Annals of Operations Research
https://doi.org/10.1007/s10479-024-06064-8

ORIGINAL RESEARCH

Multicriteria decision support under uncertainty: combining outranking methods with Bayesian networks

Melodi Cebesoy[1] · Ceren Tuncer Şakar[1] · Barbaros Yet[2]

Received: 11 May 2023 / Accepted: 20 May 2024

Abstract

Assessing alternative solutions that have uncertain evaluations in conflicting multiple criteria is not straightforward. Probabilistic models such as Bayesian networks (BNs) can effectively model and represent the uncertainty in such problems, but they do not include built-in mechanisms to guide different decision makers (DMs) with varying preferences toward the final decision. We propose a systematic approach to combine outranking methods with BNs to provide decision support for solutions with multiple and conflicting criteria under uncertainty. The proposed approach is based on Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE), and it offers different levels of precision and flexibility to the DMs in assessing the solutions. Our approach includes graphical tools and summary metrics to enhance the presentation of its results to the DMs. We test our approach with a case study on supplier selection where the uncertainty in supplier performances is modeled with a BN. We demonstrate that the proposed approach can enable the joint use of multiple criteria techniques with probabilistic modeling techniques like BNs to provide decision support in complex environments including uncertainty.

Keywords Multicriteria decision making · Bayesian networks · Outranking · Uncertainty · PROMETHEE · Supplier selection

✉ Ceren Tuncer Şakar
cerents@hacettepe.edu.tr

Melodi Cebesoy
melodicebesoy@hacettepe.edu.tr

Barbaros Yet
byet@metu.edu.tr

[1] Department of Industrial Engineering, Hacettepe University, 06800 Ankara, Turkey

[2] Department of Cognitive Science, Graduate School of Informatics, Middle East Technical University, 06800 Ankara, Turkey

Published online: 11 June 2024

1 Introduction

A rational decision maker (DM) under uncertainty chooses the decision alternative with the maximum expected value. Modeling this choice requires an understanding of the DM's preferences, and reasoning with uncertainty to estimate the expected values of interest. Neither of these tasks are trivial when multiple criteria are involved in decisions and preferences. Usually, the criteria conflict with each other, and no solution has the best evaluations in all of them. When uncertainty is involved, computing the joint probability distribution of criteria evaluations is also challenging. Previous research on multicriteria decision making (MCDM) offers various methods to deal with the former challenge (see Greco et al., 2016; Cinelli et al., 2020), but often with limited or no support to deal with uncertainty. On the other hand, probabilistic graphical models, in particular Bayesian Networks (BNs), offer a general representation and powerful inference algorithms to deal with the latter. However, although BNs have been extended to deal with simple and sequential decision making problems, there has been little focus on providing a general approach to deal with multiple and conflicting criteria.

This paper focuses on decision making under uncertainty by combining the advances from the MCDM and BN domains. The main contribution of this paper is to present a general method to provide multicriteria decision support with BNs. We propose a systematic approach that combines outranking-type MCDM techniques with BN models. The proposed approach keeps the decision and uncertainty models separate, thus enabling modifications in either the uncertainty or the decision model without disturbing the other.

In particular, we extend Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE) (Brans et al., 1986) to work with BNs. Although there have been several attempts to combine MCDM approaches with BNs (Barton et al., 2020; Dohale et al., 2021; Fan et al., 2020; Kaya et al., 2023), previous studies lack a systematic and general approach to combine PROMETHEE with BNs. Extending the BN modelling toolbox with PROMETHEE is valuable as PROMETHEE offers several advantages compared to other MCDM methods including its flexibility in using different types of criteria, ranking procedures and preference functions. First of all, the criteria values in a decision problem may have different data types such as continuous, ordinal and binary. Unlike other widely-used MCDM methods such as weighted sum and Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), PROMETHEE does not require normalization which makes it easier to work with both continuous and discrete criteria. Secondly, PROMETHEE has different variations that can provide complete or partial ranking of solutions. We work with PROMETHEE I and II to provide partial and complete rankings. Thirdly, the DM can express detailed preferences through a variety of preference functions and fine-tune those preferences with customized indifference and preference thresholds. Furthermore, the thresholds of PROMETHEE are not abstract values like the concordance and discordance thresholds used by the other popular outranking method, Elimination and Choice Translating Reality (ELECTRE).

Applying outranking methods under uncertainty, such as combining PROMETHEE with BNs, leads to partial and complete rankings that also involve uncertainty. For example, when a complete ranking is sought with uncertain criteria evaluations, a solution can have probabilities of being ranked the first, second and so on, rather than having a specific rank. This uncertainty makes it challenging to interpret the results. Another contribution of this paper is graph, plot and score-based tools to present the results of the proposed approach, and to provide decision support in a concise way.

We demonstrate the proposed approach both with a simple example and a supplier selection case study that involves several suppliers and criteria. The case study pairs the proposed approach with a BN that has been previously developed for multicriteria supplier evaluation (Kaya & Yet, 2019). To account for the preferences of the DMs regarding the importance level of criteria, we work with different weight elicitation techniques and evaluate the sensitivity of the results to changes in preference weights.

In the remainder of this paper, Sect. 2 reviews the studies that handle uncertainty in MCDM, and also studies that use BNs and MCDM methods together. Section 3 presents the proposed approach, and Sect. 4 applies it to the supplier selection case study. We conclude the paper with discussions and future work in Sect. 5.

2 Related work

2.1 Handling uncertainty in MCDM

Monte Carlo Simulation (MCS) has been widely used for incorporating uncertainty in MCDM problems. Previous studies used MCS to sample from the probability distributions of weights or criteria values. For example, Baudry et al. (2018) employed MCS in Multi Actor Multiple Criteria Analysis to decide on a biofuel option. They used Analytic Hierarchy Process (AHP) to derive the weights of criteria and select the best option. Betrie et al. (2013) studied the problem of selecting alternative mine sites when criteria weights have uncertainty. They used AHP and PROMETHEE to rank the alternatives, and uncertainty in weights was handled by fitting probability distributions and sampling from them using MCS. Dorini et al. (2011) focused on uncertainties in model inputs and weights of different DMs when comparing two alternative solutions. They applied compromise programming and MCS on a case study about comparing two sustainable electricity generation options. Baležentis and Streimikiene (2017) used additive ratio assessment, weighted aggregated sum and TOPSIS to evaluate effective energy planning scenarios. In their study, they made sensitivity analysis with different criteria weights using MCS.

Stochastic Multicriteria Acceptability Analysis (SMAA) (Lahdelma et al., 1998) is an approach based on MCS to handle uncertain or insufficient information in multicriteria problems. It considers probability distributions to provide measures for decision making. SMAA does not strictly rank the solutions, but determines the ones that can be candidates for selection. It calculates acceptability indices between 0 and 1 for the solutions, and the best solutions are accepted as the ones with the highest acceptability indices for the top ranks. Lahdelma and Salminen (2001) extended the SMAA approach to consider any possible rank. This extended method, SMAA-2, aggregates rank acceptability indices with metaweights to present an overall acceptability measure called holistic acceptability index. There are also different versions of SMAA that are combined with other MCDM methods; readers are referred to Pelissari et al. (2020) and Tervonen and Figueira (2008) for further information and a detailed literature review.

Fuzzy sets and logic are used to handle vagueness and ambiguity regarding linguistic information in decision problems (Özkan et al., 2014). Montazar et al. (2013) proposed an approach that uses fuzzy triangular numbers and AHP in order to make performance assessments of irrigation alternatives with fuzzy parameters. Kilic et al. (2014) performed a two-phase decision making method to select an appropriate Enterprise Resource Planning (ERP) system. They applied fuzzy AHP to handle varying criteria preferences, and then employed TOPSIS to rank the alternatives. Pitchipoo et al. (2013) introduced an integrated model of fuzzy AHP and Grey Relational Analysis (GRA). Vagueness about criteria was modeled with fuzzy set theory and the alternatives were ranked with the help of GRA. Venkatesh et al. (2019) used fuzzy AHP to derive criteria weights, and fuzzy TOPSIS to find the best supplier for humanitarian supply chains.

The readers are referred to Durbach and Stewart (2012), Broekhuizen et al. (2015) and Pelissari et al. (2021) for more comprehensive reviews on MCDM approaches under uncertainty. Since our approaches are based on PROMETHEE, we provide a more specific review of studies that handled uncertainty in that family of methods in the next section.

2.1.1 Handling uncertainty in PROMETHEE

Early work in PROMETHEE with uncertainty focuses on modeling the uncertainty and variation among multiple domain experts. Mareschal (1986) used PROMETHEE I and II with the average ranking of different experts to reach the final decision in a project selection problem. In more recent studies, Yuen and Ting (2012) proposed to use a fuzzy approach in PROMETHEE II to cope with uncertainty in solution evaluations. Kuang et al. (2015) used Grey Theory with PROMETHEE II to produce a ranking of alternative source water protection strategies. Maghrabie et al. (2019) focused on problems with small amount of data and uncertain information on criteria weights. They used maximizing deviation method and Grey Systems Theory to estimate the weights, and PROMETHEE II and degrees of possibility to rank the solutions. The stochastic dominance approaches in Zhang et al. (2010) and Liu et al. (2011) were also applied with PROMETHEE. They created dominance matrices based on stochastic dominance degrees, and used these matrices with PROMETHEE to find the rank order of solutions.

MCS has been a widely used approach to include uncertainty in PROMETHEE calculations too. Hyde et al. (2003) used probability distributions to model criteria weights and performance in PROMETHEE II scores. They used MCS to run the models and reported the probability of solutions occupying different ranks. Doumpos and Zopounidis (2010) made sensitivity analysis of PROMETHEE II parameters and criteria preferences using MCS. Shakhsi-Niaei et al. (2011) proposed a two-phase method for project selection where the first phase evaluates the projects with PROMETHEE and MCS, and the second phase utilizes MCS with integer programming to select a final project portfolio. Gervásio and Simões Da Silva (2012) fitted probability distributions to criteria and used MCS and PROMETHEE II to calculate the probabilities of alternatives holding possible ranks.

As discussed in the previous section, SMAA is used with uncertain data and unknown preferences of the DMs, and different versions combined with outranking approaches are available in the literature. Corrente et al. (2014) combined SMAA-2 with PROMETHEE I and II methods. Their method asks the DM to provide partial preference information on criteria and preference thresholds. For each solution i, SMAA-PROMETHEE I calculates the frequency of solution i being preferred to every other solution k (denoted UP(i,k)) and the frequency of solution k being preferred to i (denoted DOWN(i,k)). Taking all available solutions into account, solutions with the highest UP and lowest DOWN values are reported as the best ones. In addition, SMAA-PROMETHEE II provides rank acceptability indices by deriving the probabilities of solutions occupying each rank, the central weight vector, and the confidence factor. However, although the SMAA-PROMETHEE II method is based on SMAA-2 rules, it does not use the holistic acceptability index of SMAA-2 or introduce an overall measure to generate a ranking of solutions. Hubinont (2016) proposed a SMAA version of the Geometrical Analysis for Interactive Aid (GAIA) (Brans & Mareschal, 1994), which was developed to project the locations of alternatives according to all criteria on a plane. Another study proposed an extension of SMAA to work with nonmonotonic criteria in a group decision making setting (Liao et al., 2022).

Among the studies reviewed, Gervásio and Simões Da Silva (2012) and Corrente et al. (2014) are the most related to our approach since we also work with PROMETHEE II formulas and calculate the probability of solutions occupying different ranks based on criteria distributions. However, we combine PROMETHEE with BNs to evaluate the probability of solutions being preferred to others, and use SMAA-2 rules. Additionally, our approach provides a systematic methodology to use PROMETHEE with BNs in a comprehensive decision making process. We present graph, plot and score-based tools to present the results, and introduce an extension of the approach for applying data-driven outranking.

2.2 BN models and MCDM methods

Influence Diagrams (IDs), which are extensions of BNs, have been proposed to deal with decision making problems (Howard & Matheson, 2005). An ID has additional types of nodes for decision and utility variables. The computation of IDs is more complex than BNs as the model needs to calculate the utility distribution for all decision alternatives, and several algorithms are available (Jensen et al., 1994; Zhang, 1998). In MCDM problems, the preferences of the DMs can be encoded in the parameters of the utility nodes of IDs. This requires eliciting a utility value for all states of utility nodes, or an additive or multiplicative function for the utility nodes. For example, Watthayu and Peng (2004) built an ID for an MCDM problem about commuting. Delcroix et al. (2013) used IDs for a recurrent MCDM problem where the decision is made by a group of experts. Their model predicted both criteria importance and distributions within the model. Barton et al. (2020) integrated multi-attribute value functions in a BN model in order to evaluate environmental design alternatives, and modeled the problem as an ID. Shahzad (2022) studied on IDs to overcome interdependency and uncertainty issues of traditional MCDM problems, and used AHP to calculate utility values of nodes. In general, the use of IDs in complex MCDM problems requires the creation of large utility tables, which is not a straightforward task.

Combining MCDM approaches with BNs can provide more flexibility as it can exploit a wider selection of BN inference algorithms rather than using algorithms specifically designed for solving IDs. In this case, a suitable MCDM technique, such as sorting or ranking, can be selected depending on the properties of the decision problem. The use of BNs with MCDM methods in the literature is limited. As examples, Fenton and Neil (2001) illustrated how BNs can be used to calculate posterior probability distributions and constraints that could be paired with MCDM techniques. Mazaheri et al. (2010) used a BN of a questionnaire to predict fuzzy preference weights of a DM, and then used a fuzzy version of TOPSIS to recommend a decision based on the weights. Several studies combined BNs and TOPSIS to provide decision support and risk assessment for maritime safety (Fan et al., 2020, 2023; Yang et al., 2021). Fan et al. (2020) presented a method using BNs and TOPSIS to determine the best strategies for preventing marine accidents. Yang et al. (2021) combined BNs and TOPSIS to aid ship detention decision in port state control inspections. Fan et al. (2023) used BNs and TOPSIS to assess the risk of maritime piracy and to guide prevention actions. Accident types were treated as criteria, and accident occurrence probabilities were used as criteria weights. Kaya and Yet (2019) worked on integrating BNs with the Decision Making Trial and Evaluation Laboratory Method (DEMATEL). Kaya et al. (2023) extended the study of Kaya and Yet (2019) to provide a ranking of suppliers based on their overall performances.

To achieve this, they applied TOPSIS to the evaluation matrix presented by the BN model whose causal graph was determined with DEMATEL. The evaluation matrix derived from the BN model included deterministic criteria values, and they used AHP to derive criteria weights. In their case study, Kaya et al. (2023) built a BN model for supplier selection with 5 criteria. They conducted a sensitivity analysis to evaluate the effect of having missing information regarding suppliers on the rankings obtained from their approach. Quan and Cho (2014) used a combination of BNs and AHP to recommend television programs. BNs were used to learn user preferences from data, and AHP was used as a weighted combination of these preferences to make recommendations. Weight elicitation and consistency evaluation elements of AHP were not included in the study. Dohale et al. (2021) presented a three-phase approach to select a production system. In the first stage, decision criteria are determined using Delphi method, and their weights are derived by Voting AHP in the second stage. In the last stage, a BN model is constructed for each alternative using relative weights and prior probabilities in order to calculate its selection probability.

As the reviewed studies show, previous attempts to pair MCDM approaches with BNs are generally made for specific decision problems, and do not offer general guidelines about how to use multicriteria techniques with BN posteriors. In addition, majority of them focus on AHP, which is impractical for large problems, or TOPSIS, which requires normalization of criteria values and does not enable detailed preference expressions. As opposed to those studies, we propose a systematic method that uses the advantages of outranking techniques in developing a general decision support framework on BN outputs.

3 Methodology

In this paper, we consider a decision problem where discrete alternatives evaluated with multiple criteria are to be partially or fully ranked, depending on the needs of the DM. The DM can judge the importance of criteria differently, thus criteria have importance weights. The criteria can be measured in different types of scales such as continuous, nominal and ordinal. In addition, we assume that the evaluations of alternatives are subject to uncertainty. In order to provide decision support in this problem, we propose a systematic approach to combine outranking approaches with BNs.

In Sect. 3.1, we provide background information on the methods we use in our approach. We present our approach to combine BNs and PROMETHEE in Sects. 3.2 and 3.3, our tools to present the results of this approach in Sect. 3.4, and an extension of this approach to data-driven PROMETHEE in Sect. 3.5. We conclude this section with a small illustrative example in Sect. 3.6.

3.1 Background information

This section includes technical information on BNs, PROMETHEE I and II methods, and the weight elicitation techniques we use.

3.1.1 Bayesian networks

A BN is a probabilistic graphical model that represents the conditional dependencies and joint probability distribution of a set of variables. The graphical structure of a BN is a directed acyclic graph in which nodes represent random variables, and edges represent direct causal and associational relations between those variables. When a directed edge connects nodes A and B as in $A \rightarrow B$, A is called the parent of B, and B is called the child of A. Each node X_i in the BN has a local probability distribution conditioned on its parents $pa(X_i)$ denoted by $P(X_i|pa(X_i))$. The graphical structure encodes conditional independency assertions that enable the joint probability distribution to be defined in a compact and factorized way based on these local probability distributions. Each node is conditionally independent of its non-descendants given their parents. The joint probability distribution of the random variables can be decomposed into the product of the conditional probabilities of each node given their parents as in (1).

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i|pa(X_i)) \tag{1}$$

Figure 1 shows a BN example where the local conditional probability distributions are shown as probability tables. The joint probability distribution of this BN can be computed by the product of those distributions as in (2).

$$\begin{aligned} &P(B, C1, C2, C3, F1, F2) \\ &\quad = P(C3|F2)\,P(C2|B, F2)\,P(C1|F1)\,P(F2|B, F1)\,P(F1|B)\,P(B) \end{aligned} \tag{2}$$

Efficient algorithms are available to compute the posterior probabilities when any subset of the nodes is instantiated by exploiting the conditional independence assertions (Lauritzen & Spiegelhalter, 1988).

In MCDM problems, BNs offer a suitable tool to model the causal and associational relations, and joint probability distribution between the criteria based on a combination of domain knowledge and data (Kaya & Yet, 2019; Yet et al., 2014). However, BNs do not incorporate DM preferences or decision making algorithms to reach a final solution in MCDM problems; they only compute the posterior probability distributions of their variables.

3.1.2 PROMETHEE

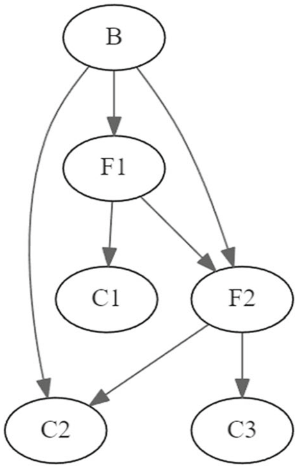
The PROMETHEE methodology proceeds by making pairwise comparisons between all solutions in terms of each criterion involved in the problem. It assigns a preference value between 0 and 1 to each comparison; and using these preference values and the importance weights of criteria, it calculates aggregated measures of preferability of solutions. Take two alternative solutions a_i and a_k, which have evaluations in m maximization-type (without loss of generality) criteria; $a_i = (a_{i1}, a_{i2}, \ldots, a_{im})$, $a_k = (a_{k1}, a_{k2}, \ldots, a_{km})$. To determine the preference strength of a_i over a_k with respect to criterion j, first the difference between their evaluations is found by $d_{ikj} = a_{ij} - a_{kj}$. A preference value is assigned to this difference using one of the six available preference functions. In Type I function, the preference value is 1 for all positive differences. In Type II function, the preference value is 0 until a positive indifference threshold is reached; after this threshold, the value is 1. In Type III function, the preference value linearly increases from 0 to 1 for positive differences, and after a preference threshold is exceeded, it remains at 1. There are three possible preference values in Type IV function; 0, 0.5 and 1, which are determined by two thresholds. Type V function resembles Type III in using a linear function from 0 to 1, but a threshold must be exceeded to start the climb. Lastly, Type VI function features a Gaussian function that goes from 0 to 1 nonlinearly. The first three functions are the most commonly used ones. Let the preference value of a_i over a_k with respect to criterion j be $P_j(d_{ikj})$. The aggregated preference index of a_i over

B	
0	1
0.40	0.60

P(B)

		F2	
F1	B	0	1
0	0	0.65	0.35
	1	0.40	0.60
1	0	0.20	0.80
	1	0.50	0.50

P(F2|B, F1)



	F1	
B	0	1
0	0.90	0.10
1	0.30	0.70

P(F1|B)

		C2		
B	F2	0	1	2
0	0	0.45	0.05	0.50
	1	0.10	0.70	0.30
1	0	0.20	0.10	0.70
	1	0.80	0.10	0.10

P(C2|B, F2)

	C3		
F2	0	1	2
0	0.30	0.50	0.20
1	0.03	0.07	0.90

P(C3|F2)

	C1		
F1	0	1	2
0	0.05	0.05	0.90
1	0.20	0.10	0.70

P(C1|F1)

Fig. 1 Bayesian network example

a_k is calculated as in (3), where w_j is the weight of criterion j. PROMETHEE works with pre-specified criteria weights that sum up to 1, so v_{ik} values are between 0 and 1. Values closer to 0 represent weak global preference of a_i over a_k and values closer to 1 represent a strong global preference of a_i over a_k.

$$v_{ik} = \sum_{j=1}^{m} P_j\left(d_{ikj}\right)w_j \tag{3}$$

After pairwise comparisons are conducted, PROMETHEE calculates overall preference indices for all solutions. In the presence of n available solutions, positive flow of a_i, which implies how strongly a_i outranks all other solutions, is denoted by φ_i^+ and calculated as in

(4). Negative flow φ_i^- of a_i is calculated as in (5), and it implies how strongly other solutions outrank a_i. A solution would be considered preferable if it has high positive flow and low negative flow.

$$\varphi_i^+ = \sum_{k=1}^{n} v_{ik} \tag{4}$$

$$\varphi_i^- = \sum_{k=1}^{n} v_{ki} \tag{5}$$

There are two classic PROMETHEE methods that work with the flows calculated in (4) and (5), PROMETHEE I and PROMETHEE II. In PROMETHEE I, a_i is preferred to a_k if one of the following conditions hold:

(i) $\varphi_i^+ > \varphi_k^+$ and $\varphi_i^- < \varphi_k^-$
(ii) $\varphi_i^+ = \varphi_k^+$ and $\varphi_i^- < \varphi_k^-$
(iii) $\varphi_i^+ > \varphi_k^+$ and $\varphi_i^- = \varphi_k^-$

Solutions a_i and a_k are incomparable if one of the following conditions hold:

(i) $\varphi_i^+ > \varphi_k^+$ and $\varphi_i^- > \varphi_k^-$
(ii) $\varphi_i^+ < \varphi_k^+$ and $\varphi_i^- < \varphi_k^-$

Lastly, there is an indifference relationship between the solutions if $\varphi_i^+ = \varphi_k^+ \text{and} \varphi_i^- = \varphi_k^-$.

As a result of PROMETHEE I, it may not be possible to achieve full ranking of solutions since there can be incomparability or indifference between some pairs of solutions. In PROMETHEE II, a final score is calculated as in (6) that can be used to achieve full ranking.

$$\varphi_i = \varphi_i^+ - \varphi_i^- \tag{6}$$

This φ_i value is called the net flow of solution a_i, and solutions are ranked in decreasing order of their net flows; r_i is the resulting rank of a_i. Due to the aggregation in (6), PROMETHEE II loses some level of outranking information given by positive and negative flows, but this allows it to produce a full rank list.

3.1.3 Elicitation of criteria weights

Many MCDM approaches, including PROMETHEE, represent the preferences of DMs for criteria in the form of weights, but they do not have a built-in weight derivation procedure. Among different types of weight derivation approaches, rank-based methods come forward as simple and practical alternatives. In those methods, the DM provides an importance ranking of criteria. These ranks are then used to calculate their weights, which are suitable to be used in PROMETHEE. Rank sum (RS), rank reciprocal (RR) and rank order centroid (ROC) are the most common rank-based formulas. Among those, ROC has been shown to perform best with respect to matching the true preferences of the DM and identifying the true best solution (Ahn, 2011; Roszkowska, 2013).

The formulas for deriving the weight of the criterion that occupies the t^{th} importance rank in RS, ROC and RR methods are given by (7), (8) and (9), respectively. ROC derives the weights by minimizing the maximum error of each weight from the centroid of all possible weights satisfying the given rank order of importance. RS assumes equal distance between weights of consecutive ranks while RR and ROC increase the difference between consecutive weights as the rank positions get higher. RR puts more weight on the first rank than the other methods.

$$w_t(RS) = \frac{n - t + 1}{\sum_{k=1}^{n} n - k + 1} \tag{7}$$

$$w_t(ROC) = \frac{1}{n} \sum_{k=t}^{n} \frac{1}{k} \tag{8}$$

$$w_t(RR) = \frac{1/t}{\sum_{k=1}^{n} 1/k} \tag{9}$$

3.2 Overview of combining PROMETHEE with BNs

Figure 2 shows an overview of the proposed approach for combining BNs and PROMETHEE. Under uncertainty, a solution a_i is a random variable and its evaluations for criteria have a probability distribution $p(a_{i1}, a_{i2}, \ldots, a_{im})$. We model this probability distribution in a BN model representing the problem, and query this model to get joint probabilities of evaluations under different conditions (Fig. 2a). For example, in a supplier selection problem, we build a causal BN model of the suppliers including the criteria variables, and use this model to get joint probability of criteria evaluations under different scenarios (see Sect. 4). Criteria weights are elicited from domain experts (Fig. 2b) and PROMETHEE operations for computing positive and negative flows are applied on these probability distributions (Fig. 2c). These operations result in a probability distribution of positive $p(\phi_i^+)$ and negative flows $p(\phi_i^-)$. The ranking

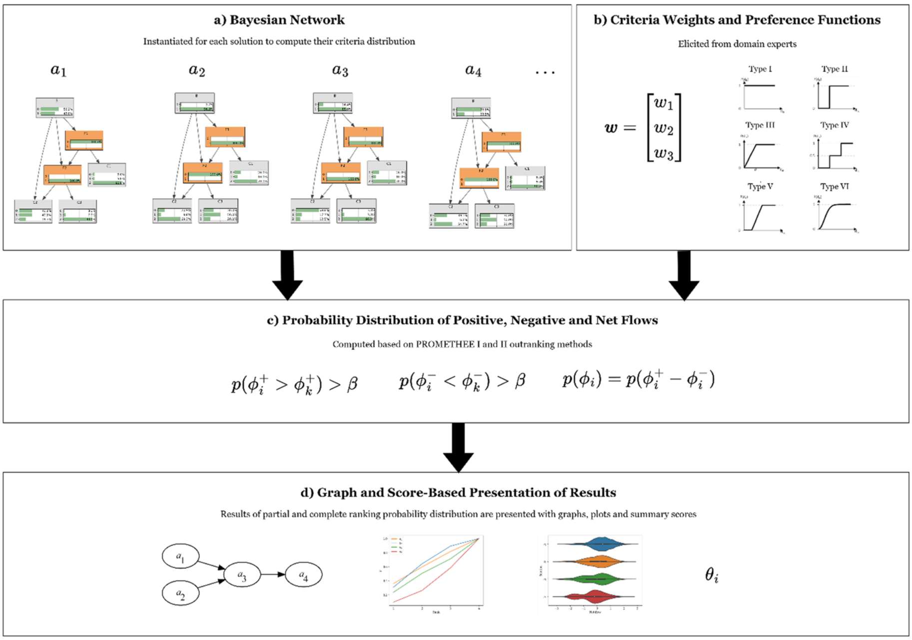


Fig. 2 Overview of method for combining BNs and outranking MCDM approaches

results obtained from these distributions are presented to the DM with graphs, plots and summary metrics (Fig. 2d).

In PROMETHEE I, the preference between a_i and a_k is determined based on these probability distributions and a probability threshold value β.

Solutions a_i and a_k are incomparable if one of the following conditions holds:

(i) $p(\phi_i^+ > \phi_k^+) > \beta$ and $p(\phi_i^- > \phi_k^-) > \beta$

(ii) $p(\phi_i^+ < \phi_k^+) > \beta$ and $p(\phi_i^- < \phi_k^-) > \beta$

If solutions are not incomparable, a_i is preferred to a_k if the following condition holds:

(i) $p(\phi_i^+ > \phi_k^+) > \beta$ or $p(\phi_i^- < \phi_k^-) > \beta$

Otherwise, there is indifference relationship between a_i and a_k.

The threshold value β is defined by the DM. A typical threshold value of $\beta = 0.5$ indicates that the condition is more likely than its inverse. Higher β values lead to more conservative preference options.

In PROMETHEE II, the probability distributions of net flows $p(\phi_i)$ are calculated with (10).

$$p(\phi_i) = p\left(\phi_i^+ - \phi_i^-\right) \tag{10}$$

As a result, a solution does not have a specific rank under uncertainty; it has different rank probabilities $p(r_i)$ according to its net flow distribution $p(\phi_i)$.

3.3 Computing positive, negative and net flow distributions from BNs

BNs are generative models. Once we have a BN model to compute the criteria evaluation distributions, the probability distributions $p\left(\phi_i^+\right)$, $p\left(\phi_i^-\right)$ and $p(r_i)$ can be estimated by generating samples from the BN. Various options are available to obtain samples from the posterior distribution of a BN model. The posterior distribution of a BN model with evidence can be calculated with exact algorithms such as junction tree, and the samples can be obtained from this posterior. Alternatively, rejection or importance sampling can be used to generate samples from the posteriors of an unpropagated BN. Following one of these approaches, we get samples a_{is} for each solution a_i and sample s. The PROMETHEE formulas (3)–(5) can be applied on those samples to calculate ϕ_{is}^+ and ϕ_{is}^- values. In PROMETHEE I, the probability distributions for the preference conditions can be estimated by counting the samples that the condition holds. Let S be the set of all samples, and $S_{\phi_i^+ > \phi_k^+}$ be the subset of S where the positive flow of solution a_i is greater than a_k, and $S_{\phi_i^- < \phi_k^-}$ be the subset of S where the negative flow of solution a_i is less than a_k.

$$S_{\phi_i^+ > \phi_k^+} = \left\{ s \in S | \phi_{is}^+ > \phi_{ks}^+ \right\} \tag{11}$$

$$S_{\phi_i^- < \phi_k^-} = \left\{ s \in S | \phi_{is}^- < \phi_{ks}^- \right\} \tag{12}$$

We estimate the probability that a_i has a higher positive flow than a_k as in (13).

$$\hat{p}\left(\phi_i^+ > \phi_k^+\right) = \frac{\left| S_{\phi_i^+ > \phi_k^+} \right|}{|S|} \tag{13}$$

Similarly, the probability that a_i has a smaller negative flow than a_k is estimated as in (14).

$$\hat{p}\left(\phi_i^- < \phi_k^-\right) = \frac{\left|S_{\phi_i^- < \phi_k^-}\right|}{|S|} \tag{14}$$

In PROMETHEE II, we calculate the sample net flows ϕ_{is} as in (15).

$$\phi_{is} = \phi_{is}^+ - \phi_{is}^- \tag{15}$$

Let r_{is} be the rank of a_i in sample s, and $S_{r_{it}}$ be the subset of S where a_i has the tth rank as shown in (16).

$$S_{r_{it}} = \{s \in S | r_{is} = t\} \tag{16}$$

The probability that a_i has rank t is estimated from the samples as in (17).

$$\hat{p}(r_i = t) = \frac{\left|S_{r_{it}}\right|}{|S|} \tag{17}$$

An alternative approach to compute PROMETHEE I and II results based on BN solutions would be to extend a BN model with the nodes representing deterministic PROMETHEE operations in (3)–(5), and to compute the posteriors of positive, negative and net flows directly within the BN model by using a hybrid BN algorithm such as dynamic discretization (Neil et al., 2007). The sampling approach, however, is simpler and offers the advantage of keeping the BN and MCDM models separate, thus allowing each model to be modified without affecting the other.

3.4 Presenting outranking results under uncertainty

The results of PROMETHEE I and II under uncertainty can be challenging for the DM to interpret. For example, the probability distributions of ranks can be overwhelming if there is a large number of alternatives. In this section, we propose graphical approaches and summary metrics (Fig. 2d) to present these results.

3.4.1 Presenting PROMETHEE I results under uncertainty

PROMETHEE I provides a partial ranking of solutions that is suitable to be presented in a directed graph, where nodes represent solutions and edges represent preferences. If there are multiple directed paths between two nodes a_i and a_j, and if one of those paths is composed of only one edge, i.e. $a_i \rightarrow a_j$, we remove this edge from the graph as it is redundant in presenting partial ranking. Its removal will present the partial ranking in a more concise way. In addition, the edges can be weighted in terms of the probability of the preference they represent, with their color and width adjusted according to these weights. The algorithm for generating the outranking graph is shown below, and an example is shown in Sect. 3.6.

Algorithm 1 Construction of outranking relation graph

Start with empty graph G composed of nodes $a_1, \ldots, a_n$
for each solution pair a_i and a_k:
  if a_i is preferred to a_k:
    add edge $a_i \rightarrow a_k$ to G
  end if
end for
for each edge $a_i \rightarrow a_k$:
  if there are multiple directed paths from a_i to a_k:
    remove edge $a_i \rightarrow a_k$
  end if
end for

3.4.2 Presenting PROMETHEE II results under uncertainty

Combining PROMETHEE II with BNs provides the probability distributions of net flows and solution rankings, which can be challenging to interpret. We use two graphical approaches to present these results. Firstly, we show the cumulative distribution plots of rankings as this highlights the dominating solutions for different ranks. Secondly, we use violin plots to show the probability distribution of net flows. The solutions are ranked in decreasing order of median net flows in the violin plot to highlight the ranking of solutions, and the net flow distributions shown in the plot highlight the uncertainty regarding this ranking. Examples of both plots are shown in Sect. 3.6.

3.4.3 Summary scores for PROMETHEE II results under uncertainty

Since PROMETHEE II under uncertainty provides a probability distribution of ranks rather than a definite ranking, multiple solutions will be ranked the highest with different probabilities in most circumstances. DMs may have different preferences; they may focus on a specific range of ranks, or consider some rank positions similar. To account for these cases, a weighted approach can be used to summarize the probability distribution of rankings into a single score, θ_i, for each solution a_i. The score θ_i is calculated as in (18) where c_t is the weight of the tth rank. Note that these weights are not about the importance of criteria; they are about assigning coefficients to the probabilities for different ranks so that we can obtain an aggregated measure.

$$\theta_i = \sum_{t=1}^{n} c_t p_{it} \tag{18}$$

3.5 Data-driven PROMETHEE

The proposed approach for computing PROMETHEE I and II from BN posteriors could be naturally expanded to apply PROMETHEE on a dataset. In this approach, we assume that a historical dataset of solutions is available, but the data-generating process of these solutions is not modeled or available. In this case, we can use statistical tests for the hypothesis $\phi_i^+ > \phi_k^+$. However, conducting multiple pairwise tests between the positive and negative flows of solutions is prone to false discoveries due to family-wise error rate. Tukey's test corrects for these errors when making pairwise comparison of the means. Dunn's test is a non-parametric alternative for this purpose. These tests only show statistical significance of the difference, but they do not provide information about the amount of difference between flows. Therefore, rather than a fully automated analysis based on statistical significance, we recommend the DMs to evaluate the confidence intervals of flow differences and assess the magnitude of the difference alongside the statistical tests.

3.6 Simple example

This section demonstrates the proposed approach by applying it to a simple example that has 3 criteria and 4 solutions based on the BN shown in Fig. 1. In the BN, nodes C1, C2 and C3 represent decision criteria, F1 and F2 represent the features that will be instantiated for different solutions, B represents a latent variable that will remain unobserved for all solutions. The feature values are instantiated for each solution and the posterior probabilities of criteria are computed using the junction tree algorithm. Table 1 shows the instantiated feature values and posterior criteria distributions of each solution; all criteria have 3 possible states. Our aim is to maximize all criteria and we use Type I preference function. The criteria weights are 0.5, 0.3 and 0.2 for C1, C2 and C3, respectively. The posterior distributions of positive, negative and net flows are computed by generating 10,000 samples from the posterior criteria distributions as described in Sect. 3.3.

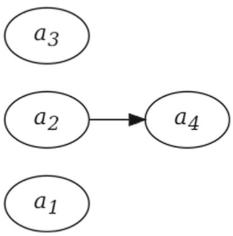
The partial ranking of solutions is obtained by using a threshold value of $\beta = 0.5$ on the posterior distributions of positive and negative flows as described in Sect. 3.3. Figure 3 shows the partial ranking graph prepared by the algorithm shown in Sect. 3.4.1. Solution a_2 is preferred to a_4, and there is no preference relation between the other solutions.

The probability distributions of rankings were computed based on the net flow distributions as described in Sect. 3.3. Figure 4a shows the cumulative distribution of rankings of each alternative. In this example, solution a_1 has the first, second, third and fourth rank with

Table 1 Posterior criteria distributions of solutions

Solution	Feature 1	Feature 2	Criterion 1			Criterion 2			Criterion 3		
			0	1	2	0	1	2	0	1	2
a_1	0	0	0.50	0.30	0.20	0.37	0.30	0.33	0.70	0.20	0.10
a_2	0	1	0.50	0.30	0.20	0.36	0.07	0.57	0.03	0.07	0.90
a_3	1	0	0.20	0.10	0.70	0.54	0.30	0.16	0.70	0.20	0.10
a_4	1	1	0.20	0.10	0.70	0.68	0.09	0.23	0.03	0.07	0.90

Fig. 3 Partial ranking of solutions



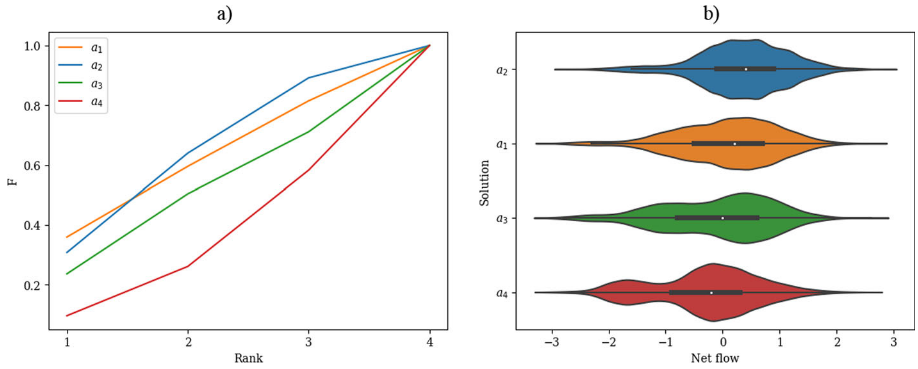


Fig. 4 **a** Cumulative distributions of solution rankings, **b** Probability distributions of net flows

0.36, 0.24, 0.22 and 0.18 probability. Solution a_4 has the first, second, third and fourth rank with 0.09, 0.17, 0.32 and 0.42 probability. Figure 4b shows the net flow distributions of the solutions. Solution a_2 has the highest mean net flow.

Figure 4a shows that a_1 and a_4 have the highest and lowest probability of having the highest rank, respectively. However, this probability alone is not sufficient to rank the solutions. For instance, even though a_2 has a lower probability of being the first than a_1, the probability of it being in top two or three ranks is higher than a_1. Therefore, a_2 can be considered as a better alternative for DMs who wish to consider multiple possibilities, whereas a_1 can be better for DMs who focus on the top rank and assign a relatively higher weight to the best possible outcome. The weighted score θ_i described in Sect. 3.4.3 summarizes the performance of solution a_i in a single measure. Table 2 shows the weight c_t of each rank t, the score θ_i of each solution a_i, and its rank in parentheses, for different rank weighting methods. If the DM assigns rank weights with RS, which assumes equal distance between the weights of different ranks, a_2 will have the highest rank, followed by a_1. However, suppose the DM considers that the first rank is considerably more important than others, and uses RR or ROC, which assigns about twice as much as weight to the first rank than the second one. In these cases,

Table 2 θ_i scores and rank weights with RS, ROC and RR methods

		Weighting method for ranks of solutions			
		RS	ROC	RR	ROC *first 2 ranks*
c_t	c_1	0.40	0.52	0.48	0.75
	c_2	0.30	0.27	0.24	0.25
	c_3	0.20	0.15	0.16	0
	c_4	0.10	0.06	0.12	0
θ_i	θ_1	0.27 (2)	0.30 (1)	0.29 (1)	0.33 (1)
	θ_2	0.29 (1)	0.29 (2)	0.28 (2)	0.31 (2)
	θ_3	0.25 (3)	0.24 (3)	0.24 (3)	0.25 (3)
	θ_4	0.19 (4)	0.17 (4)	0.19 (4)	0.11 (4)

a_1 will have the highest rank, followed by a_2. Solutions a_3 and a_4 has the third and fourth rank in all weighting methods in this example.

Suppose a DM is only interested in the first N ranks that corresponds to having more tolerance for risk than the case of considering all ranks. In this case, this DM can apply the rank-based weighting methods for only the first N ranks and use 0 weights for the others. This would require changing n with N in (7)–(9), and assigning 0 weights to $c_{N+1}, \ldots, c_n$. For example, the last column of Table 2 shows the weights and scores for the ROC approach when the DM considers only the two highest ranks.

4 Case study and results

This section applies the proposed approach to a case study of supplier selection, which is an MCDM problem that has been widely studied in the literature (see Govindan et al., 2015; Zimmer et al., 2016; Chai & Ngai, 2020; Rashidi et al., 2020; Saputro et al., 2022; Cui et al., 2023). We use a BN model for multicriteria supplier evaluation that was previously developed with domain experts by Kaya and Yet (2019). The BN model used here aims to evaluate suppliers based on seven criteria: product quality, cost, delivery performance, quality system certifications, flexibility, cooperation, and reputation. Among these criteria, cost and quality system certifications can be directly observed, so alternatives' performances are deterministic on these criteria. The other criteria are latent variables that are estimated by indirect measurements. The cost criterion is minimized whereas all other criteria are maximized. All criteria variables have five ordinal states; VL—very low, L—low, M—medium, H—high and VH—very high. The BN model was instantiated with the data of 10 different suppliers, and the posterior probability distributions of criteria were computed (see Table 3). The preference functions of the criteria and the corresponding threshold values were determined with a domain expert who was also involved in development of the BN. For product quality, cost and reputation, Type I function was selected. For quality system certifications, Type II function was selected with an indifference threshold of 1. So, for this criterion, only differences of 2 levels or higher are accounted for. For delivery performance, flexibility and cooperation, Type III function was used with preference thresholds of 3, 2 and 2 levels, respectively.

The weights of the criteria were obtained by asking the domain expert to rank the criteria based on their importance and applying the ROC formula. As a result, the weights of product quality, delivery performance, cost, quality system certifications, flexibility, cooperation, and reputation were realized as 0.370, 0.228, 0.156, 0.109, 0.073, 0.044 and 0.020, respectively. We also employed RS and RR in conducting sensitivity analysis of the weights.

We generated 1000 samples from the posteriors of the BN model, and computed the probability distribution of positive, negative and net flows based on these samples as described in Sect. 3.3. The partial ranking of solutions was obtained by using a threshold value of $\beta = 0.5$ on the posterior distributions of positive and negative flows as described in Sect. 3.3. Figure 5 shows the partial ranking graph. A directed arc from a supplier to another means that the former one outranks the latter.

Product quality, delivery performance and cost are the most important criteria for the DM, and the best suppliers S2 and S6 have good performances in those criteria. S3 is also a good alternative as it has acceptable performance in all criteria. S4 and S8 cannot be differentiated since they have advantages in different criteria. S4 is better at product quality and delivery performance, and S8 is better at cost and delivery performance. S10 is the worst alternative because it offers a medium-quality product (with 0.999 probability) with a very high cost.

Table 3 Probability distributions of the suppliers in each criterion

	Product quality					Cost					Delivery performance					Quality system certifications				
	VL	L	M	H	VH	VL	L	M	H	VH	VL	L	M	H	VH	VL	L	M	H	VH
S1	0.000	0.001	0.068	0.825	0.106	0.000	0.000	0.000	1.000	0.000	0.001	0.029	0.679	0.291	0.000	0.000	0.000	0.000	1.000	0.000
S2	0.000	0.001	0.189	0.682	0.128	0.000	0.000	1.000	0.000	0.000	0.020	0.164	0.419	0.322	0.075	0.000	0.000	0.000	0.000	1.000
S3	0.000	0.001	0.013	0.626	0.360	0.000	0.000	0.000	1.000	0.000	0.000	0.025	0.601	0.374	0.000	0.000	0.000	0.000	1.000	0.000
S4	0.000	0.001	0.280	0.706	0.013	0.000	0.000	0.000	1.000	0.000	0.000	0.007	0.132	0.487	0.374	0.000	0.000	1.000	0.000	0.000
S5	0.000	0.000	0.788	0.207	0.005	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.015	0.899	0.086	0.000	0.000	1.000	0.000	0.000
S6	0.001	0.003	0.336	0.610	0.050	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.003	0.673	0.324	0.000	1.000	0.000	0.000	0.000
S7	0.000	0.001	0.026	0.628	0.345	0.000	0.000	0.000	1.000	0.000	0.010	0.864	0.126	0.000	0.000	0.000	1.000	0.000	0.000	0.000
S8	0.000	0.028	0.972	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.005	0.822	0.173	0.000	0.000	0.000	0.000	1.000
S9	0.000	0.000	0.437	0.558	0.005	0.000	0.000	0.000	1.000	0.000	0.002	0.064	0.390	0.443	0.101	0.000	0.000	0.000	1.000	0.000
S10	0.000	0.000	0.999	0.001	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.004	0.126	0.530	0.340	0.000	0.000	0.000	1.000	0.000

	Flexibility					Cooperation					Reputation				
	VL	L	M	H	VH	VL	L	M	H	VH	VL	L	M	H	VH
S1	0.186	0.801	0.013	0.000	0.000	0.000	0.005	0.994	0.000	0.000	0.008	0.328	0.663	0.001	0.000
S2	0.027	0.127	0.295	0.342	0.209	0.000	0.000	0.108	0.820	0.072	0.000	0.000	0.997	0.003	0.000
S3	0.012	0.085	0.266	0.379	0.258	0.000	0.001	0.078	0.831	0.090	0.001	0.102	0.896	0.001	0.000
S4	0.001	0.096	0.655	0.248	0.000	0.000	0.001	0.002	0.323	0.674	0.000	0.000	0.999	0.001	0.000
S5	0.226	0.766	0.008	0.000	0.000	0.002	0.106	0.562	0.317	0.013	0.001	0.015	0.141	0.433	0.411
S6	0.000	0.000	0.002	0.475	0.523	0.001	0.014	0.844	0.141	0.000	0.011	0.368	0.621	0.000	0.000
S7	0.038	0.205	0.399	0.282	0.076	0.001	0.001	0.253	0.708	0.037	0.000	0.000	0.999	0.000	0.000
S8	0.000	0.076	0.622	0.302	0.000	0.000	0.000	0.034	0.519	0.447	0.002	0.154	0.843	0.001	0.000
S9	0.190	0.796	0.014	0.000	0.000	0.000	0.000	0.478	0.501	0.021	0.000	0.000	0.998	0.002	0.000
S10	0.176	0.809	0.015	0.000	0.000	0.000	0.000	0.477	0.503	0.020	0.010	0.370	0.619	0.001	0.000

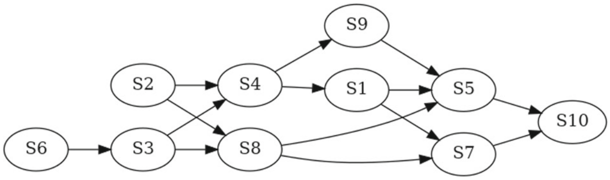


Fig. 5 Partial ranking of suppliers

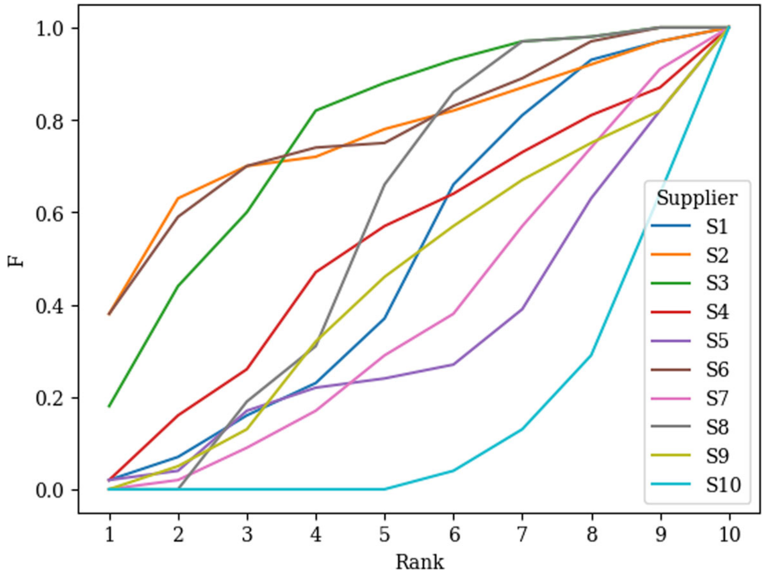


Fig. 6 Cumulative probability distribution of supplier ranking

The probability distributions of complete supplier rankings were computed based on the net flow distributions as described in Sect. 3.3. Figure 6 shows the cumulative distribution of rankings of each alternative, and Fig. 7 shows the net flow distributions of the solutions. In Fig. 6, we see that S2 and S6 both have the highest probability for the first rank, but their plots intersect at several points in Fig. 6, showing that the best between these two alternatives changes for different ranks. We can also see that the plot of S3, which has a lower probability for the first rank than S2 and S6, manages to lie above the plots of these alternatives starting from the fourth rank. Figure 7 is also helpful in comparing suppliers since it illustrates the full distributions of net flows including density and range.

In general, we observe that suppliers have varying levels of performance throughout their plots and it is not straightforward to achieve a final ranking of them. We compute the summary score (θ_i) of the suppliers as described in Sect. 3.6 by using RS, ROC or RR weights for ranks. Table 4 shows these weight vector alternatives; the last three columns also show the weights when just the first three ranks are considered by the DM. The weights for the top three ranks are provided as an example, the DM of the problem can select a different cut-off point as well. Table 5a, b show θ_i scores and rankings of the suppliers based on these rank weights.

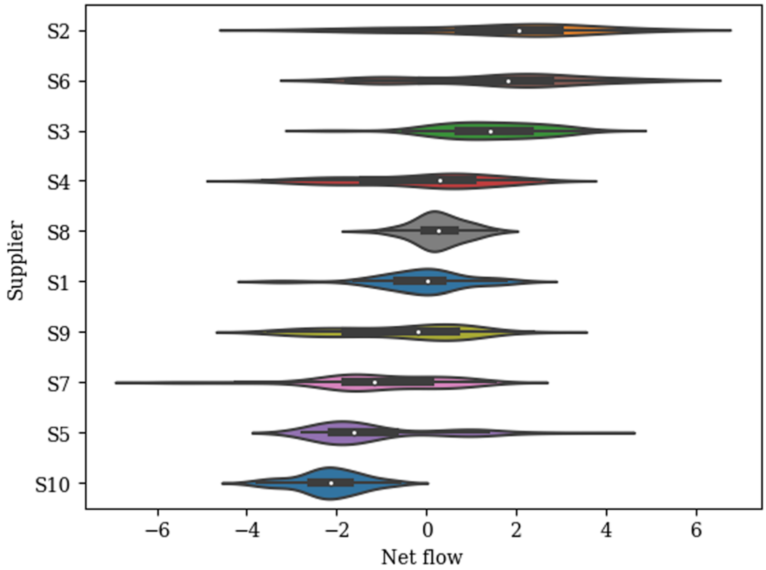


Fig. 7 Probability distribution of net flows

Table 4 Rank weights for RS, ROC and RR approaches for supplier selection

Rank (t)	c_t (RS)	c_t (ROC)	c_t (RR)	c_t (RS) *first 3 ranks*	c_t (ROC) *first 3 ranks*	c_t (RR) *first 3 ranks*
1	0.182	0.293	0.341	0.500	0.611	0.545
2	0.164	0.193	0.171	0.333	0.278	0.273
3	0.145	0.143	0.114	0.167	0.111	0.182
4	0.127	0.110	0.085	0.000	0.000	0.000
5	0.109	0.085	0.068	0.000	0.000	0.000
6	0.091	0.065	0.057	0.000	0.000	0.000
7	0.073	0.048	0.049	0.000	0.000	0.000
8	0.055	0.034	0.043	0.000	0.000	0.000
9	0.036	0.021	0.038	0.000	0.000	0.000
10	0.018	0.010	0.034	0.000	0.000	0.000

When we examine the ranking of the suppliers with ROC-weighted θ_i scores in Table 5a as an example, we observe that S2 is the best supplier followed by S6; these are the suppliers with the highest probabilities for the first rank. They are followed by S3 and S4 since they have high probabilities for the first two ranks. The rest of the rank list is not straightforward though. S8, for example, has 0 probability for occupying the first or the second rank, whereas S1 and S5 both have positive probabilities there. However, S8 is positioned higher than S1 and S5 since it improves its performance noticeably in the lower ranks. When we compare S5 and S7, we see that S5 has higher probabilities for the first three ranks, but S7 performs better for the remaining four ranks and it outranks S5. Such patterns can be observed from

Table 5 θ_i scores and supplier rankings based on θ_i scores

Rank	Supplier	θ_i(RS)	Supplier	θ_i(ROC)	Supplier	θ_i(RR)
(a) All ranks are considered						
1	S6	0.143	S2	0.185	S2	0.196
2	S3	0.142	S6	0.184	S6	0.194
3	S2	0.142	S3	0.161	S3	0.153
4	S8	0.109	S4	0.093	S4	0.085
5	S4	0.101	S8	0.089	S8	0.074
6	S1	0.095	S1	0.079	S1	0.073
7	S9	0.087	S9	0.072	S9	0.067
8	S7	0.076	S7	0.058	S5	0.062
9	S5	0.069	S5	0.057	S7	0.058
10	S10	0.038	S10	0.023	S10	0.039
(b) The first three ranks are considered						
1	S2	0.285	S2	0.309	S2	0.288
2	S6	0.278	S6	0.303	S6	0.285
3	S3	0.203	S3	0.200	S3	0.198
4	S4	0.073	S4	0.062	S4	0.067
5	S1	0.042	S1	0.036	S1	0.041
6	S5	0.038	S5	0.032	S5	0.040
7	S8	0.032	S9	0.023	S8	0.035
8	S9	0.030	S8	0.021	S9	0.028
9	S7	0.018	S7	0.013	S7	0.018
10	S10	0.000	S10	0.000	S10	0.000

Fig. 6. In Table 5a, we can see that the best two suppliers are S2 and S6 with ROC and RR rank weighting methods, but S6 and S3 with RS. This is due to the fact that RS puts less importance on the first rank than ROC and RR, and assumes equal distance between consecutive ranks. The performance of S2 is higher than S3 at the first rank, but lower at the second, third and fourth ranks. A conservative DM in the face of uncertainty can prefer the RS ranking since it treats the possible outcomes more evenly. On the other hand, a DM who is more interested in the probabilities for the first rank can choose the ranking of ROC or RR.

4.1 Data-driven outranking

The proposed approach can be used for data-driven outranking when data about previous solutions are available as described in Sect. 3.5. We generated a dataset of 100 samples from the BN to simulate a case of learning preferences from limited data, and analysed this dataset assuming that the data-generating process was not known. We conducted Tukey's Test with 95% confidence level for this purpose and tested the following hypotheses for all $a_i - a_k$ pairs.

$$H_0^{\mathrm{T1}}:\ \varphi_i^+ = \varphi_k^+$$

$$H_0^{T2}:\ \varphi_i^- = \varphi_k^-$$

Table 6 shows the mean difference between the pairs and the corresponding confidence intervals. The differences are calculated with the supplier in the row minus the supplier in the column, and the asterisk symbols show the statistically significant differences corresponding to a p value of less than 0.05. For example, the S2-S1 cell for positive flows shows that the mean positive flow of S2 is significantly higher than that of S1. The same cell for negative flows shows that the mean negative flow of S2 is significantly lower than that of S1 this time, so we conclude that S2 outranks S1. Figure 8 shows the results of this approach. In this example, the only difference between the data-driven and model-driven outranking approach was in the preference between S1 and S4. While S4 is preferred over S1 in partial ranking of the suppliers obtained from the BN model, the data-driven approach was not able to identify a statistically significant difference between the positive and negative flows of those alternatives in the given sample as the differences were small.

4.2 Sensitivity analysis on criteria weights

Since the results of our approach, as in many MCDM methods, depend on the weights of criteria, we make sensitivity analysis to see if our results are robust to small changes in criteria weights. Rather than changing the ROC weights randomly, we apply RS and RR methods to the criteria ranking of our supplier selection expert. With RS, the weights of product quality, delivery performance, cost, quality system certificates, flexibility, cooperation and reputation are realized as 0.250, 0.214, 0.179, 0.143, 0.107, 0.071 and 0.036, respectively. With RR, these weights are 0.386, 0.193, 0.129, 0.096, 0.077, 0.064 and 0.055. When we switch from ROC to RS, the most evident difference is in the weight of product quality, which becomes quite lower. On the other hand, when we use RR, this weight increases, and the weights of delivery performance and cost decreases. There are some other differences as well.

Figure 9a illustrates the outranking relations of PROMETHEE I calculated with RS criteria weights. With ROC weights, S2 and S6 were the best alternatives, and S8 was outranked by S6. With RS weights, S3, S6 and S8 have a tie. There are also other differences such as some strict preference relations turning into indifference or incomparability, or vice versa. However, the general picture is not substantially different from Fig. 5, and there are no reversals in outranking relations. In Fig. 9b, we see the outranking relations of the suppliers when RR method is performed for criteria weights. This time, the best suppliers are S2, S3 and S6 together. S4 and S8 cannot outrank each other, similar to the ROC results. Different from the ranking with ROC criteria weights, S9 outranks S1, and S7 outranks S5 in Fig. 9b. Again, we observe that there are no substantial differences.

Since PROMETHEE II provides complete rankings rather than partial, we can compare the similarity of its rankings with ROC and RS/RR weights for criteria using Kendall rank correlation coefficient (Kendall's Tau). Kendall's Tau measures the similarity between two rankings of the same elements. It can take values between -1 and 1; with values -1, 1 and 0 corresponding to perfectly opposite rankings, exactly the same rankings and no relationship between the rankings, respectively. Kendall's Tau between the rankings of ROC criteria weights and RS criteria weights when we use RS, ROC and RR rank weights are 0.644, 0.777 and 0.777, respectively. On the other hand, when we compare the rankings of ROC and RR criteria weights, Kendall's Tau coefficients are 0.822, 0.777 and 0.733, respectively. We can conclude that there is an acceptable level of similarity between the rankings with different weighting methods, especially ROC and RR. This is expected since they both put greater emphasis on the top ranks compared to RS.

Table 6 Tukey's test results for data-driven positive and negative flows of supplier pairs

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
+ Flows										
S1	–									
S2	(0.80;1.52) 1.16*	–								
S3	(0.51;1.23) 0.87*	(−0.66;0.06) −0.30	–							
S4	(−0.36;0.36) 0.00	(−1.53; −0.81) −1.17*	(−1.23; −0.51) −0.87*	–						
S5	(−0.81; −0.09) −0.45*	(−1.97; −1.25) −1.61*	(−1.67; −0.95) −1.31*	(−0.80; −0.08) −0.44*	–					
S6	(0.98;1.70) 1.34*	(−0.18;0.54) 0.18	(0.12;0.83) 0.47*	(0.99;1.71) 1.35*	(1.43;2.15) 1.79*	–				
S7	(−0.18;0.54) 0.18	(−1.34; −0.63) −0.99*	(−1.05; −0.33) −0.69*	(−0.18;0.54) 0.18	(0.26;0.98) 0.62*	(−1.52; −0.81) −1.16*	–			
S8	(0.42;1.14) 0.78*	(−0.74; −0.02) −0.38*	(−0.44;0.27) −0.08	(0.43;1.15) 0.79*	(0.87;1.59) 1.23*	(−0.92; −0.20) −0.56*	(0.25;0.96) 0.61*	–		
S9	(−0.64;0.07) −0.28	(−1.81; −1.09) −1.45*	(−1.51; −0.79) −1.15*	(−0.64;0.08) −0.28	(−0.20;0.52) 0.16	(−1.99; −1.27) −1.63*	(−0.82; −0.10) −0.46*	(−1.43; −0.71) −1.07*	–	
S10	(−1.14; −0.43) −0.79*	(−2.31; −1.59) −1.95*	(−2.01; −1.29) −1.65*	(−1.14; −0.42) −0.78*	(−0.70;0.02) −0.34	(−2.49; −1.77) −2.13*	(−1.32; −0.60) −0.96*	(−1.93; −1.21) −1.57*	(−0.86; −0.14) −0.50	–
−Flows										
S1	–									
S2	(−0.97; −0.33) −0.65*	–								
S3	(−0.98; −0.34) −0.66*	(−0.32;0.32) 0.00	–							
S4	(−0.31;0.33) 0.01	(0.34;0.98) 0.66*	(0.35;0.99) 0.67*	–						
S5	(0.28;0.92) 0.60*	(0.93;1.57) 1.25*	(0.93;1.58) 1.26*	(0.27;0.91) 0.59*	–					
S6	(−0.68; −0.03) −0.35*	(−0.02;0.62) 0.30	(−0.02;0.62) 0.30	(−0.68; −0.04) −0.36*	(−1.27; −0.63) −0.95*	–				
S7	(0.80;1.44) 1.12*	(1.45;2.09) 1.77*	(1.46;2.10) 1.78*	(0.79;1.43) 1.11*	(0.20;0.84) 0.52*	(1.15;1.79) 1.47*	–			
S8	(0.10;0.74) 0.42*	(0.75;1.39) 1.07*	(0.76;1.40) 1.08*	(0.09;0.73) 0.41*	(−0.50;0.14) −0.18	(0.45;1.09) 0.77*	(−1.02; −0.38) −0.70*	–		
S9	(−0.14;0.50) 0.18	(0.51;1.15) 0.83*	(0.52;1.16) 0.84*	(−0.15;0.49) 0.17	(−0.74; −0.10) −0.42*	(0.21;0.85) 0.53*	(−1.26; −0.62) −0.94*	(−0.56;0.08) −0.24	–	
S10	(0.98;1.62) 1.30*	(1.63;2.27) 1.95*	(1.64;2.28) 1.96*	(0.97;1.61) 1.29*	(0.38;1.02) 0.70*	(1.33;1.97) 1.65*	(−0.14;0.50) 0.18	(0.56;1.20) 0.88*	(0.80;1.44) 1.12*	–

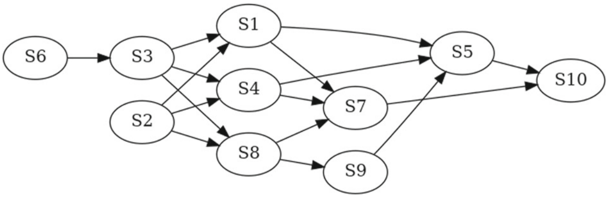


Fig. 8 Partial ranking of suppliers with data-driven outranking approach

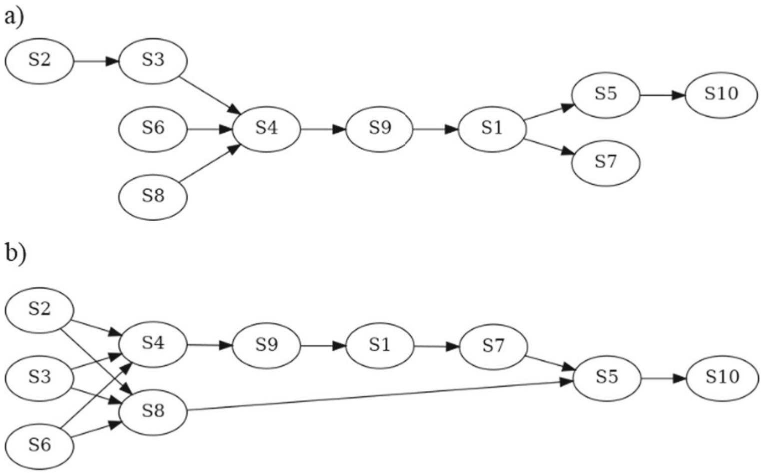


Fig. 9 Outranking relation of suppliers with **a** RS-weights, **b** RR-weights

For the complete rankings provided by PROMETHEE II, we conduct further sensitivity analysis to determine allowable ranges for criteria weights for the results to remain stable. We use the weight stability intervals procedure by Mareschal (1988) for this task by modifying it to be applicable with uncertain criteria evaluations. This procedure is developed to find intervals for criteria weights so that the given ranks of solutions according to additive utility functions do not change, and it can be applied with PROMETHEE II scores. It uses the differences in criteria values between successive solutions in the rank list to find these intervals. The details of the procedure can be seen in Mareschal (1988).

The weight stability intervals procedure can only work with a single sample, so we need to enhance it to work for PROMETHEE II under uncertainty. Firstly, using the basic procedure, we construct the interval of each criterion weight in each sample. Each sample produces its separate ranking of solutions, so we arrive at 1000 intervals for each criterion weight. Next, these intervals need to be aggregated into an overall interval for each criterion. We form these intervals with the values that appear in at least a given percentage of all samples. Since all the intervals in the samples are formed around the original weights, these original weights appear in 100% of the intervals. As we move away from the original weights, the percentage of samples that contain the value in consideration gets smaller. In line with the logic of confidence intervals, we use 95% as the cut-off value. Taking ROC weights as the original criteria weights, Table 7 reports the resulting aggregated intervals we obtain. These results suggest that the weights of product quality, delivery performance, and cost should be

Table 7 Weight stability intervals for ROC weights

Criterion	ROC weight	95% Weight stability interval
Product quality	0.370	[0.352–0.391]
Delivery performance	0.228	[0.175–0.262]
Cost	0.156	[0.119–0.196]
Quality system certificates	0.109	[0.018–0.199]
Flexibility	0.073	[0.000–0.171]
Cooperation	0.044	[0.000–0.223]
Reputation	0.020	[0.000–0.352]

set carefully since they have relatively narrow ranges. On the other hand, the ranking list is not so sensitive to changes in the weights of other criteria, so uncertainties in those areas can be tolerated better.

5 Conclusions

This paper proposes a systematic approach to combine BNs and outranking approaches to support multicriteria decision problems under uncertainty. The proposed approach is applied to a BN model that has been developed for supporting supplier selection decisions in an automobile manufacturer. Our approach enhances PROMETHEE I and II to provide partial and complete ranking with the probability distributions of decision criteria obtained from BNs. The result of partial ranking can be useful for DMs who need to stratify alternative solutions according to their performance without forcing a strict ranking. On the other hand, the result of complete ranking gives the DMs the chance to observe the overall performance of the solutions, as well as the performance in the best-case and worst-case scenarios. The results of partial and complete rankings are shown in graphs, cumulative distribution figures, and violin plots to demonstrate the preferences regarding solutions and the associated uncertainty in a concise way. We also present a summary score to summarize the performance of solutions based on their ranking distributions.

Our approaches provide a systematic and flexible way to combine a widely used MCDM method with probabilistic generative models such as BNs. They can work with different types of criteria like continuous, nominal and ordinal. In addition, the results of the sensitivity analysis on the weights of criteria can be used to determine which criteria need the most careful evaluation. The proposed approaches also extend to applying data-driven PROMETHEE when only samples from solutions are available, but the data generating process is not modeled.

Combining PROMETHEE and BNs overcomes major limitations of outranking-type MCDM approaches and BNs in modeling decisions under uncertainty. Traditional outranking approaches such as PROMETHEE provide limited or no support when criteria values involve uncertainty. BNs can model complex probability distributions in a concise way, but are not designed for decision making problems. Extensions of BNs for decision making, such as IDs, require elicitation of complex utility tables in decision problems with multiple criteria. The proposed approach overcomes these limitations by having the ability to model complex criteria distributions in a BN model, and computing outranking relations from these distributions based on DM preferences using PROMETHEE.

Limitations of our approach include dependence on the criteria weights obtained from the DMs. We use sensitivity analysis to assess the robustness of the results to changes in the weights. As future work, indirect and interactive elicitation approaches can be implemented to provide more robust weight elicitation. Implementation of the proposed approach to BN software and packages can enable a wider use of BNs for MCDM problems.

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK). This study was supported by an Institutional Links Grant (ID 352394702) under the Turkey-Newton-Katip Çelebi Fund partnership. The Grant is funded by the UK Department of Business, Energy and Industrial Strategy (BEIS) and The Scientific and Technological Research Council of Turkey (TUBITAK, ID 217M518) and delivered by the British Council.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Ahn, B. S. (2011). Compatible weighting method with rank order centroid: Maximum entropy ordered weighted averaging approach. *European Journal of Operational Research, 212*(3), 552–559. https://doi.org/10.1016/j.ejor.2011.02.017

Baležentis, T., & Streimikiene, D. (2017). Multi-criteria ranking of energy generation scenarios with Monte Carlo simulation. *Applied Energy, 185*, 862–871. https://doi.org/10.1016/j.apenergy.2016.10.085

Barton, D. N., Sundt, H., Bustos, A. A., Fjeldstad, H. P., Hedger, R., Forseth, T., Köhler, B., Aas, Ø., Alfredsen, K., & Madsen, A. L. (2020). Multi-criteria decision analysis in Bayesian networks—Diagnosing ecosystem service trade-offs in a hydropower regulated river. *Environmental Modelling and Software, 124,* 104604. https://doi.org/10.1016/j.envsoft.2019.104604

Baudry, G., Macharis, C., & Vallée, T. (2018). Range-based Multi-Actor Multi-Criteria Analysis: A combined method of Multi-Actor Multi-Criteria Analysis and Monte Carlo simulation to support participatory decision making under uncertainty. *European Journal of Operational Research, 264*(1), 257–269. https://doi.org/10.1016/j.ejor.2017.06.036

Betrie, G. D., Sadiq, R., Morin, K. A., & Tesfamariam, S. (2013). Selection of remedial alternatives for mine sites: A multicriteria decision analysis approach. *Journal of Environmental Management, 119*, 36–46. https://doi.org/10.1016/j.jenvman.2013.01.024

Brans, J. P., & Mareschal, B. (1994). The PROMCALC & GAIA decision support system for multicriteria decision aid. *Decision Support Systems, 12*(4–5), 297–310. https://doi.org/10.1016/0167-9236(94)90048-5

Brans, J. P., Vincke, Ph., & Mareschal, B. (1986). How to select and how to rank projects: The PROMETHEE method. *European Journal of Operational Research, 24*, 228–238. https://doi.org/10.1016/0377-2217(86)90044-5

Broekhuizen, H., Groothuis-Oudshoorn, C. G., Oudshoorn, G., Til, J., Hummel, J. M., & Ijzerman, M. J. (2015). A review and classification of approaches for dealing with uncertainty in multi-criteria decision

analysis for healthcare decisions. *PharmacoEconomics, 33*(5), 445–455. https://doi.org/10.1007/s40273-014-0251-x
Chai, J., & Ngai, E. W. (2020). Decision-making techniques in supplier selection: Recent accomplishments and what lies ahead. *Expert Systems with Applications, 140*, 112903. https://doi.org/10.1016/j.eswa.2019.112903
Cinelli, M., Kadziński, M., Gonzalez, M., & Słowiński, R. (2020). How to support the application of multiple criteria decision analysis? Let us start with a comprehensive taxonomy. *Omega (United Kingdom), 96*, 102261. https://doi.org/10.1016/j.omega.2020.102261
Corrente, S., Figueira, J. R., & Greco, S. (2014). The SMAA-PROMETHEE method. *European Journal of Operational Research, 239*(2), 514–522. https://doi.org/10.1016/j.ejor.2014.05.026
Cui, L., Wu, H., & Dai, J. (2023). Modelling flexible decisions about sustainable supplier selection in multitier sustainable supply chain management. *International Journal of Production Research, 61*(14), 4603–4624. https://doi.org/10.1080/00207543.2021.1924412
Delcroix, V., Sedki, K., & Lepoutre, F. X. (2013). A Bayesian network for recurrent multi-criteria and multi-attribute decision problems: Choosing a manual wheelchair. *Expert Systems with Applications, 40*(7), 2541–2551. https://doi.org/10.1016/j.eswa.2012.10.065
Dorini, G., Kapelan, Z., & Azapagic, A. (2011). Managing uncertainty in multiple-criteria decision making related to sustainability assessment. *Clean Technologies and Environmental Policy, 13*(1), 133–139. https://doi.org/10.1007/s10098-010-0291-7
Dohale, V., Gunasekaran, A., Akarte, M., & Verma, P. (2021). An integrated Delphi-MCDM-Bayesian Network framework for production system selection. *International Journal of Production Economics, 242,* 108296. https://doi.org/10.1016/j.ijpe.2021.108296
Doumpos, M., & Zopounidis, C. (2010). A multicriteria decision support system for bank rating. *Decision Support Systems, 50*(1), 55–63. https://doi.org/10.1016/j.dss.2010.07.002
Durbach, I. N., & Stewart, T. J. (2012). Modeling uncertainty in multi-criteria decision analysis. *European Journal of Operational Research, 223*(1), 1–14. https://doi.org/10.1016/j.ejor.2012.04.038
Fan, S., Zhang, J., Blanco-Davis, E., Yang, Z., & Yan, X. (2020). Maritime accident prevention strategy formulation from a human factor perspective using Bayesian Networks and TOPSIS. *Ocean Engineering, 210,* 107544. https://doi.org/10.1016/j.oceaneng.2020.107544
Fan, H., Lu, J., Chang, Z., & Ji, Y. (2023). A Bayesian network-based TOPSIS framework to dynamically control the risk of maritime piracy. *Maritime Policy & Management*. https://doi.org/10.1080/03088839.2023.2193585
Fenton, N., & Neil, M. (2001). Making decisions: Using Bayesian nets and MCDA. *Knowledge-Based Systems, 14*(7), 307–325. https://doi.org/10.1016/S0950-7051(00)00071-X
Gervásio, H., & Simões Da Silva, L. (2012). A probabilistic decision-making approach for the sustainable assessment of infrastructures. *Expert Systems with Applications, 39*(8), 7121–7131. https://doi.org/10.1016/j.eswa.2012.01.032
Govindan, K., Rajendran, S., Sarkis, J., & Murugesan, P. (2015). Multi criteria decision making approaches for green supplier evaluation and selection: A literature review. *Journal of Cleaner Production, 98*, 66–83. https://doi.org/10.1016/j.jclepro.2013.06.046
Greco, S., Ehrgott, M., & Figueira, J. R. (2016). *Multiple criteria decision analysis state of the art surveys*. Springer.
Howard, R. A., & Matheson, J. E. (2005). Influence diagrams. *Decision Analysis*. https://doi.org/10.1287/deca.1050.0020
Hubinont, J.-P. (2016). SMAA-GAIA: A complementary tool of the SMAA-PROMETHEE method. *International Journal of Multicriteria Decision Making, 6*(3), 237–246. https://doi.org/10.1504/IJMCDM.2016.079714
Hyde, K., Maier, H. R., & Colby, C. (2003). Incorporating uncertainty in the PROMETHEE MCDA method. *Journal of Multi-Criteria Decision Analysis, 12*(4–5), 245–259. https://doi.org/10.1002/mcda.361
Jensen, F., Jensen, F. V., & Dittmer, S. L. (1994). From Influence diagrams to junction trees. In *Uncertainty proceedings 1994*. https://doi.org/10.1016/b978-1-55860-332-5.50051-1
Kaya, R., Salhi, S., & Spiegler, V. (2023). A novel integration of MCDM methods and Bayesian networks: The case of incomplete expert knowledge. *Annals of Operations Research, 320*(1), 205–234. https://doi.org/10.1007/s10479-022-04996-7
Kaya, R., & Yet, B. (2019). Building Bayesian networks based on DEMATEL for multiple criteria decision problems: A supplier selection case study. *Expert Systems with Applications, 134*, 234–248. https://doi.org/10.1016/j.eswa.2019.05.053
Kilic, H. S., Zaim, S., & Delen, D. (2014). Development of a hybrid methodology for ERP system selection: The case of Turkish Airlines. *Decision Support Systems, 66*, 82–92. https://doi.org/10.1016/j.dss.2014.06.011

Kuang, H., Kilgour, D. M., & Hipel, K. W. (2015). Grey-based PROMETHEE II with application to evaluation of source water protection strategies. *Information Sciences, 294*, 376–389. https://doi.org/10.1016/j.ins.2014.09.035

Lahdelma, R., Hokkanen, J., & Salminen, P. (1998). SMAA - Stochastic multiobjective acceptability analysis. *European Journal of Operational Research, 106*(1), 137–143. https://doi.org/10.1016/S0377-2217(97)00163-X

Lahdelma, R., & Salminen, P. (2001). SMAA-2: Stochastic multicriteria acceptability analysis for group decision making. *Operations Research, 49*(3), 444–454. https://doi.org/10.1287/opre.49.3.444.11220

Lauritzen, S. L., & Spiegelhalter, D. J. (1988). Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society: Series B (Methodological), 50*(2), 157–194. https://doi.org/10.1111/j.2517-6161.1988.tb01721.x

Liao, Z., Liao, H., & Lev, B. (2022). Compromise solutions for stochastic multicriteria acceptability analysis with uncertain preferences and nonmonotonic criteria. *International Transactions in Operational Research, 29*(6), 3737–3757. https://doi.org/10.1111/itor.13078

Liu, Y., Fan, Z. P., & Zhang, Y. (2011). A method for stochastic multiple criteria decision making based on dominance degrees. *Information Sciences, 181*, 4139–4153. https://doi.org/10.1016/j.ins.2011.05.013

Maghrabie, H. F., Beauregard, Y., & Schiffauerova, A. (2019). Multi-criteria decision making problems with unknown weight information under uncertain evaluations. *Computers and Industrial Engineering, 133*(April), 131–138. https://doi.org/10.1016/j.cie.2019.05.003

Mareschal, B. (1986). Stochastic multicriteria decision making and uncertainty. *European Journal of Operational Research, 26*(1), 58–64. https://doi.org/10.1016/0377-2217(86)90159-1

Mareschal, B. (1988). Weight stability intervals in multicriteria decision aid. *European Journal of Operational Research, 33*, 54–64. https://doi.org/10.1016/0377-2217(88)90254-8

Mazaheri, A., Rafiee, N. & Khadivi, P. (2010). Location based targeted advertising using Bayesian Network and fuzzy TOPSIS. In *5th International symposium on telecommunications* (pp. 645–650). https://doi.org/10.1109/ISTEL.2010.5734103

Montazar, A., Gheidari, O. N., & Snyder, R. L. (2013). A fuzzy analytical hierarchy methodology for the performance assessment of irrigation projects. *Agricultural Water Management, 121*, 113–123. https://doi.org/10.1016/j.agwat.2013.01.011

Neil, M., Tailor, M., & Marquez, D. (2007). Inference in hybrid Bayesian Networks using Dynamic discretisation. *Statistics and Computing, 17*(3), 219–234. https://doi.org/10.1007/s11222-007-9018-y

Özkan, I., Burhan, I., & Türksen, T. (2014). Uncertainty and fuzzy decisions. https://doi.org/10.1007/978-94-017-8691-1_2

Pelissari, R., Oliveira, M. C., Abackerli, A. J., Ben-Amor, S., & Assumpção, M. R. P. (2021). Techniques to model uncertain input data of multi-criteria decision-making problems: A literature review. *International Transactions in Operational Research, 28*(2), 523–559. https://doi.org/10.1111/itor.12598

Pelissari, R., Oliveira, M. C., Amor, S., Kandakoglu, A., & Helleno, A. L. (2020). SMAA methods and their applications: A literature review and future research directions. *Annals of Operations Research, 293*(2), 433–493. https://doi.org/10.1007/s10479-019-03151-z

Pitchipoo, P., Venkumar, P., & Rajakarunakaran, S. (2013). Fuzzy hybrid decision model for supplier evaluation and selection. *International Journal of Production Research, 51*, 3903–3919. https://doi.org/10.1080/00207543.2012.756592

Rashidi, K., Noorizadeh, A., Kannan, D., & Cullinane, K. (2020). Applying the triple bottom line in sustainable supplier selection: A meta-review of the state-of-the-art. *Journal of Cleaner Production, 269*, 122001. https://doi.org/10.1016/j.jclepro.2020.122001

Quan, J. C., & Cho, S. B. (2014). A hybrid recommender system based on AHP that awares contexts with Bayesian Networks for Smart TV. *International Conference on Hybrid Artificial Intelligence Systems, 8480*, 527–536. https://doi.org/10.1007/978-3-319-07617-1_46

Roszkowska, E. (2013). Rank ordering criteria weighting methods—A comparative overview. *Optimum Studia Ekonomiczne, 5*(65), 14–33. https://doi.org/10.15290/ose.2013.05.65.02

Saputro, T. E., Figueira, G., & Almada-Lobo, B. (2022). A comprehensive framework and literature review of supplier selection under different purchasing strategies. *Computers & Industrial Engineering, 167*, 108010. https://doi.org/10.1016/j.cie.2022.108010

Shahzad, A. (2022). Multi-criteria decision making using Bayesian Networks. Doctor of Philosophy in Information Technology.

Shakhsi-Niaei, M., Torabi, S. A., & Iranmanesh, S. H. (2011). A comprehensive framework for project selection problem under uncertainty and real-world constraints. *Computers and Industrial Engineering, 61*(1), 226–237. https://doi.org/10.1016/j.cie.2011.03.015

Tervonen, T., & Figueira, J. R. (2008). A survey on stochastic multicriteria acceptability analysis methods. *Journal of Multi-Criteria Decision Analysis, 15*(1–2), 1–14. https://doi.org/10.1002/mcda.407

Venkatesh, V. G., Zhang, A., Deakins, E., Luthra, S., & Mangla, S. (2019). A fuzzy AHP-TOPSIS approach to supply partner selection in continuous aid humanitarian supply chains. *Annals of Operations Research, 283*, 1517–1550. https://doi.org/10.1007/s10479-018-2981-1

Watthayu, W., & Peng, Y. (2004). A Bayesıan Network based framework for multi-ciıteria decision making. In *17th International conference on multiple criteria decision analysis*.

Yang, Z., Wan, C., Yang, Z., & Yu, Q. (2021). Using Bayesian network-based TOPSIS to aid dynamic port state control detention risk control decision. *Reliability Engineering & System Safety, 213*, 107784. https://doi.org/10.1016/j.ress.2021.107784

Yuen, K. K. F., & Ting, T. O. (2012). Textbook selection using fuzzy PROMETHEE II method. *International Journal of Future Computer and Communication, 1*(1), 76–78. https://doi.org/10.7763/ijfcc.2012.v1.20

Zhang, N. L. (1998). Probabilistic inference in influence diagrams. In *UAI'98: Proceedings of the fourteenth conference on uncertainty in artificial intelligence* (pp. 514–522).

Zhang, Y., Fan, Z. P., & Liu, Y. (2010). A method based on stochastic dominance degrees for stochastic multiple criteria decision making. *Computers and Industrial Engineering, 58*, 544–552. https://doi.org/10.1016/j.cie.2009.12.001

Zimmer, K., Fröhling, M., & Schultmann, F. (2016). Sustainable supplier management—A review of models supporting sustainable supplier selection, monitoring and development. *International Journal of Production Research, 54*(5), 1412–1442. https://doi.org/10.1080/00207543.2015.1079340