

Portraying probabilistic relationships of continuous nodes in Bayesian networks with ranked nodes method

Pekka Laitila^{a,*}, Kai Virtanen^{a,b}

^a Department of Mathematics and Systems Analysis, School of Science, Aalto University, Helsinki, Finland

^b Department of Military Technology, Finnish National Defence University, Helsinki, Finland

ARTICLE INFO

Keywords:

Bayesian networks
Ranked nodes
Probability elicitation
Continuous node discretization
Conditional probability tables

ABSTRACT

This paper advances the use of the ranked nodes method (RNM) to portray probabilistic relationships of continuous quantities in Bayesian networks (BNs). In RNM, continuous quantities are represented by ranked nodes with discrete ordinal scales. The probabilistic relationships of the nodes are quantified in conditional probability tables (CPTs) generated with expert-elicited parameters. When ranked nodes are formed by discretizing continuous scales, ignorance about the functioning of RNM can lead to discretizations that make the generation of sensible CPTs impossible. While a guideline exists on this matter, it is limited by a requirement to define an equal number of ordinal states for all the nodes. This paper presents two novel discretization approaches that consider the functioning of RNM and allow the nodes to have non-equal numbers of ordinal states. In the first one, called the “static discretization approach”, the nodes can be given any desired discretizations that stay unchanged during the use of the BN. In the second one, called the “dynamic discretization approach”, the discretizations are algorithmically updated during the use of the BN to help manage the sizes of the generated CPTs. Both approaches are based on the original idea that, besides the RNM parameters, the nodes probabilistic relationship is defined by initial RNM-compatible discretizations elicited from the domain expert. Overall, the new approaches offer an easier and more versatile way of using RNM to depict the probabilistic relationships of continuous quantities. In doing so, they also facilitate the effective and diverse use of BNs in decision support systems.

1. Introduction

Numerous decision support systems utilize a Bayesian network (BN) or an influence diagram to represent uncertain knowledge and aid decision-making under uncertainty. Their application areas include, e. g., medical decision-making [1–5], risk and safety management [6–10], project management [11,12], maintenance and policy planning [13–15], financial forecasting [16], and military planning [17,18]. A BN [19] depicts probabilistic relationships between random variables both visually and numerically. The visual side of the BN is a directed acyclic graph in which nodes portray the random variables and arcs indicate their direct dependencies. The numerical side quantifies the probabilistic relationships, which are described with conditional probability tables (CPTs) of discrete nodes and indicated visually by the arcs. A CPT defines the probability distributions of a descendant, the child node, for all combinations of the states of its direct predecessors, the parent nodes. In an overall view, the BN encodes the joint probability distribution of

all the nodes in the network. When evidence is entered into the BN, i.e., certain states of selected nodes are given 100% probability, the probability distributions of the other nodes can readily be updated with effective algorithms; see [20,21]. Through this feature, known as probabilistic inference, the BN provides a means to answer probabilistic queries about the random variables. The same feature is utilized also in influence diagrams, which are decision-theory extensions of BNs [22]. In these models, nodes representing random variables are joined with nodes representing the objectives and possible actions of a decision maker. While the contributions of this paper are applicable to influence diagrams, the remaining discussion refers solely to BNs for the sake of simplicity.

BNs can be constructed on the basis of various information sources such as experimental data, historical data, and expert knowledge. If a comprehensive data collection is available, both the visual and the numerical sides of a BN can be constructed by data-fitting approaches [23]. However, in many practical applications, the data available may be too

* Corresponding author.

E-mail addresses: pekka.laitila@aalto.fi (P. Laitila), kai.virtanen@aalto.fi (K. Virtanen).

scarce or unsuitable for the needs of the BN construction [24,25]. For example, a specific fragment of a BN may consist of nodes whose probabilistic relationships have to be defined by expert elicitation. When the nodes represent continuous quantities, the ranked nodes method (RNM) [25] is a popular means of doing this. The method represents continuous quantities through ranked nodes with discrete ordinal scales and CPTs. The scales may consist of descriptive labeled states, e.g., {low, medium, high}, or they may be constructed by discretizing continuous scales, e.g., {[0 km, 2 km], [2 km, 5 km], [5 km, 10 km]}. Since its introduction, RNM has been used in several BN models for decision support. The latest applications include the risk management of epidemics [26], early weed invasions [27], supplier selection in the automobile industry [28], and improvement of teamwork quality in software development [29]. In recent years, research on the methodological properties of RNM has also started to emerge. One study elaborates the theoretical principle of RNM and investigates how well CPTs in real-life BN models can be reproduced with RNM [30]. Another study explores how well CPTs constructed with RNM are able to portray probabilistic relationships typical in human reliability analysis [31]. There also exists a study on the capability of RNM to represent the explaining away property of binary variables [32]. In addition, expert elicitation practices for RNM have also been established [33,34]. This paper further elaborates RNM by first discussing challenges concerning its application to nodes with discretized continuous scales. The paper then resolves those challenges by presenting two novel discretization approaches to be used with RNM. The approaches allow the continuous scales of the nodes to be discretized in ways that are not possible with current RNM practices. They provide a more flexible means than before of eliciting the parameters used in RNM from a domain expert. In addition, the approaches enable the probabilistic relationship of the nodes to be portrayed in a more versatile manner than the existing practices. This improvement broadens the scope of the probabilistic analyses that can be performed for the nodes with the BN.

RNM enables the construction of CPTs consisting of dozens or even hundreds of elements just with a handful of parameters elicited from a domain expert. First, the expert selects a generic rule, called the weight expression, according to which the parent nodes affect the child node. Then, the expert selects weight and variance parameters that define in more detail the relationship of the nodes within the frames of the weight expression. The weights reflect the parents' relative strengths of influence on the child node. The variance parameter describes the dispersion level of the child node's conditional probability distributions. Given these parameters, the CPT can be generated for further verification. If necessary, different parts of the CPT can also be generated using different weight expressions or values of the weights and the variance parameter.

When ranked nodes are formed by discretizing continuous scales, there are various sources of complication regarding the use of RNM. First, if the discretizations are formed in ignorance of the functioning of RNM, all the weight expressions of RNM may be unable to portray the probabilistic relationship between the nodes [33]. To deal with this problem, a guideline for the construction of RNM-compatible discretizations is provided in [33]. RNM-compatible discretization means discretizing the continuous scales of the nodes into equal numbers of ordinal states in a coordinated manner that takes into account the functioning of RNM. Defining RNM-compatible discretizations for the nodes supports the discovery of suitable RNM parameters as well as the construction of sensible CPTs. However, the requirement to use the same number of states for all the nodes can be undesirable in some applications. Furthermore, the elicitation effort required from the expert to construct RNM-compatible discretizations increases with the number of discrete ordinal states to be defined for the nodes. Therefore, the number of states of the nodes is likely to stay small. These features limit the ease and level of detail by which the probabilistic relationship of the nodes can be represented and explored. Another challenge concerns a property of RNM: CPTs generated using fixed RNM parameters, but with

alternative discretizations, generally contradict each other with regard to their probabilistic implications about the nodes. Therefore, if a CPT constructed with RNM is to be regenerated with alternative discretizations, the RNM parameters have to be elicited again. This property may confuse or be ignored by users of RNM. The additional elicitation effort also complicates the representation and exploration of the probabilistic relationship of the nodes.

To resolve the above challenges, this paper presents two new discretization approaches concerning the application of RNM to continuous nodes. Both are based on the original idea that the probabilistic relationship between a child node and its parents is completely encoded by the RNM-compatible discretizations and the RNM parameters elicited from the domain expert. In both approaches, this encoding principle is utilized to allow the nodes to have alternative discretizations, which otherwise would lead to the aforementioned problems under the existing RNM practices.

The first new approach is called the "static discretization approach", where "static" means that the discretizations of the nodes remain unaltered during the use of the BN. In this approach, initial RNM-compatible discretizations are first elicited from the expert with the existing discretization guideline [33]. Initial here refers to the fact that the nodes can be rediscritized at later steps of the approach. Yet, at this point, the RNM-compatible discretization leads to the nodes obtaining an equal number of ordinal states that partly define their probabilistic relationship. Next, the RNM parameters are elicited from the expert. To facilitate the elicitation, the expert can freely rediscritize the nodes. That is, the expert can redivide the continuous scales of the nodes into arbitrary and non-equal numbers of discretization intervals based on, e.g., the expert's natural tendency to comprehend and describe the probabilistic relationship of the nodes. Once the RNM parameters are elicited, one can generate a CPT and use the BN with the discretizations selected by the expert for the parameter elicitation. However, it is also possible to assign for the nodes any other discretizations that may better serve the analyses to be carried out with the BN. No matter what discretizations are selected, the CPT generated with them represents a probabilistic relationship that is encoded by the initial RNM-compatible discretizations and the RNM parameters. The option to freely rediscritize the nodes both before and after the elicitation of RNM parameters is possible because of this encoding principle. The rediscrization ability separates the static discretization approach from the present RNM practices. Like the existing practice [33], the static discretization approach takes into account the functioning of RNM and the related need for RNM-compatible discretizations. However, the new approach manages to do so without forcing one to operate only with a small and equal number of states for all the nodes. Furthermore, unlike with RNM at present, rediscrizing the nodes after the parameter elicitation does not lead to probabilistic inconsistencies between the original and the regenerated CPTs. Therefore, compared to the present practices, the static discretization approach eases the elicitation of RNM parameters from the expert and broadens the way RNM can be used to represent and explore probabilistic relationships between continuous quantities.

In principle, the static discretization approach enables the established probabilistic relationship of a child node and its parents to be represented through arbitrary dense discretizations. However, with high enough discretization densities or numbers of discretized nodes in the BN, the sizes of the generated CPTs may cause problems concerning the computer memory requirements for their storage or the conducting of probabilistic inference in the BN [19,35]. To mitigate this problem, the paper presents another new discretization approach, which combines the application of RNM and a dynamic discretization algorithm of Neil et al. [35]. In this "dynamic discretization approach", the discretizations of the ranked nodes are not static but updated with the dynamic discretization algorithm whenever new evidence is entered into the BN during its use. The dynamic approach can be applied once the initial RNM-compatible discretizations and the RNM parameters have been elicited from the expert by means of the static approach. The

discretizations formed by the algorithm are not uniform, but denser in those areas of the nodes' continuous scales in which the probability mass is more concentrated under the entered evidence. Therefore, if the use of the static approach causes a need to generate impractically large CPTs, the dynamic approach can provide a better means to depict the probabilistic relationship of the nodes with the desired level of detail. For instance, the dynamic approach readily allows entering point-valued evidence into the nodes. Moreover, it enables accurate statistics on the nodes' probability distributions, e.g., when the conditional mean values cover a wide scale and the conditional variances are small. The dynamic updating of the discretizations is possible because the nodes' probabilistic relationship is encoded by the initial RNM-compatible discretizations and the RNM parameters. This is in contrast to the existing practices, which do not enable combining the use of RNM with the dynamic discretization algorithm.

The paper is organized as follows. Section 2 provides an overview and comparison of existing methods that are complementary to RNM. Section 3 briefly explains the functioning of RNM and gives short descriptions of the guidelines [33] for applying RNM to nodes with continuous scales. Section 4 presents the motivation, underlying principle and application of the static discretization approach. Section 5 contains a corresponding presentation on the dynamic discretization approach. Section 6 provides concluding remarks. Sections 3–5, use an example BN to demonstrate the discussed matters.

2. Overview of parametric methods for constructing conditional probability tables

The size of a CPT grows exponentially with the number of parent nodes. Therefore, if a child node with three parent nodes each have five states, the CPT of the child node already consists of 625 elements. Assessing dozens or hundreds of conditional probabilities of a CPT is often impossible for a domain expert, due to cognitive strain or lack of time. To mitigate this problem, several parametric methods have been developed to ease the construction of CPTs by expert elicitation. These methods allow constructing a CPT through expert-assessed parameters whose number is significantly smaller than the number of elements in the CPT. In the literature, these methods have also been referred to as parametric probability distributions [23], canonical models [36], canonical distributions [21], and filling-up methods [31,24]. In the following, existing parametric methods are discussed and compared to RNM.

In the noisy-OR [36] and noisy-MAX [37,38] methods, the basic idea is that the parent nodes are individual causes for a common effect represented by the child node. In turn, the parameters elicited from the expert are CPT entries that indicate the abilities of the causes to bring about the effect individually. The remaining elements of the CPT are calculated with the assumption that, in the presence of several causes, each cause affects the child node independently of the others. Noisy-OR can handle only binary nodes whereas noisy-MAX is applicable to nodes with non-binary ordinal scales. According to [25], RNM enables a greater range of probabilistic relationships to be portrayed than does noisy-MAX. In addition, RNM has been shown to allow representing the explaining away property of binary variables more extensively than noisy-OR [32].

The EBBN method (Elicitation for Bayesian Belief Networks) [39], the weighted sum algorithm [40] and the Cain calculator [41] utilize interpolation of conditional probability distributions. In these methods, the expert first assesses the probability distributions of the child node for the so-called anchor combinations of states of the parent nodes. The other conditional probability distributions of the CPT are then derived by interpolating between the anchor distributions. The anchor state combinations and interpolation techniques used are specific to each method. Similarly to RNM, each of the methods involves the parent nodes receiving weights that reflect their strength of influence on the child node. On the other hand, unlike in RNM, the dispersion of the

derived distributions is not user-controlled, but reflects those of the anchor distributions.

The functional interpolation method [42] and the InterBeta method [43] also utilize the principle of interpolation to derive missing probability distributions of a CPT from method-specific anchor distributions assessed by the expert. However, in these methods, the interpolation does not directly focus on the probabilities of the anchor distributions. In the functional interpolation method, a normal distribution is fitted to each anchor distribution so that best-fit estimates of the mean and variance parameters are determined. The missing probability distributions of the CPT are formed through normal distributions whose mean and variance parameters are interpolated from those of the anchor distribution estimates. In the InterBeta method, the principle is similar except that Beta distributions are utilized instead of normal distributions. Furthermore, in the InterBeta method, the expert may assign weights to parent nodes, their states, or their state combinations. Increasing the weighting detail increases the elicitation effort of the expert. On the other hand, it enables portraying the probabilistic relationship of the nodes more accurately. In that regard, the use of the alternative weighting options of InterBeta is similar to the use of partitioned weight expressions in RNM.

A method presented by Røed et al. [44] is similar to RNM in the sense that the construction of a CPT is based on a functional relationship between the parents and the child node. Moreover, like in RNM, the parents get weights reflecting their strengths of influence on the child, and a single parameter defines the dispersion of the probability distributions. However, whereas RNM provides four basic weight expressions to describe the probabilistic relationship of the nodes, the method of Røed et al. uses only one function. This function is similar to a weight expression of RNM called WMEAN, in which weighted averages are taken of the states of the parent nodes. Also, in a method suggested by Hassall et al. [45], the conditional probability distributions of the child node are calculated utilizing weighted averages of the parent states. However, unlike RNM, this method does not involve the expert evaluating the dispersion of the distributions. Furthermore, for a child node with an odd number of states m , the middle state obtains the probability $1/m$ for any combination of the parent states. Therefore, the CPTs generated with this method are likely to require more manual editing than CPTs generated with RNM.

In the likelihood method [46], the idea is that in the absence of any information about the parent nodes, the child node has a so-called typical distribution that has been assessed by the expert. Different state combinations of the parent nodes then tend to move the probability distribution of the child node away from the typical distribution in systematic ways. The CPT is constructed by multiplying the typical distribution by likelihood terms consisting of weighting factors that the expert has selected for the states of the child node and the parent nodes. The presentation of the method in [46] does not include any detailed guideline for the elicitation of the weighting factors. Some instruction is provided in [47] along with a remark that the method becomes very complex if the child node has more than three states. On the other hand, with RNM, exact guidelines for the weight elicitation exist [33,34], and the number of parameters to be elicited does not increase with the number of states of the nodes.

Chin et al. [48] utilize the Analytic Hierarchy Process (AHP) for the construction of a CPT. In their method, the expert performs pairwise comparisons of the probabilities of the states of the child node given the state of an individual parent node. The pairwise comparisons enable the calculation of probability distributions of the child node conditioned to single parent nodes. Then, the final distributions of the CPT are formed by taking products of the probability distributions conditional to single parents. Contrary to this method, the elicitation guidelines for RNM ask the expert to assess the mode of the child node for specified state combinations of all the parent nodes. Thereby, the elicitation for RNM allows the expert to consider the joint effect of the parent nodes on the child node in a clear way.

To get an idea of the elicitation effort of the methods discussed above, Table 1 presents the numbers of quantitative expert assessments that each of them requires when a child node has n parent nodes and all the nodes have m states. The table presents formulas that apply for any values of n and m as well as the numerical values that apply for the case $n = 3$ and $m = 5$. In this specific case of $n = 3$ and $m = 5$, the CPT of the child node consists of $m^{n+1} = 625$ elements and its direct assessment would require $m^n(m - 1) = 500$ probabilities to be specified by an expert. Compared to this number, all the methods significantly reduce the number of quantitative assessments required from the expert. It should be noted that the Cain calculator does not actually provide a computational routine for the construction of the CPT when the child has more than three states [41]. Furthermore, regarding RNM and the InterBeta method, the numbers in Table 1 correspond to the ways they are used by default. As discussed above, both methods provide the option of specifying more parameters, thereby enabling the construction of CPTs describing a greater range of probabilistic relationships.

As a further consideration, the method of Hassall et al., the InterBeta method, the weighted sum algorithm, and the likelihood method all require the expert to set weights of importance for the parent nodes or their states without providing any exact instructions. This shortcoming may complicate the elicitation procedure with these methods. On the other hand, exact guidelines of weight elicitation have been established for RNM concerning ranked nodes formed through discretized continuous scales [33] and labeled scales [34]. The approach presented in [33] enables the determination of the weight expression and the weights of n parents with the expert estimating the mode of the child node on its continuous scale in $2n$ scenarios. In turn, the elicitation framework presented in [34] allows determining a feasible weight expression and a set of feasible weights once the expert has assessed the two most probable states of the child node for $2n$ parent state combinations. Based on these considerations and the results in Table 1, RNM requires the least amount of elicitation effort from the expert for constructing CPTs.

Besides the small number of parameters to be elicited, an advantage of RNM is that the alternative weight expressions help experts to understand and describe the probabilistic relationship between a child node and its parent nodes [25]. Furthermore, the easy deployment of RNM is supported by an implementation of the method in AgenaRisk software [49]. Of the other methods discussed, only noisy-OR and noisy-MAX are implemented in existing well-known BN software, such as GeNIe [50], Netica [51], and Hugin [52]. The likelihood method and the method of Hassall et al. have implementations available online [53, 54]. However, these implementations are not linked to a wide range of functionalities of BN analysis, unlike the aforementioned software.

Table 1
Numbers of quantitative expert assessments required in different parametric methods when a child node has n parent nodes and all the nodes have m states.

Method	Number of assessments	Number of assessments when $n = 3$ and $m = 5$
Hassall et al. [45]	n	3
RNM [25]a	$n + 1$	4
InterBeta [43]a	$2(m - 1) + n$	11
Røed et al. [44]	$(m - 1)n$	12
Weighted sum algorithm [40]	$m^2 - m + n$	23
EBBN [39]	$m^2 - m + 2n$	26
Likelihood [46]	$(n + 2)m + 1$	26
Functional interpolation method [42]	$2^n(m - 1)$	32
Noisy-MAX [37,38]	$n(m - 1)^2$	48
Cain calculator [41]b	$n(m - 1)^2$	48
Chin et al. [48]	$n(m^2 - m)$	60

^a The numbers of assessments correspond to default forms of use of the methods.

^b The method does not provide a computational routine for the construction of the CPT when the child node has more than three states, i.e., $m > 3$.

To summarize, the methodological principle of RNM is complementary to those of other parametric methods discussed. The use of weight expressions, which is unique to RNM, provides both flexibility and cognitive support for the expert in describing the probabilistic relationship of the nodes. Furthermore, the small number of parameters to be elicited, the related elicitation guidelines, and the existing software implementation are qualities that promote and support effective use of RNM in applications.

3. Ranked nodes method (RNM)

This section outlines the technical principle of RNM as well as two guidelines concerning its application to nodes with continuous scales. The topics are covered here at a level of detail that is necessary in order to understand the contributions of this paper. More thorough descriptions are found in [25,33,30].

3.1. Functioning of RNM

The example BN presented in Fig. 1 consists of ranked nodes formed by discretizing continuous scales. The BN describes how the price of a machine and its weekly amount of use determine the time it takes before the machine requires thorough maintenance. The quantities are represented by the nodes *Price*, *Weekly Usage*, and *Service Time*. Suppose that an increasing price increases the service time of the machine, while an increasing amount of weekly usage shortens it. This type of monotonic direction of influence of the parent nodes on the child node is an underlying assumption in RNM.

The basic idea in the generation of a CPT is that for any combination of states of the parent nodes, the most probable state of the child node is defined by a general rule. Within the framework of this rule, the parent nodes can have non-equal strengths of influence on the child node. The rule is called a weight expression and it is selected by the expert. The strengths of the parent nodes are expressed through weights that are also elicited from the expert. In addition, the expert also assigns a variance parameter, which describes how dispersed around the mode the probability distribution of the child node is for given states of the parent nodes.

The generation of the CPT of the child node is based on associating the states of the nodes with consecutive subintervals of the unit scale [0, 1]. The subintervals, called state intervals, are of equal width and indicate the direction of influence of the parent nodes on the child node. The state intervals in Fig. 1 indicate the monotonic influence directions of *Price* and *Weekly Usage* on *Service Time*.

By utilizing the state intervals, CPTs are generated according to the

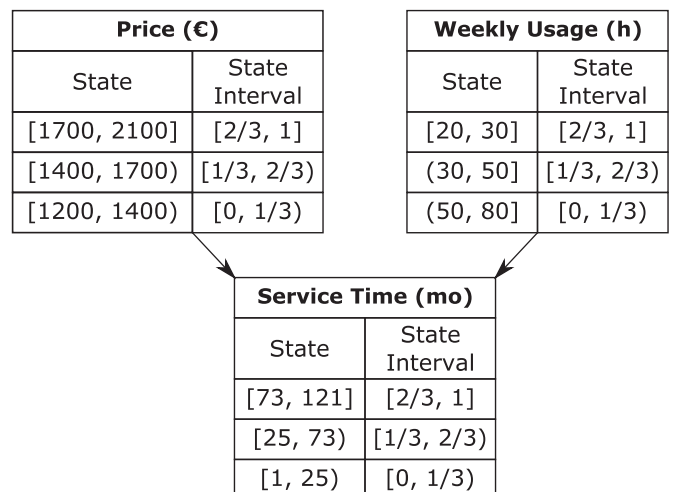


Fig. 1. Example BN.

following principle. Let there be discrete parent nodes X_1, \dots, X_n and a child node X_C that are ranked nodes. Furthermore, let there be continuous random variables χ_1, \dots, χ_n defined on the unit scale $[0, 1]$ and a random variable χ_C that depends on them according to a regression model

$$\chi_C = f(\chi_1, \dots, \chi_n, \mathbf{w}) + e, \quad e \sim N(0, \sigma^2). \tag{1}$$

The regression function $f(\cdot)$ and the regression coefficients \mathbf{w} are the weight expression and the weights elicited from the expert. The variance σ^2 of the normally distributed error term e is the variance parameter. With x_i and $[\alpha_i, \beta_i]$ denoting a given state and the associated state interval of X_i , a CPT element $P(X_C = x_C | X_1 = x_1, \dots, X_n = x_n)$ is calculated in RNM on the basis of Eq. (1) by

$$\begin{aligned} P(X_C = x_C | X_1 = x_1, \dots, X_n = x_n) \\ = P(\chi_C \in [\alpha_C, \beta_C] | \chi_1 \in [\alpha_1, \beta_1], \dots, \chi_n \in [\alpha_n, \beta_n], \chi_C \in [0, 1]). \end{aligned} \tag{2}$$

That is, knowing X_i to be in the state x_i on the ordinal scale is equivalent to knowing χ_i to lie within the state interval $[\alpha_i, \beta_i]$ on the unit scale. Based on this analogy, X_i and χ_i can both be seen to represent the same continuous quantity through the ordinal scale and the unit scale, respectively. The calculation of Eq. (2) is realized in practice by taking equidistant sample points from the state intervals $[\alpha_i, \beta_i]$, $i = 1, \dots, n$, of the parent nodes and integrating normal distributions truncated to $[0, 1]$ over the state interval $[\alpha_C, \beta_C]$ of the child node. The alternative weight expressions are discussed in more detail in [25], while the computational process as a whole is explained thoroughly in [30].

3.2. Guidelines for application of RNM to continuous nodes

The application of RNM to nodes with continuous scales is elaborated in [33] with two elicitation guidelines. The first concerns discretizing the continuous scales compatibly with the functioning of RNM. The second is about determining a feasible weight expression and feasible weight values based on the RNM-compatible discretizations.

3.2.1. Guideline for RNM-compatible discretization of continuous scales

The first guideline stems from the following property of RNM concerning a setting in which parent nodes and a child node are ranked nodes with the same number of states.

Property 1. Let the CPT of the child node be generated with RNM using any weight expression and any values of the weights and the variance parameter. Then, for every combination of states of equal rank of the parent nodes, the CPT implies the mode of the child node to be the state with the same rank.

Property 1 follows from the functional forms of the weight expressions and the way all the states of a node are identified with equisized sub-intervals of $[0, 1]$ in RNM. When RNM is applied to nodes with continuous scales, Property 1 should be taken into account. Therefore, the guideline instructs that the interval scales are to be discretized into an equal number of ordinal states in a specific manner. The resulting discretizations are said to be RNM-compatible.

The guideline begins with the expert dividing the continuous scale of each node freely into m discretization intervals, e.g., by considering descriptive labels like *Low*, *Medium*, etc. Next, the expert is asked to consider one by one $m + 1$ scenarios in each of which the values of the parent nodes on their continuous scales correspond to the same boundary point of a state interval on the unit scale $[0, 1]$. For example, with the BN in Fig. 1, there are in total $m + 1 = 4$ scenarios and in one of them, *Price* and *Weekly Usage* have the values €1700 and 30h, as they both are identified with the boundary point $2/3$ on the unit scale. With each of the scenarios, the expert is asked whether the most probable value of the child node on its continuous scale is the value that matches the given boundary point of the parent nodes. Thus, with the example scenario, the expert would be asked whether the most probable value of

Service Time is 73 months, which is identified with $2/3$ on the unit scale. If there is any scenario in which the suggested mode of the child node does not match the expert's view, the discretization intervals of one or more nodes are to be adjusted freely until the matter becomes resolved.

Once the discretizations are carried out in the above manner, any RNM-generated CPT correctly indicates the mode of the child node in the specific scenarios in which all parent nodes are in a state of equal rank. This property of the RNM-compatible discretizations helps one to find a suitable weight expression and weights for the parent nodes later in the elicitation. On the other hand, discretizations formed in ignorance of the functioning of RNM may render the weight expressions infeasible for portraying the probabilistic relationship of the nodes.

3.2.2. Guideline for elicitation of weight expression and weights

Following the RNM-compatible discretization of the continuous nodes, the weight expression, weights, and variance parameter are to be elicited from the expert. A direct but potentially laborious way to do it is through trial and error. Trial and error is laborious because the CPT of the child node must be generated repeatedly with different RNM parameters until it reflects the probabilistic views of the expert well enough. Alternatively, a feasible weight expression and feasible weight values can be elicited indirectly with the guideline presented in [33]. Here, the expert assesses the mode of the child node on the continuous scale in specific scenarios in which the values of the parent nodes on their continuous scales are given. The assessments are then evaluated with regard to various feasibility conditions to determine the feasible weight expression and the feasible weight values.

One key concept on which the guideline is based are piecewise linear mappings defined between the continuous and unit scales of the nodes in accordance with the RNM-compatible discretizations and the functioning of RNM. The graphs in Fig. 2 represent such mappings for the nodes of the example BN in accordance with the discretizations displayed in Fig. 1. These types of piecewise linear mappings are a crucial element of the discretization approaches presented in this paper. They are utilized in the approaches even if RNM parameters could be elicited without the guideline described above. Note that the mappings are directly determined by the RNM-compatible discretizations.

4. Static discretization approach

The guidelines discussed briefly in Section 3.2 extend the application of RNM to nodes with continuous scales. Still, there remain further challenges that can complicate or limit the utilization of RNM with such nodes. This section begins with a discussion of those challenges and then moves on to present the new static discretization approach for resolving them.

4.1. Motivation

When RNM is applied to nodes with continuous scales, those scales need to be discretized to form ranked nodes with discrete ordinal scales. As noted in Section 3.2.1, if the discretizations are formed without considering the functioning of RNM, none of the weight expressions of RNM may seem to be feasible options for representing the probabilistic views of a domain expert about the nodes. This problem can be mitigated by following the guideline for RNM-compatible discretizations. However, the guideline requires the expert to define for the child node and its parents an equal number of states. Moreover, in order to define m states for the nodes, the expert must evaluate the mode of the child node on its continuous scale in $m + 1$ scenarios. These properties of the guideline may give rise to the following challenges. First, if the expert naturally perceives the continuous scales of the nodes through varying numbers of ordinal states, defining suitable discretizations may be difficult. For example, if the expert is used to considering one parent node through three states and another through seven, it is not necessarily straightforward to define an equal number of states for them.

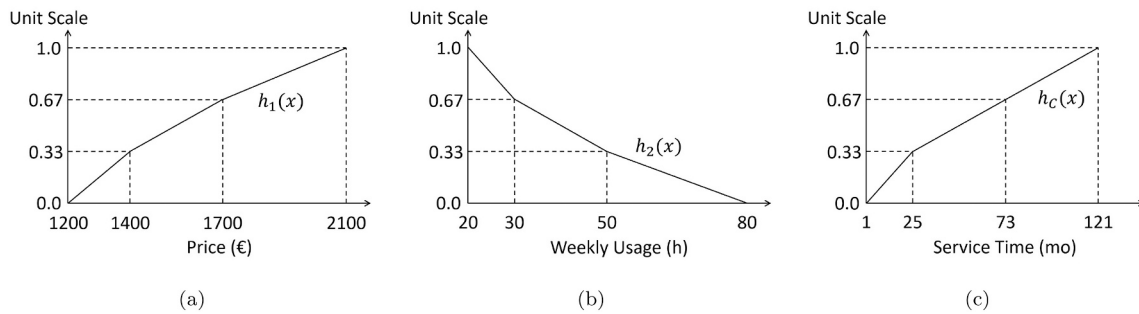


Fig. 2. Piecewise linear mappings $h_1(x)$, $h_2(x)$, and $h_c(x)$ defined for *Price*, *Weekly Usage*, and *Service Time* according to the discretizations in Fig. 1. In each graph, x denotes the quantity on the x-axis.

Second, the expert is likely to lack either the time or the cognitive resources to define several states for the nodes. For instance, defining $m = 10$ ordinal states requires the mode of the child node to be evaluated in $m + 1 = 11$ scenarios. Because of the elicitation effort, the expert may instead prefer to define only $m = 5$ states for the nodes. Yet, the fewer number of states limits the precision of analyses that can be carried out with the BN. Together, the shortcomings of the discretization guideline limit its ease of use and scope of application. Thereby, the scope of efficient use of RNM with regard to continuous nodes also remains restricted.

Another challenge concerns a specific property of RNM: the states of a ranked node are always associated with state intervals of equal width on the unit scale $[0, 1]$. By this property, when the nodes of the example BN are discretized as in Fig. 3 (a), the states $[1200, 1400]$ and $[30, 50]$ of *Price* and *Weekly Usage* are associated with the state intervals $[0, 1/3]$ and $[1/3, 2/3]$. On the other hand, if the discretizations displayed in Fig. 3 (b) are used, the associated state intervals become $[0, 1/4]$ and $[1/2, 3/4]$. Because of this difference in the state intervals, a CPT generated for *Service Time* with fixed RNM parameters has different probabilistic implications depending on the discretizations used. As an illustration, Fig. 3 (a) and (b) display probability distributions obtained for *Service Time* when *Price* and *Weekly Usage* are in states $[1200, 1400]$ and $[30, 50]$, respectively. In both cases, the CPT of *Service Time* is generated with a weight expression called WMAX in which the child node tends to follow any parent node that is in a state associated with large values on the unit scale $[0, 1]$. However, the strength of this tendency depends on the parent node in question. Now, the weights of *Price* and *Weekly Usage* are $w_1 = 1.0$ and $w_2 = 5.0$, respectively. Furthermore, the value of the variance parameter used is $\sigma^2 = 0.001$. Despite the RNM parameters and the states of the parent nodes being the same, the probability

distributions of *Service Time* are completely different from one another.

The above example demonstrates that suitable values of RNM parameters are specific to selected discretizations. If the need arises to change the discretizations after the parameter elicitation, the elicitation has to be carried out again. This feature of RNM limits the flexibility by which the probabilistic relationship of the nodes can be represented and explored after the initial expert elicitation. It may also confuse or be ignored by the users of RNM if they do not understand the functioning of the method thoroughly. For instance, in a given BN application, one might like to carry out analyses using denser discretizations than those specified by a domain expert during the elicitation for RNM. In that case, it is straightforward to rediscritize the nodes and regenerate the CPTs using the original RNM parameters. However, it is likely that phenomena of the kind demonstrated in Fig. 3 will then occur. If this is not understood, the results of the analyses carried out with the BN may turn out to be misleading.

4.2. Underlying principle

The static discretization approach is based on the novel idea that RNM-compatible discretizations and parameters of RNM completely encode the probabilistic relationship between a child node and its parents. Through this idea, the approach provides a means to apply RNM to continuous nodes so that the challenges of the existing practices discussed in Section 4.1 are resolved.

The underlying principle of the static discretization approach is as follows. Consider a child node X_C and its parent nodes X_1, \dots, X_n for which RNM-compatible discretizations have been elicited. Recall from Section 3.2.2 that the discretization of X_i implies a piecewise linear mapping h_i between its continuous scale and the unit scale $[0, 1]$.

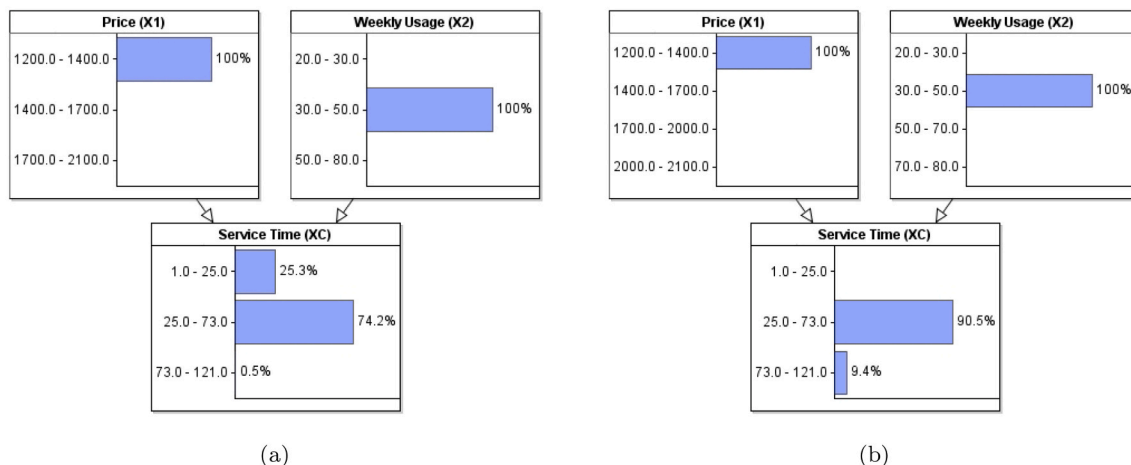


Fig. 3. Different versions of the conditional probability distribution $P(\text{ServiceTime} \mid \text{Price} = [1200, 1400], \text{WeeklyUsage} = [30, 50])$ obtained when using fixed RNM parameter values but alternative discretizations of parent nodes.

Suppose that the exact point value of node X_i on its continuous scale is known to be a_i . In the elicitation guideline presented in Section 3.2.2, this knowledge is associated with knowing that the exact point value of variable χ_i of the regression model in Eq. (1) is $h_i(a_i)$. By the same logic, knowing that the value of X_i on its continuous scale lies within interval $[a_i, b_i]$ corresponds to knowing that χ_i lies within interval $[h_i(a_i), h_i(b_i)]$ on the unit scale. Thus, it can be written

$$\begin{aligned}
 &P(X_C = [a_C, b_C] \mid X_1 = [a_1, b_1], \dots, X_n = [a_n, b_n]) \\
 &= P(\chi_C \in [h_C(a_C), h_C(b_C)] \mid \chi_1 \in [h_1(a_1), h_1(b_1)], \dots, \chi_n \in [h_n(a_n), h_n(b_n)], \chi_C \in [0, 1]).
 \end{aligned}
 \tag{3}$$

Now, Eq. (3) establishes a way to construct a CPT for X_C with arbitrary discretizations of the nodes $X_i, i = 1, \dots, n, C$. By letting $[a_i, b_i]$ represent a discretization interval of node X_i , the CPT of X_C can be calculated according to Eq. (3) by using the regression model of RNM in Eq. (1). The calculation of the right-hand side of Eq. (3) is carried out in practice with the same computational routine that is applied in RNM when calculating the right-hand side of Eq. (2) based on Eq. (1).

In order to calculate the CPT of X_C by Eq. (3), it is sufficient that one knows the piecewise linear mappings h_i as well as the RNM parameters, i.e., the weight expression f , the weights w , and the variance parameter σ^2 included in the regression model of Eq. (1). The piecewise linear mappings h_i are defined by the initial RNM-compatible discretizations. Therefore, these discretizations and the RNM parameters completely encode the probabilistic relationship that any CPT consistent with Eq. (3) represents. The roles of these encoding factors in defining the probabilistic relationship are discussed in more detail in the appendix.

The way RNM-compatible discretizations are used in the static discretization approach distinguishes it from the earlier elaboration [33] on applying RNM to continuous nodes. The differences between these two approaches are next clarified. Recall from Section 3.2 that the concept of RNM-compatible discretization is presented in [33] as a way to coordinate the discretization of nodes on the basis of Property 1. The idea is that by following the guideline in Section 3.2.1, one can form RNM-compatible discretizations that support the discovery of suitable RNM parameters and the construction of sensible CPTs. Yet, once the RNM-compatible discretizations are formed, the corresponding ordinal scales of the nodes are meant to be kept intact for the elicitation of the RNM parameters and the use of the BN [33]. The example in Fig. 3

demonstrates how changing the discretizations can cause probabilistic inconsistencies with the BN. Thus, the nodes are bound to have the same number of states identified with equisized subintervals of the unit scale $[0, 1]$. Any desire to change the discretization of a single node prompts one to define new RNM-compatible discretizations for all the nodes, along with new RNM parameters. In the new static discretization approach, RNM-compatible discretizations are also initially formed for

the nodes. However, from then on, these initial RNM-compatible discretizations serve as a basis for rediscrretization of the nodes. The ability for this rediscrretization follows directly from the novel principle that the initial RNM-compatible discretizations encode the nodes' probabilistic relationship together with the RNM parameters. Through this principle, the nodes can be rediscrretized independently of each other to non-equal numbers of states. The rediscrretization ability can be utilized in the elicitation of RNM parameters and in any analyses that are to be carried out with the BN. Unlike previously in RNM, the states of a rediscrretized single node do not necessarily correspond to equisized sub-intervals on $[0, 1]$.

4.3. Application

The construction of a BN with the static discretization approach is explained next. In what follows, it is assumed that the BN is constructed with a standard BN software and that an implementation of the CPT generation routine of RNM is available.

The static discretization approach consists of six steps, illustrated in Fig. 4. Given a child node X_C and parent nodes X_1, \dots, X_n all measured on continuous scales (see Fig. 5 (a)), Step 1 of the approach is to insert an auxiliary node Y into the BN as depicted in Fig. 5 (b). The node is needed to conduct computation of the conditional probabilities of X_C according to Eq. (3) through the basic CPT generation routine of RNM. The auxiliary node is used only during the construction of the BN and can be removed before the actual use of the BN.

In Step 2, the expert defines initial RNM-compatible discretizations for the nodes $X_i, i = 1, \dots, n, C$, according to the guideline presented in Section 3.2.1. The resulting discretization intervals are then set to be the

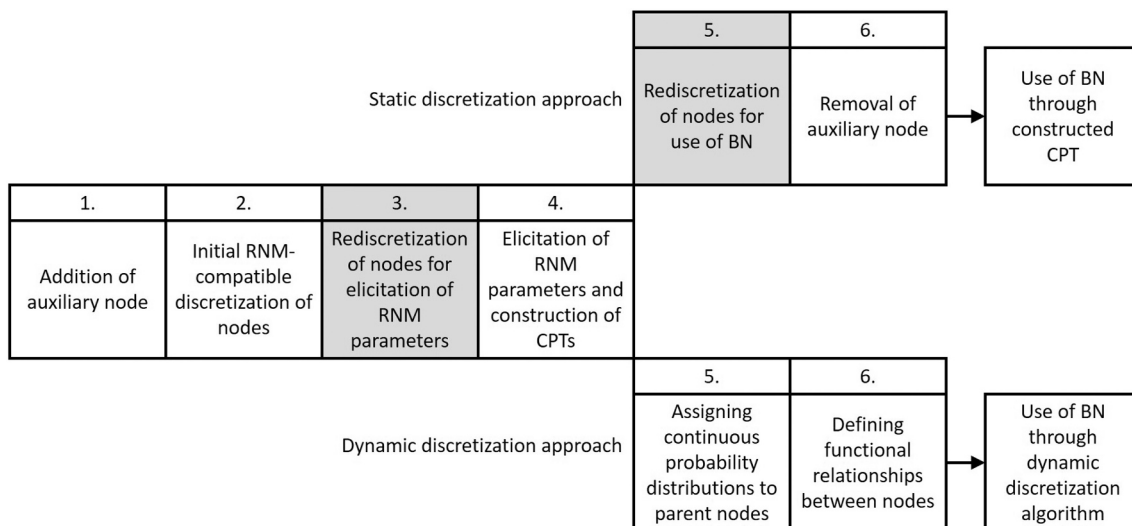


Fig. 4. Steps of the static discretization approach and the dynamic discretization approach. The optional steps are highlighted in grey.

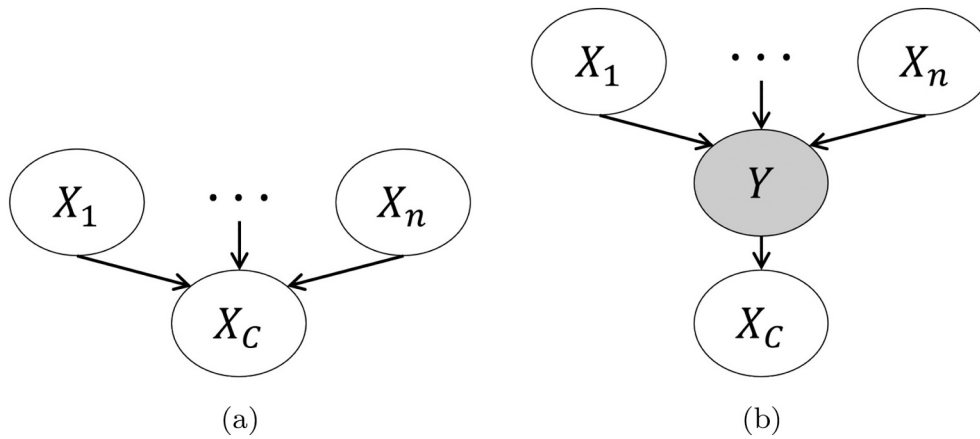


Fig. 5. (a) BN consisting of parent nodes X_1, \dots, X_n and a child node X_C with continuous scales. (b) A modified BN with an auxiliary node Y used in the static discretization approach.

states of X_i . In turn, the states of Y are defined as the state intervals of equal width on the unit scale $[0, 1]$ that are associated with the states of X_C in RNM. Referring to Section 3.2.2, the initial RNM-compatible discretizations establish piecewise linear mappings between the continuous and unit scales of the nodes X_i . These mappings are utilized later in the approach.

Step 3 is optional rediscrretization of the nodes for the elicitation of RNM parameters. At least the elicitation of the variance parameter requires a trial and error procedure, in which the expert has to evaluate CPTs generated with alternative parameter values. Therefore, it is beneficial if the states of the nodes correspond to the discretizations through which the expert naturally perceives the nodes' probabilistic relationship. To this end, Step 3 provides the expert the option to freely rediscrretize nodes X_i into node-specific numbers of consecutive discretization intervals, which can be of varying widths. If the discretization of X_C is changed, the discretization of Y is to be updated accordingly. For any discretization interval $[a_C, b_C]$ of X_C , Y should have a corresponding discretization interval of the form $[h_C(a_C), h_C(b_C)]$.

Step 4 is the elicitation of RNM parameters from the expert and the construction of CPTs for the nodes Y and X_C . It starts with the construction of the CPT of X_C according to

$$P(X_C = [a_C, b_C] \mid Y = [h_C(a_C), h_C(b_C)]) = 1, \quad (4)$$

which reflects a deterministic relationship $Y = h_C(X_C)$ concerning the continuous scales of the nodes. The elicitation of the RNM parameters is linked to the construction of the CPT of Y . One way to elicit the RNM parameters is through trial and error. This involves the expert reviewing probability distributions $P(X_C \mid X_1 = [a_1, b_1], \dots, X_n = [a_n, b_n])$ of interest generated with different RNM parameters until the expert is satisfied. With the auxiliary node Y included in the BN, the expert's evaluation of a distribution $P(X_C \mid X_1 = [a_1, b_1], \dots, X_n = [a_n, b_n])$ is performed as follows. First, the desired evidence is entered into the nodes X_1, \dots, X_n . Next, the CPT of Y is generated with selected RNM parameters in accordance with

$$\begin{aligned} P(Y = [h_C(a_C), h_C(b_C)] \mid X_1 = [a_1, b_1], \dots, X_n = [a_n, b_n]) \\ = P(\chi_C \in [h_C(a_C), h_C(b_C)] \mid \chi_1 \in [h_1(a_1), h_1(b_1)], \dots, \chi_n \in [h_n(a_n), h_n(b_n)], \chi_C \in [0, 1]), \end{aligned} \quad (5)$$

where $[a_i, b_i]$ is a discretization interval of node X_i . The right-hand side of the equation is computed in practice with the CPT generation routine of RNM. The expert then reviews the marginal probability distribution of

X_C updated in the BN based on the new CPT of Y . If necessary, the RNM parameters are adjusted, the CPT of Y is regenerated, and the marginal distribution of X_C is reviewed again.

Instead of eliciting RNM parameters by trial and error alone, the guideline in Section 3.2.2 can also be utilized to determine a suitable weight expression and weights. However, the use of the guideline is not mandatory in the application of the static discretization approach.

Step 5 is optional rediscrretization of the nodes for the use of the BN. After Step 4, the probabilistic relationship between X_C and X_1, \dots, X_n has become fully established. In this regard, the BN can already be used to conduct probabilistic analyses with the nodes by using their present discretizations. However, if these discretizations are not considered ideal for the desired analyses, they can be updated to any others in Step 5. If the discretizations of X_1, \dots, X_n, X_C are altered at this point, the rest of the BN is upgraded as follows. First, the discretization of Y is updated to correspond to that of X_C in the same manner as in Step 3. Next, the CPT of Y is regenerated as per Eq. (5).

Step 6 concerns returning the BN to its original structure, as presented in Fig. 5 (a), without auxiliary node Y . Before the removal of Y , its CPT is put aside. This CPT is then set to be the CPT of X_C . The resulting CPT portrays the probabilistic relationship that the initial RNM-compatible discretizations and the RNM parameters define between X_C and X_1, \dots, X_n .

4.4. Illustrative example

The use of the static discretization approach is demonstrated with the example BN in Fig. 1. The demonstration is carried out using AgenaRisk 5.0, which includes an implementation of RNM.

Fig. 6 (a) displays the auxiliary node Y attached to the original nodes *Price* (X_1), *Weekly Usage* (X_2), and *Service Time* (X_C) according to Step 1 of the approach. Furthermore, the states of X_i shown in the figure are the RNM-compatible discretizations elicited from the expert in Step 2. The associated state intervals are those in Fig. 1, whereas the piecewise linear mappings h_i , defined by the discretizations, are presented in

Fig. 2. The states of Y correspond to the state intervals of X_C . No specific information about the probability distributions of X_i has yet been stated. This is reflected by all the nodes having uniform distributions.

Instead of the discretizations in Fig. 6 (a), the expert finds it easier to

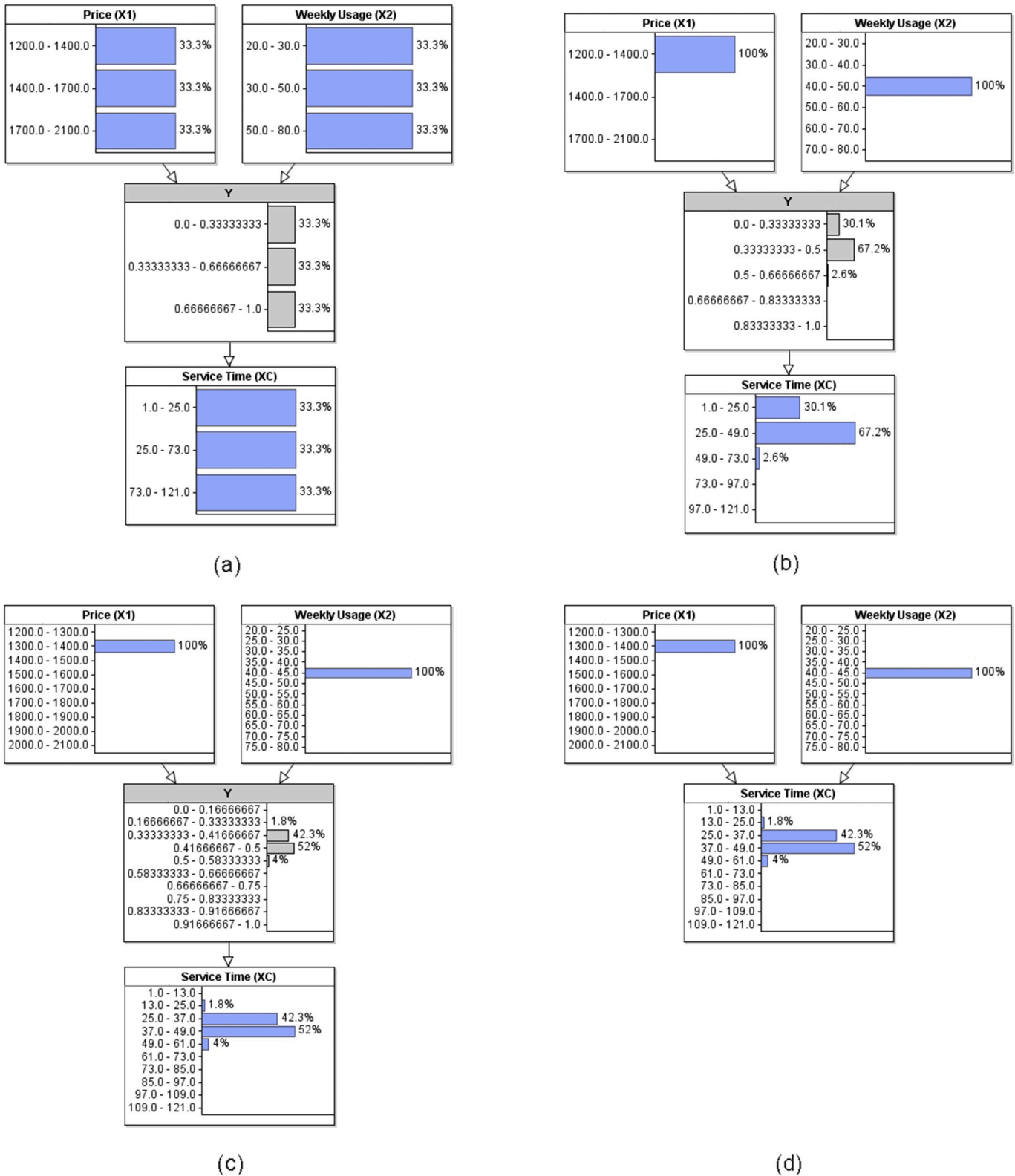


Fig. 6. AgenaRisk implementation for applying the static discretization approach to the example BN. In (a), the nodes have initial RNM-compatible discretizations and uniform probability distributions. In (b), updated discretizations used in the elicitation of RNM parameters, and the evidence $X_1 = [1200, 1400]$ and $X_2 = [40, 50]$ inserted after the construction of CPTs of Y and X_C . In (c), final discretizations for the use of the BN, and the evidence $X_1 = [1300, 1400]$ and $X_2 = [40, 45]$ inserted after updating the CPTs. In (d), the BN after the removal of the auxiliary node Y with its CPT copied to X_C .

Table 2

Elicitation scenarios and mode assessments for determining feasible weight expression and weights.

Elicitation scenario	Price (€)	Weekly usage (h)	Mode of service time (mo)
E^1	2100	80	49
E^2	1200	20	97

consider the probabilistic relationship of the nodes so that discretization intervals of widths 10 hours and 24 months are applied with *Weekly Usage* and *Service Time*, respectively. Therefore, the optional Step 3 (rediscrretizing the nodes for the elicitation of RNM parameters) is utilized (see Fig. 6 (b)). The discretization of Y is also updated by using the function h_C depicted in Fig. 2 (c). Its new discretization corresponds to mapping the discretization intervals of X_C into $[0, 1]$ by h_C .

In Step 4, a suitable weight expression and weights of the parent nodes are elicited by utilizing the guideline discussed in Section 3.2. This helps to reduce the search of RNM parameters by trial and error. The guideline requires the expert to evaluate the mode of *Service Time* in the two scenarios E^1 and E^2 that are defined by the point values of *Price* and *Weekly Usage* presented in Table 2. The example mode assessments that are shown in the table imply that WMAX is a feasible weight expression, with $w_1 = 1.0$ and $w_2 = 5.0$ being the weights of *Price* and *Weekly Usage*, respectively.

The variance parameter is determined by having the expert try different values and review the corresponding CPTs. Here, the state combination (*Price* = [1200, 1400], *Weekly Usage* = [40, 50]) is among those that the expert wants to consider. Fig. 6 (b) illustrates the probability distribution of *Service Time* obtained for this state combination when the CPT of X_C is defined in line with Eq. (4), and the CPT of Y is generated according to Eq. (5) with the variance parameter $\sigma^2 = 0.001$. This value is selected by the expert.

Suppose that for analyses to be carried out with the BN, the discretizations in Fig. 6 (b) are deemed insufficiently dense. In that case, the optional Step 5 (rediscrretizing the nodes for the use of the BN) is applied. Fig. 6 (c) displays the nodes *Price*, *Weekly Usage*, and *Service Time* rediscrretized to uniform precisions of 100 euros, 5 hours, and 12 months, respectively. The states of Y are again the intervals that are obtained when the discretization intervals of X_C are mapped onto $[0, 1]$ with h_C . The figure presents the probability distribution of *Service Time* when *Price* and *Weekly Usage* are fixed to the states [1300, 1400] and [40, 45].

Comparison of Fig. 6 (b) and (c) indicates how the denser discretizations allow portraying the probabilistic relationship of the nodes in greater detail. First, with the denser discretization of *Service Time* in Fig. 6 (c), one obtains more detailed information of the dispersion of the probability mass over the continuous scale than what is obtained in Fig. 6 (b). Second, the denser discretization enables the entering of more detailed evidence into the BN. For example, suppose one knows that the price and weekly usage of a machine are €1350 and 42 hours, respectively. In Fig. 6 (b), this knowledge is expressed with the evidence *Price* = [1200, 1400] and *Weekly Usage* = [40, 50] entered into the BN. On the other hand, in Fig. 6 (c), the corresponding evidence is *Price* = [1300, 1400] and *Weekly Usage* = [40, 45]. The more precise evidence in Fig. 6 (c) leads to a new insight about the probability distribution of *Service Time*. For instance, while the probability that the service time stays below 25 months is about 30% in Fig. 6 (b), it is less than 2% in Fig. 6 (c).

Fig. 6 (d) displays the example BN after Step 6, i.e., back in the original structure without the auxiliary node Y . Here, the CPT of *Service Time* has been defined to be the same as the CPT of Y in Fig. 6 (c). Therefore, the probability distribution of *Service Time* in Fig. 6 (d) corresponds to the probability distributions of both Y and *Service Time* in Fig. 6 (c).

To finish the demonstration, Fig. 7 displays the BN in Fig. 6 (d) next to a version where the nodes have alternative final discretizations, but where the CPT of *Service Time* is constructed with the same RNM parameters as before. Comparison of Fig. 7 (a) and (b) indicates that though the discretizations have changed, the probability distribution of *Service Time* for the given states of *Price* and *Weekly Usage* remains the same. The corresponding phenomenon does not occur when using RNM in its basic form (see Fig. 3). The consistency in the nodes' probabilistic behavior between Fig. 7 (a) and (b) is an effect of using the piecewise linear mappings h_i in the CPT generation. For example, in the CPT generation equation (3), the discretization interval [1300, 1400] of *Price* is identified with the interval $[h_1(1300), h_1(1400)] = [1/6, 1/3]$ of the unit scale. This connection holds regardless of the other discretization intervals of *Price*.

5. Dynamic discretization approach

The dynamic discretization approach enables representing probabilistic relationships established with RNM with discretizations that update automatically during the use of the BN. Like the static discretization

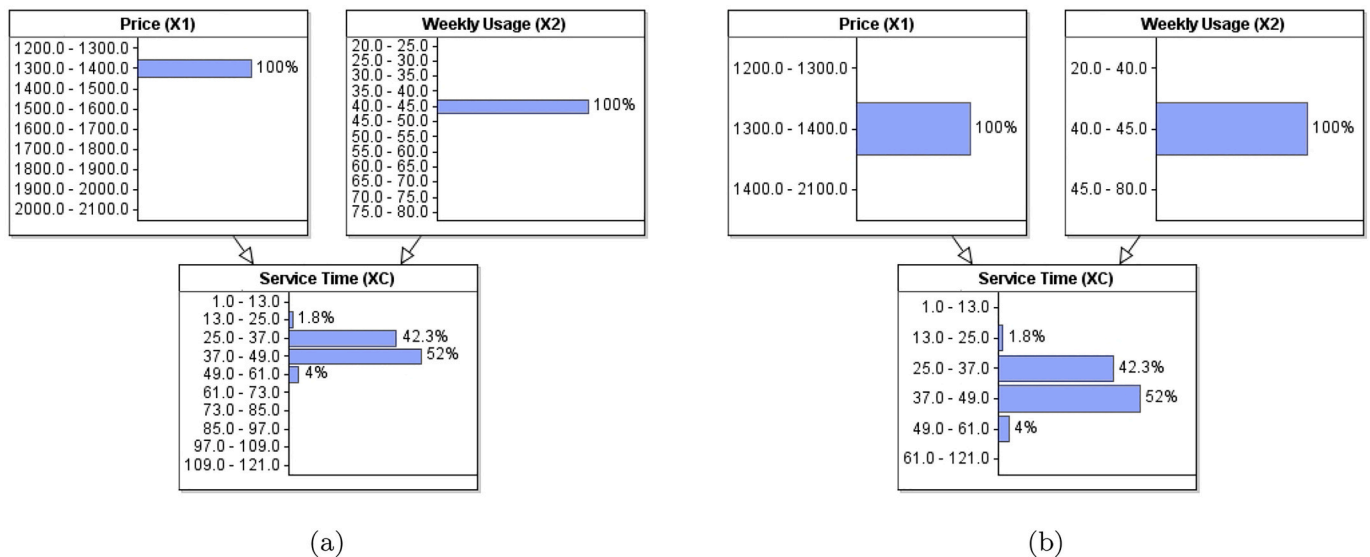


Fig. 7. Conditional probability distribution $P(\text{ServiceTime} \mid \text{Price} = [1300, 1400], \text{WeeklyUsage} = [40, 45])$ generated with fixed RNM parameters stays intact despite changes in discretizations of the parent nodes.

approach, also this creative way of using RNM is based on the principle that the nodes' probabilistic relationship is encoded together by the initial RNM-compatible discretizations and the RNM parameters. The two approaches are complementary to each other. The static approach enables the application of RNM to continuous nodes through freely selected discretizations. Furthermore, the approach takes into account the concept of RNM-compatibility in the generation of CPTs. However, increasing the granularity of the discretizations increases the sizes of the CPTs. This may cause problems concerning the computer memory requirements for storing the CPTs or conducting probabilistic inference in the BN [19,35]. In the dynamic approach, the continuous scales of the nodes are discretized non-uniformly based on the evidence entered into the BN. The granularity is always higher at areas of high probability on the nodes' continuous scales. This helps to portray their probability distributions accurately without the need for dense, uniform static discretizations. For example, suppose that the mean of a child on its continuous scale varies a lot depending on the states of the parent nodes. Then, a need for accurate statistics concerning the tails of its distribution may require impractically dense static discretizations. In such cases, the dynamic approach provides a means to get the statistics. The matter is demonstrated in Section 5.3.

5.1. Underlying principle

The dynamic discretization approach combines the uses of RNM and a dynamic discretization algorithm of Neil et al. [35]. In the algorithm, continuous nodes of a BN are discretized based on their probability distributions. A brief outline of the algorithm is provided below. The principle concerning its utilization with RNM is explained thereafter.

5.1.1. Dynamic discretization algorithm

The dynamic discretization algorithm is designed for discretizing continuous nodes in hybrid BNs that contain both discrete and continuous nodes. Its methodology has not been linked to RNM in the existing literature. The idea is that whenever new evidence is entered into the BN, the algorithm rediscrretizes the continuous nodes. The rediscrretization is carried out iteratively based on the probability distributions that the evidence imposes on the nodes.

Starting with some initial discretizations, the iteration rounds in the algorithm proceed as follows. First, CPTs of child nodes are determined based on the current discretizations and on the functional relationships defined between them and their parent nodes. After that, discrete marginal probability distributions of all nodes are computed in accordance with the evidence entered into the BN. On the basis of these distributions, entropy error values over the discretization intervals are calculated for all the nodes. With each node, the discretization interval with the largest entropy error is then split in two while consecutive discretization intervals with zero entropy errors are merged together. After this, the algorithm continues to the next iteration round. The rediscrretization of a given node stops when a convergence threshold or the maximum number of iteration rounds is reached.

The algorithm enables entering point-valued evidence about the continuous nodes into the BN. If a point value x is assigned to a continuous node X , the algorithm divides the continuous scale $[a, b]$ of X into intervals $[a, x - \delta x]$, $[x - \delta x, x + \delta x]$, and $[x + \delta x, b]$, where δx represents a selected tolerance bound on x . During the course of the algorithm, this discretization of X is kept fixed and the interval $[x - \delta x, x + \delta x]$ is given 100% probability whenever marginal probability distributions of other nodes are computed.

While the discretizations provided by the algorithm are not uniform, they are, for any evidence entered into the BN, always denser in the high probability areas of the continuous scales. Therefore, compared to uniform static discretizations, the algorithm can produce good discrete approximations of continuous distributions with a smaller number of discretization intervals. In turn, the smaller number of discretization intervals used leads to smaller CPTs, which require less computer

memory for their storage or for the conducting of probabilistic inference in the BN. Naturally, the need to rerun the algorithm whenever new evidence is entered into the BN is a source of computational burden that is absent when using static discretizations. In the algorithm, adjusting the convergence threshold is a means to control the trade-off between the accuracy reached with the discretizations and the computational burden. For a more thorough presentation of the algorithm, see [19,35, 55,56]. The algorithm is implemented in AgenaRisk software [49].

5.1.2. Use of RNM with dynamic discretization algorithm

The basis for combining RNM with the dynamic discretization algorithm is that a probabilistic relationship between parent nodes X_1, \dots, X_n and a child node X_C has been established according to Steps 1–4 of the static discretization approach (see Fig. 4). In other words, the relationship has been encoded by the initial RNM-compatible discretizations and the RNM parameters selected by the expert. Recall that in the static approach, this property enables one to generate CPTs representing the relationship with any assigned discretizations of the nodes. In the dynamic approach, the same property is utilized to discretize the nodes with the dynamic discretization algorithm.

Provided that continuous probability distributions are defined for the parent nodes, the nodes can be discretized with the algorithm as follows. Let $[a_i, b_i]$ denote a discretization interval of X_i on a specific iteration round of the algorithm. Then, the conditional probability $P(X_C = [a_C, b_C] \mid X_1 = [a_1, b_1], \dots, X_n = [a_n, b_n])$ for the CPT of X_C on that round is generated with RNM according to Eq. (3). This way, once the execution of the algorithm terminates, the CPT of X_C over its new discretization intervals depicts the probabilistic relationship encoded by the initial RNM-compatible discretizations and the RNM parameters.

5.2. Application

The application of the dynamic discretization approach is explained next. The description concerns a setting in which a BN is constructed with a software that includes implementations of the dynamic discretization algorithm and the RNM routine.

The dynamic discretization approach consists of the six steps illustrated in Fig. 4. As mentioned above, Steps 1–4 are the same as in the static discretization approach. In Step 5, continuous probability distributions are assigned to the nodes X_1, \dots, X_n . These can be normal distributions or (possibly piecewise) uniform distributions decided by an expert or estimated from the data. In Step 6, the dependencies of X_C and Y on their parent nodes are defined with functional relationships. The relationship of X_C and Y is

$$X_C = h_C^{-1}(Y), \quad (6)$$

where h_C^{-1} is the inverse piecewise linear mapping of h_C . For Y and X_1, \dots, X_n , the relationship is

$$Y \sim TNormal(\mu, \sigma^2, 0, 1), \quad \mu = f(h_1(X_1), \dots, h_n(X_n), \boldsymbol{w}), \quad (7)$$

where $TNormal(\mu, \sigma^2, a, b)$ denotes a normal distribution with mean μ and variance σ^2 truncated to the interval $[a, b]$.

After Step 6, the BN can be used by applying the dynamic discretization algorithm. Through the functional relationships defined in Step 6, the CPTs of X_C and Y are generated in accordance with Eqs. (4) and (5) during the execution of the algorithm. Any point-valued evidence entered into the nodes X_i is handled as described in Section 5.1.1. Unlike in the static discretization approach, the auxiliary node Y now remains in the BN. However, its role is purely computational.

5.3. Illustrative example

The dynamic discretization approach is demonstrated with the example BN in Fig. 1 and AgenaRisk 5.0. The demonstration begins from the point that the probabilistic relationship between *Price*, *Weekly Usage*,

and *Service Time* has been established with the static discretization approach as described in Section 4.4. This relationship is encoded by the initial RNM-compatible discretizations shown in Fig. 6 (a) and the RNM parameters $f = \text{WMAX}$, $w_1 = 1.0$, $w_2 = 5.0$, and $\sigma^2 = 0.001$.

The earlier application of the static approach means that Steps 1–4 of the dynamic approach have already been carried out. In order to proceed with the dynamic approach, the auxiliary node Y displayed in Fig. 6 (a)–(c) is first returned to the BN. This is followed by Step 5 in which *Price* (X_1) and *Weekly Usage* (X_2) are estimated to follow doubly truncated normal distributions $TNormal(\mu = 1600, \sigma^2 = 90,000, 1200, 2100)$ and $TNormal(\mu = 40, \sigma^2 = 100, 20, 80)$, respectively. Next, in Step 6,

functional relationships are established between the nodes. Referring to Eq. (6) and the mapping h_C depicted in Fig. 2 (c), *Service Time* (X_C) and Y get the functional relationship

$$X_C = h_C^{-1}(Y) = \begin{cases} 72 * Y + 1, & Y \in [0, 1/3] \\ 144 * Y - 23, & Y \in (1/3, 1] \end{cases} \quad (8)$$

In turn, by Eq. (7), the functional relationship established between Y and its parent nodes *Price* (X_1) and *Weekly Usage* (X_2) is

$$Y \sim TNormal(\mu, \sigma^2 = 0.001, 0, 1), \\ \mu = \text{WMAX}(h_1(X_1), h_2(X_2), w_1 = 1.0, w_2 = 5.0), \quad (9)$$

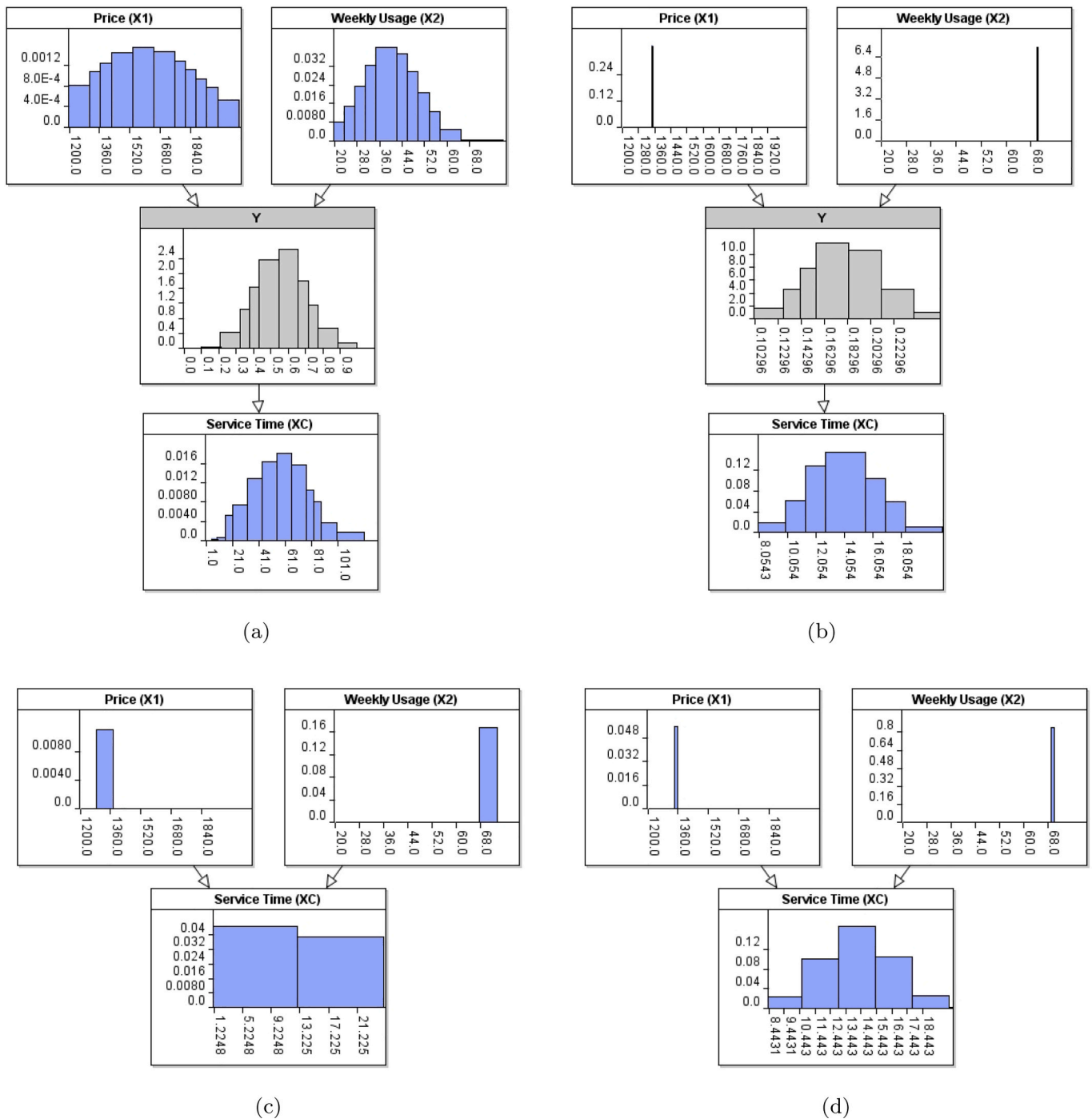


Fig. 8. In (a) and (b), AgenaRisk implementation for applying the dynamic discretization approach to the example BN. In (a), the BN is without evidence and in (b), with the point-valued evidence $X_1 = 1350$ and $X_2 = 70$. In (c) and (d), corresponding BNs with all the nodes having 10 and 50 equisized static discretization intervals, respectively. In (c), the BN is with the evidence $X_1 = [1290, 1380]$ and $X_2 = [68, 74]$. In (d), the evidence is $X_1 = [1344, 1362]$ and $X_2 = [69.2, 70.4]$.

where

$$WMAX(x_1, x_2, w_1, w_2) = \max\left\{\frac{w_1 * x_1 + x_2}{w_1 + 1}, \frac{w_2 * x_2 + x_1}{w_2 + 1}\right\} \quad (10)$$

is the functional form of the weight expression WMAX. The piecewise linear mappings h_1 and h_2 depicted in Fig. 2 (a) and (b) are

$$h_1(x) = \begin{cases} (1/600)x - 2, & x \in [1200, 1400] \\ (1/900)x - 11/9, & x \in (1400, 1700], \\ (1/1200)x - 3/4, & x \in (1700, 2100] \end{cases} \quad (11)$$

$$h_2(x) = \begin{cases} -(1/30)x + 5/3, & x \in [20, 30] \\ -(1/60)x + 7/6, & x \in (30, 50]. \\ -(1/90)x + 8/9, & x \in (50, 80] \end{cases} \quad (12)$$

Fig. 8 (a) displays the modified example BN with the auxiliary node Y when the dynamic discretization algorithm has been run with no evidence in the BN. The initially undivided continuous scales of the nodes have been discretized with 10 iteration rounds of the algorithm. The use of the algorithm is reflected by non-uniform widths of the discretization intervals. In Fig. 8 (b), point-valued evidence $X_1 = 1350$ and $X_2 = 70$ has been incorporated, and the discretizations and the probability distributions have been updated by another run of the algorithm. The percentage tolerance bound used for the evidence is 0.1% whereby the discretization intervals of *Price* and *Weekly Usage* that get a probability of 100% are [1348.65, 1351.35] and [69.93, 70.07], respectively. This evidence is pointed out by single probability peaks on their continuous scales in Fig. 8 (b). For the nodes Y and X_C , the part of the probability distribution displayed is the segment between the 1st and 99th percentiles. Thus, the majority of the discretization intervals of these nodes is concentrated to a much narrower region than in Fig. 8 (a).

To get a concrete sense of the benefit obtained with the dynamic discretization approach, Fig. 8 (c) and (d) display BNs corresponding to that in Fig. 8 (b) but constructed with the static discretization approach. In Fig. 8 (c), the nodes X_1 , X_2 , and X_C each have 10 discretization intervals of equal width. In turn, they all have 50 equisized discretization intervals in Fig. 8 (d). In both figures, the parent nodes are fixed to the states that best correspond to the point-valued evidence $X_1 = 1350$ and $X_2 = 70$ in Fig. 8 (b). The evidence in Fig. 8 (c) is $X_1 = [1290, 1380]$ and $X_2 = [68, 74]$ whereas $X_1 = [1344, 1362]$ and $X_2 = [69.2, 70.4]$ are used as evidence in Fig. 8 (d). The probability distributions of X_C are again displayed so that the x-axes cover the portion between the 1st and 99th percentiles of the distributions. There is evident variation in the accuracy by which the probability distribution of *Service Time* is portrayed in Fig. 8 (b)–(d). Whereas the displayed range of the continuous scale consists of 2 discretization intervals in Fig. 8 (c), the numbers of intervals in Fig. 8 (b) and (d) are 7 and 5, respectively. Thus, the BN in Fig. 8 (c) provides less insight about the probability distribution of X_C than the other two. On the other hand, the accuracy of the probability distributions in Fig. 8 (b) and (d) seems comparable. However, it should be noted that the result in Fig. 8 (d) is obtained with a total number of 50 discretization intervals, whereas the corresponding number in Fig. 8 (b) is only 13. In this regard, the dynamic approach reaches the same accuracy as the static approach with a considerably smaller overall discretization density.

Table 3
Statistics of probability distributions of X_C in Fig. 8 (b)–(d).

Statistic	Dynamic disc. 10 rounds Fig. 8 (b)	Static disc. 10 intervals Fig. 8 (c)	Static disc. 50 intervals Fig. 8 (d)
Mean	14.0	12.6	14.3
Median	14.0	12.2	14.3
Variance	6.1	35.8	5.9
1st percentile	8.1	1.2	8.4
5th percentile	10.1	2.1	10.2

The accuracy of the results is also reflected by the statistics of the probability distributions of X_C presented in Table 3. Especially the variance and the percentile values concerning Fig. 8 (c) are different from those of Fig. 8 (b) and (d). The static approach with 10 discretization intervals (Fig. 8 (c)) results in the variance, the 1st, and the 5th percentile values being 35.8, 1.2, and 2.1, respectively. The corresponding values with the dynamic approach (Fig. 8 (b)) are 6.1, 8.1, and 10.1, whereas the static approach with 50 intervals (Fig. 8 (d)) produces the values 5.9, 8.4, and 10.2. Thus, if one needs accurate statistics on the probabilistic relationship of the nodes, the dynamic approach provides them with a lot smaller discretization density than the static approach.

The example demonstrates that with a small number of dynamically generated intervals, the dynamic discretization approach allows portraying probabilistic relationships of ranked nodes as accurately as the static approach with a larger number of intervals. This property is useful if the sizes of CPTs resulting from static discretizations start to hamper the efficient use of the BN [35,19]. For instance, the 1st and 5th percentile values reported in Table 3 are common value at risk (VaR) measures calculated in risk analysis [19,57]. In a more complicated BN, the static discretizations required to calculate VaR values accurately enough might result in impractically large CPTs. In such a case, the dynamic discretization approach would provide the only way to solve their values. It is also worth noting that with more complicated BNs, reaching accurate estimates of, e.g., mean and median values of node probability distributions may require impractically dense static discretizations. For example, consider that the machine described with the example BN had different modes of use corresponding to very different lengths of service time. Depending on the mode of use, variances in the service time could then be very small compared to the overall width of its measurement scale. This type of situation would require dense static discretizations to accurately portray the statistics of the conditional probability distributions of the service time. Yet, the dynamic discretization approach could always provide accurate statistics with a small number of dynamically generated discretization intervals.

6. Conclusion

This paper discussed the following challenges related to the discretization of continuous nodes when their probabilistic relationships are portrayed with CPTs constructed through expert elicitation using RNM. First, ignorance of the functioning of RNM can lead to discretizations for which it is impossible to generate sensible CPTs. Second, an existing guideline for forming RNM-compatible discretizations requires defining an equal number of ordinal states for all the nodes. Moreover, the guideline is laborious for constructing dense discretizations. Third, changing the discretizations after the CPT construction demands re-elicitation of the RNM parameters in order to regenerate the CPT. Otherwise, the probabilistic implications of the new CPT become inconsistent with the original one.

To deal with the challenges, the paper presented two new discretization approaches, referred to as “static” and “dynamic”, for the application of RNM. In the static one, the desired discretizations of the nodes are selected during the construction of the BN, and the discretizations are not changed during the use of the BN. In the dynamic one, discretizations are determined by a dynamic discretization algorithm. The algorithm updates the discretizations during the use of the BN based on the entered evidence and the probabilistic relationship of the nodes established while constructing the BN. In both approaches, the functioning of RNM is taken into account while still allowing the nodes to have unequal numbers of ordinal states. The novel underlying principle is that the nodes’ probabilistic relationship is defined by both the initial RNM-compatible discretizations and the RNM parameters elicited from a domain expert. This relationship is portrayed consistently by CPTs generated with RNM, independent of the discretizations used. Besides presenting the technical idea behind the approaches, the paper explained and demonstrated how they are applied and implemented

with a suitable standard BN software.

The new approaches have several beneficial features for easy and diverse application of RNM to continuous nodes. The initial RNM-compatible discretizations form the basis for the construction of sensible CPTs and require that all the nodes get the same number of ordinal states. However, after assessing the initial discretizations, the expert can assign for the nodes any discretizations which, in the expert's view, most naturally reflect the nodes' probabilistic relationship. This feature facilitates the determination of suitable RNM parameters. In the static discretization approach, another point for optional rediscrretization is provided after the parameter elicitation. Here, the discretizations used in the elicitation can be replaced with any others that might be considered more adequate for the analyses to be carried out with the BN. The use of the BN is therefore not restricted to the discretizations preferred by the expert in the elicitation. Another beneficial feature concerns nodes with more than one child or that themselves are both a parent and a child. With these nodes, the initial RNM-compatible discretizations and the discretizations used in the parameter elicitation can be selected separately for each of their roles. This provides additional flexibility to the defining of the probabilistic relationships. It is also helpful if, e.g., different experts are used to construct different CPTs or many experts are involved in the construction of one. If the discretizations desired in the static approach cause computational memory problems, the dynamic approach enables the probabilistic relationship of the nodes to be explored with the desired level of detail. The dynamic approach is useful especially when one would like to enter point-valued evidence into the BN or explore the precise statistics of the probability

distributions. However, if there is no need for such analyses, the use of the static approach is recommendable as it spares one the computational effort of recurring rediscrretization of the nodes. All of these beneficial features of the new approaches stem from the principle that the nodes' probabilistic relationship is defined by both the initial RNM-compatible discretizations and the RNM parameters. Because of the novelty of this principle, these features are lacking from the existing practices of applying RNM to continuous nodes.

The following themes have been identified as avenues for further research. First, it could be experimentally explored how the common number of states defined for nodes with initial RNM-compatible discretizations affects CPTs constructed with the new discretization approaches. Based on the results, recommendations could be established about the suitable number of initial states. Another future theme concerns determining initial RNM-compatible discretizations and RNM parameters by data fitting. One aspect of this theme would be to compare how CPTs constructed by using the new approaches and data fitting compare to CPTs estimated from the same data through other means. A third future topic is a detailed comparison of computational properties of the static and dynamic discretization approaches. This could reveal circumstances additional to simple computer memory shortage under which it would be better to use the dynamic approach instead of the static one. However, in an application, it is straightforward to check this by trying which one of the approaches is more befitting. In that regard, the presentation of the approaches given in this paper already serves well for their deployment in decision support systems utilizing BNs.

Appendix

The appendix discusses the roles of initial RNM-compatible discretizations and RNM parameters in defining the probabilistic relationship of nodes in the discretization approaches presented in the paper.

To explain the role of the initial RNM-compatible discretizations, consider parent nodes X_1, \dots, X_n and a child node X_C for which those discretizations corresponding to m ordinal states have been determined. The discretization intervals of the nodes are associated with state intervals of the form $[\frac{k-1}{m}, \frac{k}{m}]$, $k = 1, \dots, m$ on unit scales. Furthermore, the discretization of node X_i defines a piecewise linear mapping h_i between its continuous and unit scales. Consider then the point values a_1, \dots, a_n of the parent nodes such that $h_i(a_i) = \alpha$ for all $i = 1, \dots, n$. Then, independent of the RNM parameters used, the associated mode of the child node on its continuous scale is the value a_C given by $a_C = h_C^{-1}(\alpha)$ [33].

By referring to the above result, the role of the initial RNM-compatible discretizations in defining the probabilistic relationship of the nodes can be characterized as follows. The boundary points of the discretization intervals are benchmarks regarding which the expert has verified that when the parent nodes X_i , $i = 1, \dots, n$, have values $a_i = h_i^{-1}(\frac{k}{m})$ on their continuous scales, the mode of the child node X_C on its continuous scale is $a_C = h_C^{-1}(\frac{k}{m})$, with any $k = 0, \dots, m$. Between these benchmarks, simultaneous and equally large percentage changes in the point values of the parent nodes imply an equally large percentage change in the mode of the child node. To elaborate this idea, think about discretization intervals $[a_i, b_i]$ of the nodes X_i , $i = 1, \dots, n$, C that are all associated with a state interval $[\frac{k-1}{m}, \frac{k}{m}]$. Suppose the point values of the parent nodes X_i , $i = 1, \dots, n$, on their continuous scales are z_i such that $z_i \in [a_i, b_i]$ and $\frac{z_i - a_i}{b_i - a_i} = u$. Then, the mode implied for the child node X_C is z_C , fulfilling $\frac{z_C - a_C}{b_C - a_C} = u$. Thus, the initial RNM-compatible discretizations alone imply the mode of the child node on its continuous scale for specific combinations of point values of the parent nodes.

The role of the RNM parameters is to complement the probabilistic information that the initial RNM-compatible discretizations imply about the nodes. When the weight expression f and the weights of the parent nodes w are known, one can determine the mode \hat{a}_C of the child node X_C on its continuous scale for any combination of point values a_1, \dots, a_n of the parent nodes X_1, \dots, X_n on their continuous scales. Referring to Eq. (1), the related equation is $\hat{a}_C = h_C^{-1}(f(h_1(a_1), \dots, h_n(a_n), w))$. Thereby, the role of the weight expression and the weights is to define the mode of the child node for all those point values of the parent nodes for which it is not directly implied by the initial RNM-compatible discretizations. Exact interpretations of the weights in different weight expressions are presented in [33]. The role of the variance parameter σ^2 is to define the level of dispersion of probability mass around the determined mode \hat{a}_C on the continuous scale of X_C . To summarize, the RNM parameters together with the RNM-compatible discretizations establish the probabilistic relationship of the nodes X_1, \dots, X_n, X_C so that a CPT for X_C can be generated in accordance with Eq. (3).

References

- [1] A. Hill, C.H. Joyner, C. Keith-Jopp, B. Yet, C.T. Sakar, W. Marsh, D. Morrissey, A Bayesian network decision support tool for low back pain using a RAND appropriateness procedure: proposal and internal pilot study, *JMIR Res. Protoc.* 10 (1) (2021) 1–11, e21804.
- [2] Y. Zhou, N. Fenton, C. Zhu, An empirical study of Bayesian network parameter learning with monotonic influence constraints, *Decis. Support Sys.* 87 (2016) 69–79.
- [3] A.C. Constantinou, M. Freestone, W. Marsh, J. Coid, Causal inference for violence risk management and decision support in forensic psychiatry, *Decis. Support Sys.* 80 (2015) 42–55.
- [4] B. Yet, K. Bastani, H. Raharjo, S. Lifvergren, W. Marsh, B. Bergman, Decision support system for warfarin therapy management using Bayesian networks, *Decis. Support Sys.* 55 (2) (2013) 488–498.
- [5] C. Bielza, J.A.F. del Pozo, P.J. Lucas, Explaining clinical decisions by extracting regularity patterns, *Decis. Support Sys.* 44 (2) (2008) 397–408.
- [6] K. Topuz, D. Delen, A probabilistic Bayesian inference model to investigate injury severity in automobile crashes, *Decis. Support Sys.* 150 (2021) 1–13, e113557.

- [7] L. Kaikkonen, T. Parviainen, M. Rahikainen, L. Uusitalo, A. Lehikoinen, Bayesian networks in environmental risk assessment: a review, *Integr. Environ. Asses. Manag.* 17 (1) (2021) 62–78.
- [8] E. Tosoni, A. Salo, J. Govaerts, E. Zio, Comprehensiveness of scenarios in the safety assessment of nuclear waste repositories, *Reliab. Eng. Sys. Saf.* 188 (2019) 561–573.
- [9] G. Zhang, V.V. Thai, Expert elicitation and Bayesian network modeling for shipping accidents: a literature review, *Saf. Sci.* 87 (2016) 53–62.
- [10] M. Naderpour, J. Lu, G. Zhang, An intelligent situation awareness support system for safety-critical environments, *Decis. Support Sys.* 59 (2014) 325–340.
- [11] F. Sanchez, S. Steria, E. Bonjour, J.-P. Micaelli, D. Monticolo, An approach based on Bayesian network for improving project management maturity: an application to reduce cost overrun risks in engineering projects, *Comput. Indus.* 119 (2020) 1–13, e103227.
- [12] B. Yet, A. Constantinou, N. Fenton, M. Neil, E. Luedeling, K. Shepherd, A Bayesian network framework for project cost, benefit and risk analysis with an agricultural development case study, *Expert Sys. Appl.* 60 (2016) 141–155.
- [13] A. Mancuso, M. Compare, A. Salo, E. Zio, Optimal prognostics and health management-driven inspection and maintenance strategies for industrial systems, *Reliab. Eng. Sys. Saf.* 210 (2021) 1–10, e107536.
- [14] D. Dinis, A.P. Teixeira, A. Barbosa-Póvoa, Foresim-bi: a predictive analytics decision support tool for capacity planning, *Decis. Support Sys.* 131 (2020) 1–13, e113266.
- [15] M.J. Barons, S.K. Wright, J.Q. Smith, Eliciting probabilistic judgements for integrating decision support systems, in: L.C. Dias, A. Morton, J. Quigley (Eds.), *Elicitation: The Science and Art of Structuring Judgement*, Springer, New York, 2017.
- [16] L.S. Malagrino, N.T. Roman, A.M. Monteiro, Forecasting stock market index daily direction: a Bayesian network approach, *Expert Sys. Appl.* 105 (2018) 11–22.
- [17] J. Brynielsson, Using AI and games for decision support in command and control, *Decis. Support Sys.* 43 (4) (2007) 1454–1463.
- [18] L. Falzon, Using Bayesian network analysis to support centre of gravity analysis in military planning, *Europ. J. Oper. Res.* 170 (2) (2006) 629–643.
- [19] N. Fenton, M. Neil, Risk Assessment and Decision Analysis with Bayesian Networks, CRC Press, Boca Raton, FL, USA, 2013.
- [20] R. Neapolitan, *Learning Bayesian Networks*, Pearson Prentice Hall, Upper Saddle River, NJ, USA, 2004.
- [21] S. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, Upper Saddle River, NJ, USA, 2003.
- [22] R.A. Howard, J.E. Matheson, Influence diagrams, *Decis. Anal.* 2 (3) (2005) 127–143.
- [23] M. Druzdel, L. van der Gaag, Building probabilistic networks: “Where do the numbers come from?,” *IEEE Trans. Knowl. Data. Eng.* 12 (4) (2000) 481–486.
- [24] J. Rohmer, Uncertainties in conditional probability tables of discrete Bayesian belief networks: a comprehensive review, *Eng. Appl. Artif. Intell.* 88 (2020) 1–16, 103384.
- [25] N. Fenton, M. Neil, J. Caballero, Using ranked nodes to model qualitative judgments in Bayesian networks, *IEEE Trans. Knowl. Data. Eng.* 19 (10) (2007) 1420–1432.
- [26] L. Fu, X. Wang, D. Wang, M.A. Griffin, P. Li, Human and organizational factors within the public sectors for the prevention and control of epidemic, *Saf. Sci.* 131 (2020) 1–9, 104929.
- [27] J.G. Froese, A.R. Pearce, G. Hamilton, Rapid spatial risk modelling for management of early weed invasions: balancing ecological complexity and operational needs, *Method Eco. Evolution* 10 (12) (2019) 2105–2117.
- [28] R. Kaya, B. Yet, Building Bayesian networks based on DEMATEL for multiple criteria decision problems: a supplier selection case study, *Expert Sys. Appl.* 134 (2019) 234–248.
- [29] A. Freire, M. Perkusich, R. Saraiva, H. Almeida, A. Perkusich, A Bayesian networks-based approach to assess and improve the teamwork quality of agile teams, *Inf. Software Technol.* 100 (2018) 119–132.
- [30] P. Laitila, K. Virtanen, On theoretical principle and practical applicability of ranked nodes method for constructing conditional probability tables of Bayesian networks, *IEEE Trans. Sys. Cybern.: Sys.* 50 (5) (2020) 1943–1955.
- [31] L. Mkrtychan, L. Podofilini, V.N. Dang, Methods for building conditional probability tables of Bayesian belief networks from limited judgment: an evaluation for human reliability application, *Reliab. Eng. Sys. Saf.* 151 (2016) 93–112.
- [32] T. Noguchi, N. Fenton, M. Neil, Addressing the practical limitations of noisy-OR using conditional inter-causal anti-correlation with ranked nodes, *IEEE Trans. Knowl. Data. Eng.* 31 (4) (2019) 813–817.
- [33] P. Laitila, K. Virtanen, Improving construction of conditional probability tables for ranked nodes in Bayesian networks, *IEEE Trans. Knowl. Data. Eng.* 28 (7) (2016) 1691–1705.
- [34] P. Laitila, K. Virtanen, Elicitation Framework for Parameters of Ranked Nodes Method – Further Relief to Construction of Conditional Probability Tables of Bayesian Networks, 2021 submitted for publication.
- [35] M. Neil, M. Taylor, D. Marquez, Inference in hybrid Bayesian networks using dynamic discretization, *Stat. Comput.* 17 (3) (2007) 219–233.
- [36] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Mateo, CA, USA, 1988.
- [37] F. Diez, Parameter adjustment in bayes networks. the generalized noisy OR-gate, in: *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, Washington, D.C., USA, July 9–11, 1993, pp. 99–105.
- [38] S. Srinivas, A generalization of the noisy-OR model, in: *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, Washington, D.C., USA July 9–11, 1993, pp. 208–215.
- [39] B. Wisse, S.P. van Gosliga, N.P. van Elst, A.I. Barros, Relieving the elicitation burden of Bayesian belief networks, in: *Proceedings of the Sixth UAI Bayesian Modelling Applications Workshop*, BMAW 2008, Helsinki, Finland, 2008, pp. 10–20.
- [40] B. Das, Generating Conditional Probabilities for Bayesian Networks: Easing the Knowledge Acquisition Problem, 2004 arXiv:cs/0411034 [cs.AI].
- [41] J. Cain, *Planning Improvements in Natural Resources Management*, Centre for Ecology and Hydrology, Wallingford, UK, 2001.
- [42] L. Podofilini, L. Mkrtychan, V. Dang, Aggregating expert-elicited error probabilities to build hra models, in: *Safety and Reliability: Methodology and Applications*. Proceedings of the European Safety and Reliability Conference, ESREL 2014, Wroclaw Poland, 2014, pp. 1083–1091.
- [43] M.J. Barons, S. Mascaro, A.M. Hanea, Balancing the elicitation burden and the richness of expert input when quantifying discrete Bayesian networks, *Risk Anal.* (2021) 1–39, <https://doi.org/10.1111/risa.13772>.
- [44] W. Roed, A. Mosleh, J.E. Vinnem, T. Aven, On the use of the hybrid causal logic method in offshore risk analysis, *Reliab. Eng. Sys. Saf.* 94 (2) (2009) 445–455.
- [45] K.L. Hassall, G. Dailey, J. Zawadzka, A.E. Milne, J.A. Harris, R. Corstanje, A. P. Whitmore, Facilitating the elicitation of beliefs for use in Bayesian belief modelling, *Environ. Model. Software* 122 (2019) 1–9, <https://doi.org/10.1016/j.envsoft.2019.104539>.
- [46] E. Kemp-Benedict, Elicitation techniques for Bayesian network models, Rep. WP-US-0804, Stockholm Environment Institute, Sweden, 2008 <https://www.sei.org/publications/elicitation-techniques-bayesian-network-models>.
- [47] F. Hansson, S. Sjökvist, Modelling Expert Judgement into a Bayesian belief network. A Method for Consistent and Robust Determination of Conditional Probability Tables, Centre for Mathematical Sciences, Faculty of Engineering, Lund University, Sweden, 2013. Master’s thesis, <https://core.ac.uk/download/pdf/289953185.pdf>.
- [48] K.-S. Chin, D.-W. Tang, J.-B. Yang, S.Y. Wong, H. Wang, Assessing new product development project risk by Bayesian network with a systematic probability generation methodology, *Expert Sys. Appl.* 36 (6) (2009) 9879–9890.
- [49] Agena Ltd, AgenaRisk Software, Version 10.0, 2021 <http://www.agenarisk.com>.
- [50] BayesFusion, L.L.C., GeNIe Software, Version 3.0, 2021 <https://www.bayesfusion.com>.
- [51] Norsys Software Corp., Netica Application, Version 6.08, 2021 <https://www.norsys.com/index.html>.
- [52] Hugin Expert A/S, Hugin Software, Version 9.0, 2021 <http://www.hugin.com>.
- [53] E. Luedeling, L. Goehring, decisionSupport: Quantitative Support of Decision Making Under Uncertainty, Version 1.106, 2021 <https://CRAN.R-project.org/package=decisionSupport>.
- [54] K. Hassall, ACE: Application for Conditional Probability Elicitation, Version 1.0, 2021, <https://doi.org/10.5281/zenodo> <https://github.com/KirstyLHassall/ACE>.
- [55] H. Langseth, D. Marquez, M. Neil, Fast approximate inference in hybrid Bayesian networks using dynamic discretisation, in: *Proceedings Part I of the 5th International Work-Conference on the Interplay Between Natural and Artificial Computation*, Mallorca, Spain, 2013, pp. 225–234.
- [56] B. Yet, M. Neil, N. Fenton, A. Constantinou, E. Dementiev, An improved method for solving hybrid influence diagrams, *Int. J. Approx. Reason.* 95 (2018) 93–112.
- [57] P. Jorion, *Value at Risk*, McGraw-Hill, New York City, NY, USA, 2000.

Pekka Laitila is a Ph.D. candidate at Department of Mathematics and Systems Analysis, Aalto University, Finland. His research interests include probability elicitation to Bayesian networks and the effects of probability elicitation methods to decision recommendations of influence diagrams.

Kai Virtanen is Professor of Operations Research at Department of Mathematics and System Analysis, Aalto University, Finland and at Department of Military Technology, National Defence University of Finland. His research includes decision analysis, dynamic optimization, simulation optimization and human performance in complex systems.