Bayesian best-worst method
A probabilistic group decision making model

Mohammadi, Majid; Rezaei, Jafar

DOI
10.1016/j.omega.2019.06.001

Publication date
2019

Document Version
Final published version

Published in
Omega (United Kingdom)

Omega xxx (xxxx) xxx

Contents lists available at ScienceDirect

Omega

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

Bayesian best-worst method: A probabilistic group decision making model☆

Majid Mohammadi*, Jafar Rezaei

Faculty of Technology, Policy and Management, Delft University of Technology, The Netherlands

ARTICLE INFO

Article history:
Received 10 December 2018
Accepted 3 June 2019
Available online xxx

Keywords:
Best-worst method
Multi-criteria decision-making
Bayesian hierarchical model
Generalizability
Group decision-making

ABSTRACT

The best-worst method (BWM) is a multi-criteria decision-making method which finds the optimal weights of a set of criteria based on the preferences of only one decision-maker (DM) (or evaluator). However, it cannot amalgamate the preferences of multiple decision-makers/evaluators in the so-called group decision-making problem. A typical way of aggregating the preferences of multiple DMs is to use the average operator, e.g., arithmetic or geometric mean. However, averages are sensitive to outliers and provide restricted information regarding the overall preferences of all DMs. In this paper, a Bayesian BWM is introduced to find the aggregated final weights of criteria for a group of DMs at once. To this end, the BWM framework is meaningfully viewed from a probabilistic angle, and a Bayesian hierarchical model is tailored to compute the weights in the presence of a group of DMs. We further introduce a new ranking scheme for decision criteria, called *credal ranking*, where a confidence level is assigned to measure the extent to which a group of DMs prefers one criterion over one another. A weighted directed graph visualizes the credal ranking based on which the interrelation of criteria and confidences are merely understood. The numerical example validates the results obtained by the Bayesian BWM while it yields much more information in comparison to that of the original BWM.

1. Introduction

Multi-criteria decision-making (MCDM) is a sub-discipline of Operations Research, which has growingly gained momentum since its genesis. In a typical MCDM problem, a number of alternatives are evaluated based on a handful number of criteria. The evaluation is usually performed based on the elicitation of preferences of a decision maker (DM) and commonly results in sorting, ranking, or selecting the alternative(s). In order to do the evaluation, we need to find the performance of the alternatives with respect to the criteria, which is called the performance matrix, and the importance (weight) of the criteria. Finding the performance matrix usually follows a simple yet crucial data collection approach. Weight determination is usually done based on the preferences of the actual DM. There exist several preference elicitation methods to infer the weights of the decision criteria based on the preferences of the DM, including the analytic hierarchy process (AHP) [1], the analytic network process (ANP) [2], the simple multi-attribute rating technique (SMART) [3,4], Swing [5], FARE [6], CILOS and IDOCRIW [7], to name just a few (see [8] for more MCDM methods). One of the most recently developed preference elicitation methods is the best-worst method (BWM) developed by Rezaei in 2015 [9,10], which is a pairwise comparison-based MCDM method.

When we have a single DM, the elicited preferences are directly used in the decision analysis while incorporating the elicited preferences is not a straightforward step when there are several DMs. The latter case is usually called group MCDM [11–13]. We can classify group MCDM problems into two categories. In the first category, which has a normative approach, a group of DMs seeks a solution which somehow represents the opinion of the whole group. In the second category, which is of a descriptive approach, we want to have a clear understanding of the preferences of the DMs. An example of the first category is when a number of DMs from a supply chain management department of a company decides on selecting the best suppliers for some materials used in the company [14], while an example from the second category is when a researcher tries to understand the importance of the criteria which define the quality of passenger transport transit nodes [15]. The main focus of this study is on group MCDM, where we have the preferences of a group of DMs, whether it is used for a normative or a descriptive approach.

☆ This manuscript was processed by Associate Editor Triantaphyllou.

* Corresponding author.
E-mail addresses: m.mohammadi@tudelft.nl (M. Mohammadi), J.rezaei@tudelft.nl (J. Rezaei).

https://doi.org/10.1016/j.omega.2019.06.001
0305-0483/© 2019 Published by Elsevier Ltd.

For the weight elicitation methods that are based on the pairwise comparison (PC), there are two classes of techniques which can be used to reconcile the discrepancy among DMs [16,17]. The first approach is the aggregation of individual judgment (AIJ) [18,19], in which the PCs of different DMs are first integrated into one, and the resulting aggregated PC is then treated as a single DM problem and evaluation is performed accordingly. The other class is the aggregation of individual priorities (AIP) [20–24]. In the AIP, a weight vector is first calculated for each DM, and the consequent weights are combined to result in a single weight vector. The most popular technique to find the optimal weight for the AIP is the arithmetic mean [25] (for other techniques of aggregation, see, for instance [26]). Both AIJ and AIP approaches result in a weight vector which represents the preferences of the whole group. Although both are practically simple, we lose much information due to the aggregation. That is to say, we use the centrality feature and ignore the dispersion property. On top of that, averages are sensitive to outliers. Therefore, even if one decision-maker has different preferences from the entire group, he/she will significantly influence the overall aggregated preferences of all DMs.

In this study, we propose a novel approach for group MCDM. The proposed approach is particularly presented for the BWM due to its particular features. The pairwise comparison vectors associated with each DM in the BWM contain integers only; hence, they can be modeled using the multinomial distribution. Nevertheless, the proposed approach can be extended for other MCDM methods with some efforts. More specifically, the Bayesian BWM is introduced which can solve the group MCDM problem. The inputs to the Bayesian BWM are identical to those of the original BWM, which are the pairwise comparisons. The output is, on the other hand, the optimal aggregated final weights reflecting the total preferences of all DMs along with the confidence level for ranking the criteria.

Since the Bayesian BWM is stochastic, the inputs and outputs of the method need to be modeled using probability distributions. In particular, we model the pairwise comparisons using the multinomial distribution, and the final aggregated weights by the Dirichlet distribution. We further demonstrate that such modeling, though different, is identical to what is expected in the MCDM, and the BWM in particular.

Based on the inputs and required outputs, a Bayesian hierarchical model is developed to find the optimal weights of all DMs and the aggregated final weight at once. The proposed model is distinct from that of the AIJ in which various PCMs are combined to reach a consensus matrix. In the AIJ, one needs to accept that some DMs compromise in order to get a unanimous ranking. However, we merely view various DMs as statistical samples based on which the criteria are probabilistically evaluated. The *credal ranking* is further introduced in which each pair of criteria has a relation, e.g., $<$ or $>$, with a confidence level. The confidence level represents the extent to which one can be certain about the superiority of a criterion over one another. The confidence level is computed based on the Bayesian test that is especially-tailored based on the proposed hierarchical model. A weighted directed graph visualizes the outcome of the credal ranking.

The main contribution of this study is to propose a novel approach in group MCDM and to apply Bayesian statistics to MCDM. This approach is used for the BWM, which is a significant empowerment for the method for its use in the context of group decision-making. The proposed Bayesian BWM is particularly very powerful when the goal is to describe the preferences of a group of DMs (who can be the actual DMs, experts, or users).

The remainder of this article is structured as follows. Section 2 contains the original best-worst method and the corresponding optimization problem to obtain the optimal weights of the criteria for one DM only. In Section 3, we provide the probabilistic interpretation of the BWM inputs and outputs and justify that such an interpretation would preserve the underlying ideas in the original BWM. Section 4 is dedicated to the proposed Bayesian model, and we present the credal ranking in Section 5. The numerical example regarding the proposed model is given in Section 6, and the article is concluded in Section 7.

2. Best-worst method

The BWM is a relatively new MCDM method [9,10]. One of the most popular pairwise comparison-based MCDM methods is the AHP [1] which needs to have the pairwise comparison of all the n decision criteria together, i.e., $n(n-1)/2$ pairwise comparisons. In contrast, the BWM needs only the so-called reference pairwise comparisons, i.e., $2n-3$ pairwise comparisons. Other than this feature of the BWM, which makes it a more data efficient method compared to AHP, it has several other interesting features. By first selecting the best and the worst criteria and then comparing all the other criteria with these two criteria, it gives a structure to the problem. Such structure helps the DM to provide more reliable pairwise comparisons [9]. Furthermore, the particular structure of the BWM leads to two vectors containing only integers, which prevents a fundamental distance problem associated with the use of fractions in pairwise comparisons [27]. The original BWM is presented as a non-linear optimization problem [9], while there exists a linear approximation [10], a multiplicative version [28], and some hybrid versions such as BWM-MULTIMOORA [29] and BWM-VIKOR. The method has also been extensively used in many real-world applications including, but not limited to, transport and logistics [30–32], supply chain management [33–39], technology management [40], risk management [41], science and research assessment [42,43], and energy [44,45] (see [46] for more recent advances in the BWM and its applications).

Since the two vectors provided by each DM in the BWM might represent different comparisons (with different bests and worsts), the AIJ is not a proper way of aggregating the preferences of a group of DMs for this method. Almost all applications presented in existing literature use the AIP for the aggregation, i.e., the arithmetic mean of the weights of the criteria obtained from the individual DMs. There exists a number of researchers who have proposed different ways for the case of group decision-making with the BWM [47,48]. However, none of them has proposed a way to find the overall weights of the group in a probabilistic environment.

The steps required for the original BWM are as follows [9].

Step 1: The DM needs to provide a set of decision criteria $C = \{c_1, c_2, \ldots, c_n\}$.

Step 2: The DM selects the best (c_B) and the worst (c_W) criteria from C.
In this step, the DM only selects the best and the worst from the criteria set C identified in the first step. The DM does not conduct any pairwise comparison in this stage. Based on the DM's preference, the best criterion is the most important or the most desirable while the worst criterion is the least important or the least desirable criterion among others.

Step 3: The DM conducts the pairwise comparison between the best (c_B) and the other criteria from C.
In this step, the DM calibrates his/her preferences of the best criterion to the other criteria by a number between one and nine, where *one* means equally important and *nine* means extremely more important. The pairwise comparison leads to the "Best-to-Others"

Please cite this article as: M. Mohammadi and J. Rezaei, Bayesian best-worst method: A probabilistic group decision making model, Omega, https://doi.org/10.1016/j.omega.2019.06.001

vector A_B as

$$A_B = \left(a_{B1}, a_{B2}, \ldots, a_{Bn}\right) \quad (1)$$

where a_{Bj} represents the preference of the best (c_B) to the criterion $c_j \in C$.

Step 4: The DM conducts the pairwise comparison between the worst (c_W) and the other criteria from C.
Similar to Step 3, the DM needs to calibrate his/her preferences of the other criteria over the worst criterion by a number between one and nine. The result of this step is the "Others-to-Worst" vector A_W as

$$A_W = \left(a_{1W}, a_{2W}, \ldots, a_{nW}\right)^T \quad (2)$$

where a_{jW} represents the preference of the criterion $c_j \in C$ over the worst (c_W).

Step 5: Obtaining the optimal weights $w^* = (w_1^*, w_2^*, \ldots, w_n^*)$.
Given A_B and A_W, a weight vector w^* must be computed. The weight vector must be in the neighborhood of the equations $w_B/w_j = a_{Bj}$ and $w_j/w_W = a_{jW}$ for $j = 1, 2, ., n$. Thus, one can minimize the maximum absolute differences $|\frac{w_B}{w_j} - a_{Bj}|$ and $|\frac{w_j}{w_W} - a_{jW}|$ for all $j = 1, 2, \ldots, n$. Besides, the non-negativity and sum-to-one property of the weight vector must be fulfilled. As a result, the following optimization problem can find the optimal weight vector w^* [9]

$$\min_{w} \max_{j} \left\{ \left| \frac{w_B}{w_j} - a_{Bj} \right|, \left| \frac{w_j}{w_W} - a_{jW} \right| \right\}$$
$$s.t. \quad \sum_{j=1}^{n} w_j = 1, \quad w_j \geq 0 \quad \forall j = 1, 2, \ldots, n. \quad (3)$$

Similarly, the weight vector can also be calculated by the following problem [10]

$$\min_{\xi, w} \quad \xi$$
$$s.t. \quad \left| \frac{w_B}{w_j} - a_{Bj} \right| \leq \xi \quad \forall j = 1, 2, \ldots, n$$
$$\left| \frac{w_j}{w_W} - a_{jW} \right| \leq \xi \quad \forall j = 1, 2, \ldots, n$$
$$\sum_{j=1}^{n} w_j = 1, \quad w_j \geq 0 \quad \forall j = 1, 2, \ldots, n. \quad (4)$$

To check the reliability of the optimal weights, the veracity between the input pairwise comparisons and their associated weight ratios are checked using the following consistency ratio (CR):

$$CR = \frac{\xi^*}{CI} \quad (5)$$

where ξ^* is the optimal objective value of model (4), and CI (consistency index) is a fixed value per a_{BW}, which can be read from Table 1.
CR is a number between 0 and 1, where 0 means full consistency and by increasing the value of CR the consistency of the pairwise comparison system is decreasing.

3. Probabilistic interpretation of BWM

In this section, we provide a probabilistic interpretation of the BWM inputs and outputs, and then review two schools of thoughts in the probability estimation, e.g., frequentist and Bayesian, in the context of the BWM.

3.1. Modeling inputs and outputs: multinomial and Dirichlet distributions

As stated before, the typical outcome of the MCDM methods is the weight vector $w = [w_1, \ldots, w_n]$ such that $w_j \geq 0, \sum_{j=1}^{n} w_j = 1$. The magnitude of each w_j indicates the importance of the corresponding criteria c_j.

From a probabilistic perspective, the criteria are seen as the random events, and their weights are thus their occurrence likelihoods. Mathematically speaking, such an interpretation is in line with the MCDM since $w_j \geq 0$ and $\sum_{j=1}^{n} w_j = 1$ according to the probability theory as well. It is further of the essence to illustrate that probabilistic modeling makes sense from a decision-making point of view.

For the probabilistic reasoning, one needs to model all the inputs and the outputs as the probability distributions. First, consider A_B and A_W which are the inputs to the BWM. From a mathematical point of view, the multinomial distribution can model the vectors since all of their elements are integers. The probability mass function (PMF) of the multinomial distribution for a given A_w is [49]

$$P(A_W|w) = \frac{\left(\sum_{j=1}^{n} a_{jW}\right)!}{\prod_{j=1}^{n} a_{jW}!} \prod_{j=1}^{n} w_j^{a_{jW}} \quad (6)$$

where w is the probability distribution.

In the multinomial distribution, the weight vector is the probability distribution and A_W contains the number of occurrence of each event. Apparently, it is completely different from what is expected for the BWM represented in Section 2. Interestingly, we show that modeling with multinomial would fulfill the underlying idea of the BWM.

Based on the multinomial distribution, the probability of the event j is proportionate to the number of occurrence of the event to the total number of trials, i.e.,

$$w_j \propto \frac{a_{jW}}{\sum_{i=1}^{n} a_{iW}}, \quad \forall j = 1, \ldots, n. \quad (7)$$

Similarly, one can write the same equation for the worst criterion as

$$w_W \propto \frac{a_{WW}}{\sum_{i=1}^{n} a_{iW}} = \frac{1}{\sum_{i=1}^{n} a_{iW}} \quad (8)$$

Using Eqs. (7) and (8), one obtains

$$\frac{w_j}{w_W} \propto a_{jW}, \quad \forall j = 1, \ldots, n, \quad (9)$$

which is precisely the relation we seek in the original BWM presented in Step 5 of Section 2.

Similarly, A_B can be modeled using the multinomial distribution. However, A_B is different from A_W since the former represents the preferences of the best over the other criteria while the latter denotes the preferences of the others over the worst. Thus, A_B

Table 1
Consistency Index (CI) Table [9].

a_{BW}	1	2	3	4	5	6	7	8	9
Consistency Index (CI)	0.00	0.44	1.00	1.63	2.30	3.00	3.73	4.47	5.23

yields the inverse of the weight, i.e.,

$$A_B \sim multinomial(1/w) \tag{10}$$

where w is the probability distribution, and / represents the element-wise division operator. Identical to the worst criterion, one can write

$$\frac{1}{w_j} \propto \frac{a_{Bj}}{\sum_{i=1}^n a_{Bi}}, \quad \frac{1}{w_B} \propto \frac{a_{BB}}{\sum_{i=1}^n a_{Bi}} = \frac{1}{\sum_{i=1}^n a_{Bi}} \\ \Rightarrow \frac{w_B}{w_j} \propto a_{Bj}, \quad \forall j = 1, \ldots, n, \tag{11}$$

which is again the exact relation we seek in the BWM.

So far, we showed that the multinomial distribution could meaningfully model the inputs to the BWM. The problem of finding the weights in the MCDM problem is thus transferred to the estimation of a probability distribution. Therefore, one can use the statistical inference techniques to find w in the multinomial distribution.

A weight vector for the MCDM must satisfy the non-negativity and sum-to-one properties. Therefore, an appropriate distribution to model the weights is to use the Dirichlet distribution. Given a parameter $\alpha \in R^n$, the Dirichlet distribution of the weights w is defined as [49]

$$Dir(w|\alpha) = \frac{1}{B(\alpha)} \prod_{j=1}^{n} w_j^{\alpha_j - 1}. \tag{12}$$

The distribution has only a vector parameter α, and w meets the constraints of an optimal weight vector of MCDM since it is a probability distribution.

3.2. Estimation of the weight vector: statistical inference

For a moment, assume that there is only A_W in the BWM, then we consider two widely-accepted inference techniques: frequentist and Bayesian. The underlying idea of the frequentist approach is that there is a precise yet unknown optimal point, and the effort is to estimate it based on the observations. As a result, the outcome of the frequentist inference is a precise weight vector for a set of criteria. The maximum likelihood estimation (MLE) is arguably the most popular inference technique which finds the optimal weight vector using the following optimization

$$w = arg \max_{w, \sum_{j=1}^n w_j = 1} P(A_W|w). \tag{13}$$

The optimum of (13) yields at

$$w_j^* = \frac{a_{jW}}{\sum_{i=1}^n a_{iW}}, \quad \forall j = 1, \ldots, n, \tag{14}$$

which is indeed the normalized A_W. The same solution will be obtained by the BWM if the preferences of the DM are fully consistent. Thus, (14) shows that the MLE bears the same result as the BWM under specific circumstances.

The second approach is the Bayesian estimation, in which the parameters are approximated by using a distribution rather than a precise point as is in the MLE. Thus, we first need to specify a prior distribution for the weight vector. In the Bayesian inference, the Dirichlet distribution is used as the prior to the multinomial. The Dirichlet distribution can perfectly represent the weight vector since it satisfies both its non-negativity and sum-to-one properties. Using Dirichlet as the prior and multinomial as the likelihood, the posterior distribution would also be Dirichlet with the posterior parameter $\alpha_{post} = \alpha + A_W$. Since the prior should be uninformative so that its impact on the posterior is minimal, we set the prior parameter $\alpha = 1$.

As a result of the Bayesian estimation, the values of w is shown by a Dirichlet distribution. The mode of the posterior distribution $\mu \in R^n$ with the parameter α_{post}

$$\mu_j = \frac{\alpha_{post_j} - 1}{\sum_{i=1}^n \alpha_{post_i} - n} \\ = \frac{1 + a_{jW} - 1}{\sum_{i=1}^n (a_{iW} + 1) - n} \\ = \frac{a_{jW}}{\sum_{i=1}^n a_{iW}}, \quad \forall j = 1, \ldots, n. \tag{15}$$

Thus, the mode of the posterior distribution would provide the exact MLE. As a result, the Bayesian paradigm would yield more information regarding the events under study since its outcome is a distribution, not a point. The standard deviation of such a distribution, for instance, is an indicator of uncertainty regarding the inference problem, which can have distinct interpretations with respect to the problem under study.

So far, we merely considered A_W for estimating the weights; however, it is critical to use both A_B and A_W according to the BWM. The MLE inference containing both A_B and A_W does not bear an analytical solution due to the complexity of the corresponding optimization problem. Further, the simple Dirichlet-multinomial conjugate cannot encompass the A_B and A_W together. The problem compounds when it comes to having the preferences of multiple DMs. Considering these issues, a Bayesian hierarchical model is presented in the next section to estimate the optimal weight of the criteria considering both A_B and A_W of multiple DMs.

4. Bayesian best-worst method

This section presents a Bayesian hierarchical model to find the optimal weights of a set of criteria based on the preferences of multiple DMs using the best-worst framework.

4.1. Group decision-making: a joint probability distribution

Assume that the k^{th} DM, $k = 1, \ldots, K$, evaluates the criteria $c_1, \ldots, c_n$ by providing the vectors A_B^k and A_W^k. We show the set of all vectors of K DMs by $A_B^{1:K}$ and $A_W^{1:K}$. From now on, the superscript $^{1:K}$ would indicate the total of all vectors in the base. We also represent the overall optimal weight by w^{agg}.

The estimation of w^{agg} entails using several auxiliary variables. In particular, w^{agg} is computed based on the optimal weights of K DMs shown by w^k, $k = 1, \ldots, K$. Thus, the proposed Bayesian model would simultaneously compute w^{agg} and $w^{1:K}$. Prior to conducting any statistical inference, it is required to write the joint probability distribution of all random variables given the available data. In the group decision-making within the BWM, $A_B^{1:K}$ and $A_W^{1:K}$ are given, and $w^{1:K}$ and w^{agg} must be estimated accordingly. Thus, the following joint probability distribution is sought

$$P\left(w^{agg}, w^{1:K} \,\middle|\, A_B^{1:K}, A_W^{1:K}\right). \tag{16}$$

If the probability in (16) is computed, then the probability of each individual variable can be computed using the following probability rule

$$P(x) = \sum_y P(x, y) \tag{17}$$

where x and y are two arbitrary random variables.

4.2. Bayesian hierarchical model

To develop a Bayesian model, we first need to identify the independence and conditional independence of variables. Fig. 1 plots

Please cite this article as: M. Mohammadi and J. Rezaei, Bayesian best-worst method: A probabilistic group decision making model, Omega, https://doi.org/10.1016/j.omega.2019.06.001

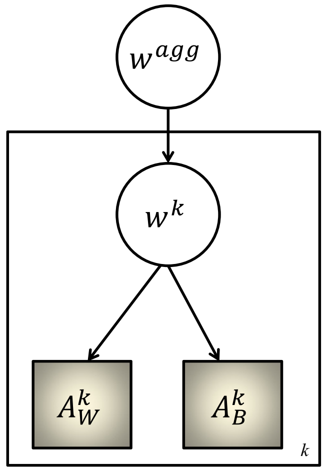

Fig. 1. The probabilistic graphical model of the Bayesian BWM.

the graphical model corresponding to the proposed method. The nodes in the graph are the variables. As a convention, the rectangles are the observed variables, which are the inputs to the original BWM. The circular nodes are the variables which must be estimated. Also, arrows denote that the node in origin is dependent on the node at the other end. That is to say, the value of w_k is dependent on A_W^k and A_B^k, and the value of w^{agg} is also dependent on w_k.

The plate, which covers a set of variables, means that the corresponding variables are iterated for each DM, and w^{agg} is not in the plate since there is only one w^{agg} for all DMs.

The conditional independence between various variables is clear based on Fig. 1. For instance, A_W^k is independent of w^{agg} given w_k, i.e.,

$$P\left(A_W^k \,\middle|\, w^{agg}, w^k\right) = P\left(A_W^k \,\middle|\, w^k\right) \tag{18}$$

Considering all independence among different variables, applying the Bayes rule to the joint probability (16) follows

$$P\left(w^{agg}, w^{1:K} \,\middle|\, A_B^{1:K}, A_W^{1:K}\right) \propto P\left(A_B^{1:K}, A_W^{1:K} \,\middle|\, w^{agg}, w^{1:K}\right) P\left(w^{agg}, w^{1:K}\right)$$
$$= P(w^{agg}) \prod_{k=1}^{K} P\left(A_W^k \,\middle|\, w^k\right) P\left(A_B^k \,\middle|\, w^k\right) P\left(w^k \,\middle|\, w^{agg}\right). \tag{19}$$

where the last equality is obtained using the probability chain rule and the conditional independence of different variables, and the fact that each DM provides his/her preferences independently. Since the estimation of the parameters in Eq. (19) is reliant on the estimation of other variables, there is a chain between different parameters. The existence of the chain is the reason that the model is called hierarchical.

We now need to specify the distributions of each and every element in Eq. (19). We have already shown that A_B and A_W can be perfectly modeled using the multinomial distribution in the sense that it preserves the underlying idea of the BWM. There is only one difference between A_B and A_W since the former shows the preference of all the criteria over the worst, while the latter contains the preference of the best over all the other criteria. Thus, one can model them as

$$A_B^k | w^k \sim multinomial\left(1/w^k\right), \quad \forall k = 1, \ldots, K,$$
$$A_W^k | w^k \sim multinomial\left(w^k\right), \quad \forall k = 1, \ldots, K. \tag{20}$$

Given w^{agg}, one can expect that each and every w^k be in its proximity. To this end, we reparameterize the Dirichlet distribution with respect to its mean and a concentration parameter. The models of w^k given w^{agg} are

$$w^k | w^{agg} \sim Dir(\gamma \times w^{agg}), \quad \forall k = 1, \ldots, K, \tag{21}$$

where w^{agg} is the mean of the distribution and γ is the concentration parameter.

The equation in (21) says that the weight vector w^k associated with each DM must be in the proximity of w^{agg} since it is the mean of the distribution, and their closeness is governed by the non-negative parameter γ. Such a technique is used in different Bayesian models as well [50]. The concentration parameter also needs to be modeled using a distribution. A reliable option is the gamma distribution which satisfies the non-negativity constraints, i.e.,

$$\gamma \sim gamma(a, b), \tag{22}$$

where a and b are the shape parameters of the gamma distribution.

We finally supply the prior distribution over w^{agg} using an uninformative Dirichlet distribution with the parameter $\alpha = 1$ as

$$w^{agg} \sim Dir(\alpha). \tag{23}$$

The specified model does not bear a closed-form solution. As a result, Markov-chain Monte Carlo (MCMC) techniques [51] must be used to compute the posterior distribution. For the MCMC sampling, we use the "just another Gibbs sampler" (JAGS) [52], which is one of the best available probabilistic languages to date, to sample and compute the posterior determined in (19). The useful outcome of the model is the posterior distribution of weights for every single DM and the aggregated w^{agg}.

The proposed Bayesian model will replace Step 5 of the original BWM explained in Section 2. In fact, the optimization problem is substituted with a probabilistic model while the inputs to both methods are identical. However, the proposed model would provide more information regarding the confidence of the relation between each pair of criteria. The excessive information is obtained by devising a new Bayesian test based on the approximated distribution from the model, which is explained in the next section.

5. Credal ranking

The modus operandi in the MCDM is to say one criterion is more important than one another merely if its weight, or the weight average for the group decision-making, is higher than one another. Assume that there are three criteria c_1, c_2, and c_3 with the weight vector $w = [0.49, 0.50, 0.01]$. According to the MCDM, c_2 is superior to both c_1 and c_3. However, the confidence of the superiority cannot be determined by solely comparing two figures. This is even much more important when the weight vector represents the preferences of a group of DMs. To date, there are various ranking schemes such as interval-based ranking [53], fuzzy ranking [54,55], and ranking based on grey relational analysis (GRA) [56].

The notion of credal ranking is now introduced, which can calibrate the degree to which one criterion is superior to one another. Having the posterior distribution of weights would help gauge the confidence of the relations between various criteria. The difference between the credal ranking and other ranking schemes is that a confidence is computed in the credal ranking based on one distribution, i.e., the Dirichlet distribution of w^{agg}, while other ranking methods usually take two numbers/intervals and try to find the extent to which one is superior.

We first define the credal ordering, which is the building-block of credal ranking.

Definition 5.1 (Credal Ordering)**.** For a pair of criteria c_i and c_j, the credal ordering O is defined as

$$O = (c_i, c_j, R, d) \tag{24}$$

Table 2
Comparison of the original BWM and the Bayesian BWM on the mobile phone selection criteria based on the preferences of 50 students.

	Basic	Physical char.	Tech feat.	Func	Brand	Customer
BWM	0.1945	0.1623	0.2014	0.2467	0.1277	0.0673
Bayesian BWM	0.1929	0.1776	0.2052	0.2376	0.1277	0.0591

where

- R is the relation between the criteria c_i and c_j, i.e., $<$, $>$, or $=$;
- $d \in [0, 1]$ represents the confidences of the relation.

Definition 5.2 (Credal Ranking)**.** For a set of criteria $C = (c_1, c_2, \dots, c_n)$, the credal ranking is a set of credal orderings which includes all pairs (c_i, c_j), for all $c_i, c_j \in C$.

The confidence in the credal ordering can provide the DMs with more information which can significantly improve their decisions. We now devise a new Bayesian test based on which we can find the confidence of each credal ordering. The test is predicated on the posterior distribution of w^{agg}. The confidence that c_i being superior to c_j is computed as

$$P(c_i > c_j) = \int I_{(w_i^{agg} > w_j^{agg})} P(w^{agg}). \quad (25)$$

where $P(w^{agg})$ is the posterior distribution of w^{agg} and I is one if the condition in the subscript holds, and zero otherwise. This integration can be approximated by the samples obtained via the MCMC. Having Q samples from the posterior distribution, the confidence can be computed as

$$P(c_i > c_j) = \frac{1}{Q} \sum_{q=1}^{Q} I(w_i^{agg_q} > w_j^{agg_q})$$
$$P(c_j > c_i) = \frac{1}{Q} \sum_{q=1}^{Q} I(w_j^{agg_q} > w_i^{agg_q}) \quad (26)$$

where w^{agg_q} is the q^{th} sample of w^{agg} from the MCMC samples. Thus, for each pair of criteria, one can compute the confidence that one is superior to one another. The credal ranking can be merely changed into the traditional ranking. In this regard, it is evident that $P(c_i > c_j) + P(c_j > c_i) = 1$. Therefore, c_i is more important than c_j if and only if $P(c_i > c_j) > 0.5$. As a result, the traditional ranking of criteria is obtainable by applying a threshold of 0.5 to the credal ranking.

6. Numerical examples

In this section, a real-world example is analyzed using the Bayesian BWM, and the corresponding credal ranking is computed and visualized by using a weighted directed graph. The MATLAB implementation of the proposed model can be found at http://bestworstmethod.com/software/.

The real-world application is the selection of the mobile phone, to which the BWM has been already applied [9]. The problem implicates the selection of one from a set of mobile phones based on different criteria. Six different criteria that studied and found in the literature is used to evaluate the mobile phone alternatives. The criteria are basic requirement, physical characteristics, technical features, functionality, brand choice, and customer excitement.

The data collected in [9] was from 50 university students who completely got familiar with different selected criteria through a provided document. Various characteristics (e.g., price, dimension, weight, display) of four particular mobile phones were given to the participants. Through a questionnaire, students filled in a form to get the information required for the original BWM, i.e., A_B and A_W.

The first approach, which was employed in [9], was to find the optimal weight vector separately for each student, and then aggregate them using the arithmetic mean to compute the final aggregated weight vector. The first row of Table 2 tabulates the final aggregated weights obtained by the BWM. We particularly consider the arithmetic mean to validate the proposed Bayesian model since the average of 50 participants is a reliable measure according to *central limit theorem*.

The obtained inputs from 50 participants in this experiment were also given to the Bayesian BWM, and the outcome of the test is obtained. Since the output of the aggregated weight is a distribution in the Bayesian BWM, it is not possible to compare the two methods directly and verify if the output of the Bayesian BWM is valid. To have a meaningful discussion and validation, however, we use the average of the Dirichlet distribution of w^{agg} to be able to compare the two methods. The second row of Table 2 shows the average of the final aggregated weight computed by the Bayesian BWM.

Table 2 indicates that the estimation based on the proposed Bayesian model yields a meaningful result since the average of the estimated distribution is centered around the overall average of each individual preferences. The example shows that the output of the Bayesian BWM is valid and makes perfect sense. Keep in mind that the BWM obtains the weight of each individual first and then aggregate them by the arithmetic mean, while the Bayesian BWM computes the aggregated distribution and all the individual preferences at once using probabilistic modeling. The essential difference between the BWM and the Bayesian BWM is that we can compare the criteria *colorfully*. The current practice is to say a criterion is more important than one another if its average weight has a higher value; therefore, it is a *black and white* (or zero-one) decision.

We compare all pairs of criteria with each other using the credal ranking and visualize its outcome using a weight directed graph. Fig. 2 displays the credal ranking of criteria for selecting the cell phones. The nodes in this graph are the criteria and each edge $A \xrightarrow{d} B$ tells that criterion A is more important than B with the confidence d.

Based on Fig. 2, functionality is the most important criterion based on the opinions of all participants. At the other extreme, customer excitement and brand choice are the least desirable features for the participants in this experiment. Further, the degree of certainty about the relation of criteria is also evident. For instance, *technical features* is certainly more important than *customer excitement*, but it is more desirable than *basic requirement* with the confidence of 0.71.

As mentioned before, the credal ranking visualized in Fig. 2 can be changed into the conventional ranking merely by applying the threshold of 0.5 to the obtained confidences. The threshold can vary from one problem to one another, and it is entirely to the DMs' volition to opt for a particular threshold value. For instance, the confidence 0.71 between *technical feature* and *basic requirement* might be strong enough for the mobile selection problem, but it could not be regarded as strong if the study was on another problem. In other words, credal ranking could be shaped to represent

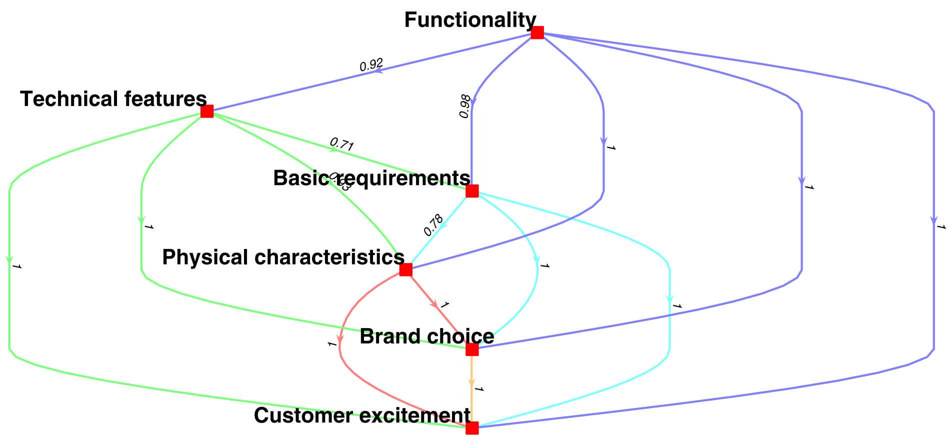

Fig. 2. The visualization of the credal ranking for the example of mobile phone selection criteria.

the ranking of criteria in different problems based on the DMs' desired confidence.

7. Conclusion

This paper presents a Bayesian model for the group decision-making within the BWM. The proposed method models the inputs of the BWM using the multinomial distribution and it is demonstrated that such a distribution would preserve the underlying idea of the original BWM. Further, the weight vector is modeled using the Dirichlet distribution. The proposed Bayesian model is able to compute the weight distribution of each individual in the group decision-making, and an aggregated final distribution representing the overall preferences of all DMs. The credal ranking of criteria is developed based on which each pair of criteria are assigned a relation and a confidence. The confidence which is computed based on a proposed Bayesian model shows the extent to which one is certain about the relation of the corresponding pair of criteria. In addition, the credal ranking is visualized using a weighted directed graph which shows the interrelation of criteria clearly.

The proposed Bayesian BWM is a promising method in the context of group decision making where one is interested in the collective opinion of a group, but at the same time, one could check the ranking of the weights in a probabilistic sense. The group will be more certain about the relation of two criteria if it is associated with a high confidence level while the relations with low confidence level should be interpreted more carefully.

There are several avenues to extend the current research. We aim to apply such modelings in other important MCDM methods. It is also interesting to investigate the role of outliers in different group BWM. Finally, it would be interesting to work on some other features of the Bayesian BWM such as consistency measure.

References

[1] Saaty TL. A scaling method for priorities in hierarchical structures. J Math Psychol 1977;15(3):234–81.
[2] Saaty TL. Decision making for leaders: the analytic hierarchy process for decisions in a complex world. RWS publications; 1990.
[3] Edwards W. How to use multiattribute utility measurement for social decisionmaking. IEEE Trans Syst Man Cybern 1977;7(5):326–40.
[4] Edwards W, Barron FH. Smarts and smarter: improved simple methods for multiattribute utility measurement. Organ Behav Hum Decis Process 1994;60(3):306–25.
[5] Mustajoki J, Hämäläinen RP, Salo A. Decision support by interval smart/swing-incorporating imprecision in the smart and swing methods. Decis Sci 2005;36(2):317–39.
[6] Ginevičius R. A new determining method for the criteria weights in multicriteria evaluation. Int J Inf TechnolDecis Making 2011;10(06):1067–95.
[7] Zavadskas EK, Podvezko V. Integrated determination of objective criteria weights in mcdm. Int J Inf TechnolDecis Making 2016;15(02):267–83.
[8] Triantaphyllou E. Multi-criteria decision making methods. In: Multi-criteria decision making methods: A comparative study. Springer, Boston, MA; 2000. p. 5–21.
[9] Rezaei J. Best-worst multi-criteria decision-making method. Omega 2015;53:49–57.
[10] Rezaei J. Best-worst multi-criteria decision-making method: some properties and a linear model. Omega 2016;64:126–30.
[11] Davey A, Olson D. Multiple criteria decision making models in group decision support. Group Decis Negotiat 1998;7(1):55–75.
[12] Wu Z, Xu J. Managing consistency and consensus in group decision making with hesitant fuzzy linguistic preference relations. Omega 2016;65:28–40.
[13] Yan H-B, Ma T, Huynh V-N. On qualitative multi-attribute group decision making and its consensus measure: a probability based perspective. Omega 2017;70:94–117.
[14] Sanayei A, Mousavi SF, Yazdankhah A. Group decision making process for supplier selection with vikor under fuzzy environment. Expert Syst Appl 2010;37(1):24–30.
[15] Groenendijk L, Rezaei J, Correia G. Incorporating the travellers' experience value in assessing the quality of transit nodes: a rotterdam case study. Case Stud Transp Policy 2018;6(4):564–76.
[16] Forman E, Peniwati K. Aggregating individual judgments and priorities with the analytic hierarchy process. Eur J Oper Res 1998;108(1):165–9.
[17] Forman E, Peniwati K. Aggregating individual judgments and priorities with the analytic hierarchy process. Eur J Oper Res 1998;108(1):165–9.
[18] Blagojevic B, Srdjevic B, Srdjevic Z, Zoranovic T. Heuristic aggregation of individual judgments in ahp group decision making using simulated annealing algorithm. Inf Sci 2016;330:260–73.
[19] Forman E, Peniwati K. Aggregating individual judgments and priorities with the analytic hierarchy process. Eur J Oper Res 1998;108(1):165–9.
[20] Abel E, Mikhailov L, Keane J. Group aggregation of pairwise comparisons using multi-objective optimization. Inf Sci 2015;322:257–75.
[21] Herrera-Viedma E, Herrera F, Chiclana F. A consensus model for multiperson decision making with different preference structures. IEEE Trans Syst ManCybern-Part A 2002;32(3):394–402.
[22] Morais DC, de Almeida AT. Group decision making on water resources based on analysis of individual rankings. Omega 2012;40(1):42–52.
[23] Ramanathan R, Ganesh L. Group preference aggregation methods employed in ahp: an evaluation and an intrinsic process for deriving members' weightages. Eur J Oper Res 1994;79(2):249–65.
[24] Vinogradova I, Podvezko V, Zavadskas E. The recalculation of the weights of criteria in mcdm methods using the bayes approach. Symmetry 2018;10(6):205.
[25] Ishizaka A, Labib A. Review of the main developments in the analytic hierarchy process. Expert Syst Appl 2011;38(11):14336–45.
[26] Grošelj P, Stirn LZ, Ayrilmis N, Kuzman MK. Comparison of some aggregation techniques using group analytic hierarchy process. Expert Syst Appl 2015;42(4):2198–204.
[27] Salo AA, Hämäläinen RP. On the measurement of preferences in the analytic hierarchy process. J Multi-Criteria Decis Anal 1997;6(6):309–19.
[28] Brunelli M, Rezaei J. A multiplicative best-worst method for multi-criteria decision making. Oper Res Lett 2019;47(1):12–15.
[29] Hafezalkotob A, Hafezalkotob A, Liao H, Herrera F. Interval multimoora method integrating interval borda rule and interval best-worst-method-based weighting model: case study on hybrid vehicle engine selection. IEEE Trans Cybern 2019.
[30] Rezaei J, van Wulfften Palthe L, Tavasszy L, Wiegmans B, van der Laan F. Port performance measurement in the context of port choice: an mcda approach. Manag Decis 2018;57(2):396–417.
[31] Groenendijk L, Rezaei J, Correia G. Incorporating the travellers' experience value in assessing the quality of transit nodes: a rotterdam case study. Case Stud Transp Policy 2018;6(4):564–76.

[32] Rezaei J, Hemmes A, Tavasszy L. Multi-criteria decision-making for complex bundling configurations in surface transportation of air freight. J Air Transp Manag 2017;61:95–105.
[33] Gupta H, Barua MK. A novel hybrid multi-criteria method for supplier selection among smes on the basis of innovation ability. Int J Logist ResAppl 2018;21(3):201–23.
[34] Kusi-Sarpong S, Gupta H, Sarkis J. A supply chain sustainability innovation framework and evaluation methodology. Int J Prod Res 2019;57(7):1990–2008.
[35] Vahidi F, Torabi SA, Ramezankhani M. Sustainable supplier selection and order allocation under operational and disruption risks. J Clean Prod 2018;174:1351–65.
[36] Ahmad WNKW, Rezaei J, Sadaghiani S, Tavasszy LA. Evaluation of the external forces affecting the sustainability of oil and gas supply chain using best worst method. J Clean Prod 2017;153:242–52.
[37] Ahmadi HB, Kusi-Sarpong S, Rezaei J. Assessing the social sustainability of supply chains using best worst method. Resour Conserv Recycl 2017;126:99–106.
[38] Rezaei J, Nispeling T, Sarkis J, Tavasszy L. A supplier selection life cycle approach integrating traditional and environmental criteria using the best worst method. J Clean Prod 2016;135:577–88.
[39] Rezaei J, Wang J, Tavasszy L. Linking supplier development to supplier segmentation using best worst method. Expert Syst Appl 2015;42(23):9152–64.
[40] Gupta H, Barua MK. Identifying enablers of technological innovation for indian msmes using best–worst multi criteria decision making method. Technol Forecast Soc Change 2016;107:69–79.
[41] Torabi SA, Giahi R, Sahebjamnia N. An enhanced risk assessment framework for business continuity management systems. Saf Sci 2016;89:201–18.
[42] Salimi N. Quality assessment of scientific outputs using the bwm. Scientometrics 2017;112(1):195–213.
[43] Salimi N, Rezaei J. Measuring efficiency of university-industry ph. d. projects using best worst method. Scientometrics 2016;109(3):1911–38.
[44] Ren J. Multi-criteria decision making for the prioritization of energy systems under uncertainties after life cycle sustainability assessment. Sustain Prod Consump 2018;16:45–57.
[45] Gupta H. Evaluating service quality of airline industry using hybrid best worst method and vikor. J Air Transp Manag 2018;68:35–47.
[46] Mi X, Tang M, Liao H, Shen W, Lev B. 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 2019.
[47] Hafezalkotob A, Hafezalkotob A. A novel approach for combination of individual and group decisions based on fuzzy best-worst method. Appl Soft Comput 2017;59:316–25.
[48] Mou Q, Xu Z, Liao H. An intuitionistic fuzzy multiplicative best-worst method for multi-criteria group decision making. Inf Sci 2016;374:224–39.
[49] Forbes C, Evans M, Hastings N, Peacock B. Statistical distributions. John Wiley & Sons; 2011.
[50] Kruschke J. Doing Bayesian data analysis: a tutorial with R, JAGS, and Stan. Academic Press; 2014.
[51] Gilks WR, Richardson S, Spiegelhalter D. Markov chain Monte Carlo in practice. CRC press; 1995.
[52] Plummer M. Jags: just another gibbs sampler. 2004.
[53] Moore RE, Kearfott RB, Cloud MJ. Introduction to interval analysis, 110. Siam; 2009.
[54] Yatsalo BI, Martínez L. Fuzzy rank acceptability analysis: a confidence measure of ranking fuzzy numbers. IEEE Trans Fuzzy Syst 2018;26(6):3579–93.
[55] Dubois D, Prade H. Ranking fuzzy numbers in the setting of possibility theory. Inf Sci 1983;30(3):183–224.
[56] Xie N-m, Liu S-f. Novel methods on comparing grey numbers. Appl Math Model 2010;34(2):415–23.

Please cite this article as: M. Mohammadi and J. Rezaei, Bayesian best-worst method: A probabilistic group decision making model, Omega, https://doi.org/10.1016/j.omega.2019.06.001