau *et al.*, Performance-based assessment of expertise: How to decide if someone is an expert or not, *European Journal of Operational Research* **136**(2) (2002) 253–263.
61. J. Rezaei, A concentration ratio for nonlinear best worst method, *International Journal of Information Technology & Decision Making* **19**(3) (2020) 891–907.