



**HAL**  
open science

# Uncertainties in conditional probability tables of discrete Bayesian Belief Networks: A comprehensive review

Jeremy Rohmer

## ► To cite this version:

Jeremy Rohmer. Uncertainties in conditional probability tables of discrete Bayesian Belief Networks: A comprehensive review. *Engineering Applications of Artificial Intelligence*, 2020, 88, 10.1016/j.engappai.2019.103384 . hal-02386579

**HAL Id: hal-02386579**

**<https://brgm.hal.science/hal-02386579>**

Submitted on 29 Nov 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Uncertainties in conditional probability tables of discrete**  
2 **Bayesian Belief Networks: a comprehensive review**

3 **Jeremy Rohmer<sup>1</sup>**

4  
5 [1]{BRGM, 3 av. C. Guillemin - 45060 Orléans Cedex 2 - France}

6 Correspondence to: J. Rohmer (j.rohmer@brgm.fr)

7  
8 **Abstract**

9 Discrete Bayesian Belief Network (BBN) has become a popular method for the analysis of  
10 complex systems in various domains of application. One of its pillar is the specification of the  
11 parameters of the probabilistic dependence model (i.e. the cause-effect relation) represented via  
12 a Conditional Probability Table (CPT). Depending on the available data (observations, prior  
13 knowledge, expert-based information, etc.), CPTs can be populated in different manners, i.e.  
14 different assumptions can be made and different methods are available, which might lead to  
15 uncertain BBN-based results. Through an extensive review study of the past ten years, we aim  
16 at addressing three questions related to the CPT uncertainties. First, we show how to constrain  
17 these uncertainties either using elicitation of expert inputs, or using a combination of scarce  
18 data and expert-derived information. Second, we show how to integrate these uncertainties in  
19 the BBN-based analysis through propagation procedures either using probabilities or imprecise  
20 probabilities within the setting of credal or evidential networks. Finally, we show how to test  
21 the robustness of the BBN-based results to these uncertainties via sensitivity analysis  
22 specifically dedicated to BBNs. A special care was paid to describe the best practices for the  
23 implementation of the reviewed methods and the remaining gaps.

24  
25 **Keywords:** Bayesian Belief Network; Conditional Probability Table; Expert Elicitation;  
26 Sensitivity Analysis; Credal Network; Evidential Network.

27

## 28 **1 Introduction**

29 Bayesian Belief Network (BBN) has become an increasingly popular method for the analysis  
30 of complex systems in various domains of application, like ecosystems (Milns et al., 2010),  
31 genetics and biology (Scutari et al., 2014), agriculture (Drury et al., 2017), industry (Weber et  
32 al., 2012), finance forecasting (Malagrino et al., 2018), marine safety (Hänninen et al., 2014),  
33 human reliability assessment (Mkrtchyan et al., 2015), nuclear power plants (Kwag and Gupta,  
34 2017), aviation risk analysis (Brooker, 2011), coastal systems (Jäger et al., 2018), structure  
35 reliability assessments (Langseth and Portinale, 2007), multi-hazard risk assessments (Gehl and  
36 D’Ayala, 2016), etc.

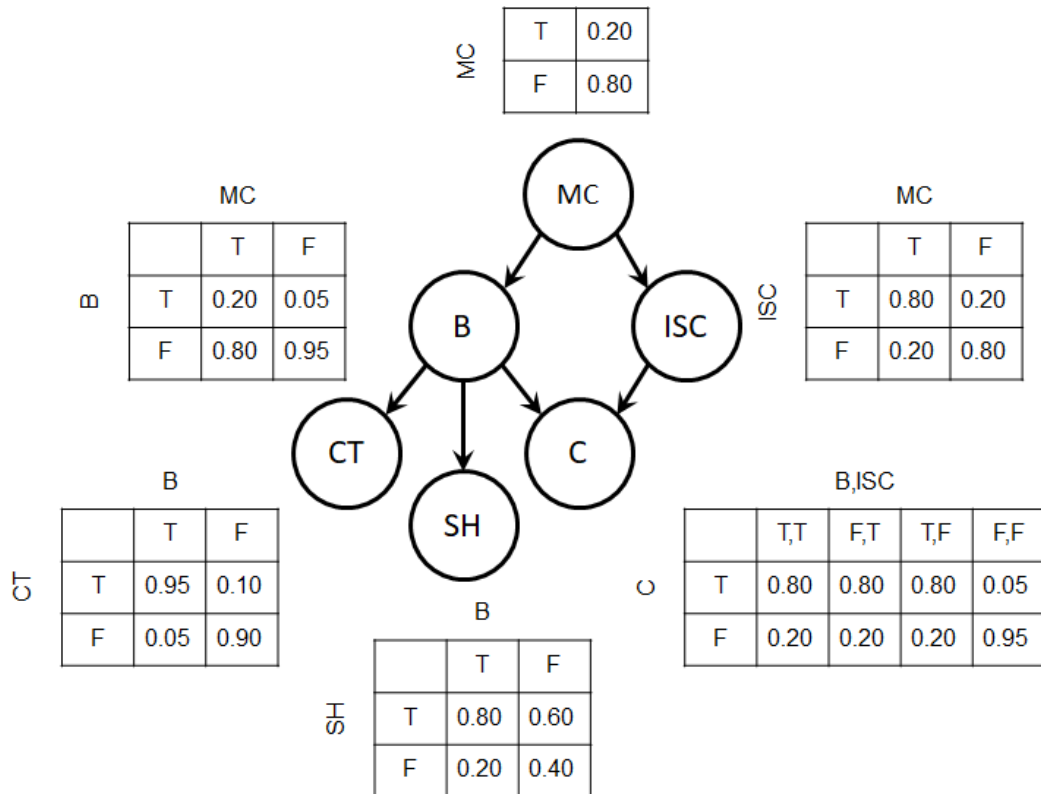
37 Its benefits are: (1) its high flexibility to model any causal relationships; (2) its capability to  
38 integrate information from any kind of sources, including experimental data, historical data,  
39 and prior expert opinion, and (3) its capability to answer probabilistic queries about them and  
40 to find out updated knowledge of the state of a subset of variables when other variables (i.e. the  
41 evidence variables) are observed.

42 Formally, a Bayesian belief Network (BBN) is a class of graphical model (see Jensen, 2001 for  
43 a complete and detailed introduction to BBNs), which allows to synthetically represent relations  
44 among random variables by means of a directed acyclic graph (DAG) composed of nodes (i.e.  
45 the states of the random variables) and arcs (i.e. dependency between nodes). The value of the  
46 nodes may be discrete or continuous, and we focus here on the former case, which is the most  
47 widely used. For instance, a Boolean node representing the state of a system component can be  
48 either “True” or “False”. The nodes connected by an arc are called the parent nodes and child  
49 nodes respectively. One child node may have several parent nodes, meaning that this node is  
50 affected by several factors. Similarly, a parent node could have several child nodes, meaning  
51 that this factor may have influences on several other factors. Conditional probabilities are the  
52 probabilities that reflect the degree of influence of the parent nodes on the child node. For BBNs  
53 with discrete nodes, the probabilistic dependence (i.e. the cause-effect relation) is often  
54 represented via a table called a Conditional Probability Table (CPT).

55 As an illustration, Fig. 1 depicts the binary BBN adapted by van der Gaag et al. (2013) from  
56 Cooper (1984) in the field of oncology. The network is composed of 6 nodes and 6 arcs. Node  
57 MC refers to metastatic cancer, which may potentially lead to the development of a brain tumor  
58 (node B) and may give rise to an increased level of serum calcium (node ISC). The presence of  
59 a brain tumour can be established from a CT scan (node CT). Another indicator of the presence  
60 of a brain tumour can be related to severe headaches (node SH). A brain tumour or an increased

61 level of serum calcium are both likely to cause a patient to fall into a coma (the node C is  
 62 connected to node B and node ISC). The conditional probabilistic relationships between the  
 63 nodes (CPT entries) are provided in Fig. 1 next to the corresponding nodes. For instance, the  
 64 probability that a patient falls into coma given brain tumor and increased level of serum calcium  
 65 corresponds to the first entry of the table (1<sup>st</sup> row, 1<sup>st</sup> column), namely  
 66  $P(C=True|C=True,ISC=True)=0.80$ .

67



68

69 Figure 1. Binary BBN adapted by van der Gaag et al. (2013) from Cooper (1984) in the field  
 70 of oncology. The tables (called CPT) next to the nodes provide the conditional probabilities  
 71 values.

72

[Figure 1 about here]

73

74 Two key ingredients are necessary to build a BBN, namely (1) the graph structure with the  
 75 direction of the arcs, i.e. the DAG; (2) the states of nodes and the strength of the relationships  
 76 between nodes, i.e. the CPT. In the present study, we assume that the DAG model has already  
 77 been determined and restrict the analysis to the quantification of the BBN relationships. The  
 78 process of deriving the CPTs and its associated uncertainties is recognized in the literature as  
 79 one of the most delicate part of the BBN development (e.g., Chen and Pollino, 2012; Druzdzal

80 and van der Gaag, 2000; Marcot et al., 2006; Cain, 2001, etc.). It should, however, be noted  
81 that the process of DAG derivation (i.e. building the graph structure plus the directions; also  
82 known as causal structure learning) has its own challenges as well, in particular when the  
83 learning is based on data (see e.g., a comprehensive review by Heinze-Deml et al., 2018).

84 Depending on the available data (observations, prior knowledge, expert-based information,  
85 etc.), CPTs can be evaluated in different manners, i.e. different assumptions can be made and  
86 different methods are available leading to different BBN-based results, hence resulting in  
87 uncertain BBN-based results. This raises the following questions: (1) how to constrain the  
88 uncertainties related to CPT derivation, i.e. what are the methods that are available to minimize  
89 these uncertainties? (2) how to integrate these uncertainties in the BBN-based analysis, i.e. what  
90 are the methods for propagating these uncertainties? (3) how to test the robustness of the BBN-  
91 based results to these uncertainties, i.e. what are the methods for identifying the most influential  
92 uncertainties? These questions are addressed below through an extensive review of studies  
93 performed in the past ten years by focusing on discrete BBNs that can be used for modelling  
94 complex causal relationships, for merging different information sources, for prediction, and for  
95 belief/evidence propagation (i.e. probabilistic queries). Continuous BBNs (i.e. BBNs with  
96 continuous nodes) and dynamic BBNs (i.e. BBNs adapted to model systems evolving over  
97 time) are out of the scope of the review.

98 The paper is organized as follows. The first section describes more specifically the problem of  
99 populating the CPT parameters directly from data/observations. This first part highlights the  
100 necessity for overcoming the lack of data by complementing with additional sources of  
101 information. Sect. 3 explores an expert-based option for constraining the uncertainties related  
102 to data incompleteness, namely by completing with different expert-based sources of  
103 information. Sect. 4 provides an overview of the different approaches embedded in different  
104 uncertainty analysis settings for evaluating the impacts of CPT uncertainties, either using  
105 probabilities or imprecise probabilities. Sect. 5 further addresses the problem of screening these  
106 uncertainties by describing sensitivity analysis techniques. Finally, Sect. 6 summarizes the main  
107 findings and discusses the open questions.

## 108 **2 Learning CPT from data**

109 In this section, we address the issues related to deriving the CPT entries from data. Sect. 2.1  
110 first discusses the problem of performing this task by using only data. Sect 2.2 and 2.3 further  
111 discuss two practical difficulties, namely: (1) the presence of missing values and (2) the  
112 problem of translating observations related to continuous variables into a limited number of

113 discrete states. Finally, Sect. 2.4 describes methods that make the most out of scarce data while  
114 exploiting qualitative information provided by experts.

## 115 **2.1 A pure data-driven approach**

116 Let us consider a BBN composed of  $n$  discrete nodes  $X_{i=1,\dots,n}$ . Let us denote  $r_i$  the cardinality of  
117  $X_i$  and  $q_i$  the one of the parent set of  $X_i$ , denoted  $pa(X_i)$ . The  $k^{\text{th}}$  probability value of the  
118 conditional probability distribution is  $\theta_{ijk} = P(X_i = k | pa(X_i) = j)$  where  $i=1,\dots,n$ ;  $j=1,\dots,q_i$ ;  
119  $k=1,\dots,r_i$ .

120 In data rich contexts, CPT parameters can be evaluated by computing the appropriate  
121 frequencies from data. An example is provided by Chojnacki et al. (2019) for fire safety analysis  
122 where more than 1 million of numerical simulation results are used. This method corresponds  
123 to the maximum likelihood estimation (MLE), which is described below.

124 Let us consider a dataset  $D$  where a total number  $N_{ij}$  of data records are available for which  
125  $pa(X_i)$  is in the state  $j$  and where  $N_{ijk}$  data records are available for which  $X_i$  is in the state  $k$  and  
126  $pa(X_i)$  is in the state  $j$ . MLE aims at maximizing the log-likelihood function  $l(\cdot)$  of  $\theta$  given  $D$  as  
127 follows

$$128 \quad l(\theta|D) = \log(P(D|\theta)) = \sum_{ijk} N_{ijk} \log(\theta_{ijk}) \quad (\text{Eq. 1})$$

129 The solution is then  $\frac{N_{ijk}}{N_{ij}}$ .

130 The MLE method however fails to find good estimates due to data scarcity when  $N_{ij} \approx 0$ , i.e.  
131 when training data are not sufficient in number in some specific variable state configurations.  
132 Examples of such contexts are not rare in practice; see e.g. rare disease diagnostic (Seixas et  
133 al., 2014), accident prevention (e.g., Hänninen, 2014), reliability analysis (e.g., Musharraf et  
134 al., 2014), etc. This problem is even worsened when the number of nodes increases. Recall that  
135 the number of conditional probabilities is exponential with the number of its parent nodes, i.e.  
136 for a node with  $i$  states and  $k$  parent nodes and if each parent node has  $n$  states,  $(i-1) \times n^k$  CPT  
137 entry values have to be specified. For instance, a binary node with 2 binary parent nodes  
138 imposes to specify 4 entries, whereas for a ternary node with 2 ternary nodes, this number  
139 reaches 18.

## 140 **2.2 Dealing with missing values**

141 The process for parameter learning of discrete BBNs may be complicated in the presence of  
142 missing values. This can be handled by means of different algorithms. The most popular ones  
143 are Expectation Maximization (Dempster et al., 1977) and Gibbs sampling (Geman and Geman,  
144 1984). Yet, they both assume that the values are missing at random. This hypothesis may not  
145 always be true in practice. Alternative methods have been proposed to overcome this  
146 disadvantage, like AI&M procedure (Jaeger, 2006), the RBE algorithm (Ramoni and  
147 Sebastiani, 2001), and the maximum entropy method (Cowell, 1999). Other methods have also  
148 been developed to speed up the learning process, like generalized conjugate gradient algorithm  
149 by Thiesson (1995) or the online updating of rules (Bauer et al., 1997). To deal with both  
150 missing data and qualitative influences (as described in Sect. 2.4), some initiatives have been  
151 proposed like the one of Masegosa et al. (2016), who further improved the combined Isotonic  
152 Regression - EM approach.

## 153 **2.3 Discretising continuous variables**

154 A second practical difficulty for parameter learning of discrete BBNs is inherent to the main  
155 assumption introduced by discrete BBNs, namely that data should be represented by a limited  
156 numbers of outcomes. This imposes to discretize continuous variables. This process might,  
157 however, lead to a loss of information, and potentially to an increase of the associated  
158 computational effort, because the size of discrete BBNs increases approximately exponentially  
159 with the number of discrete states of its nodes. Nojavan et al. (2017) investigated the  
160 implications of several mathematical methods for constructing discrete distributions in an  
161 unsupervised manner. Using a simple 3-node BBN describing chlorophyll concentrations in  
162 Finnish lakes, the authors evaluated the impact on the developed BBNs of the number of  
163 intervals and of the choice of the type of discretization methods. Three techniques were  
164 investigated, namely in which the data are divided into groups: (1) of equal length; (2) of equal  
165 sample size; (3) for which the moments of the discretized distribution match with the moments  
166 of the continuous data. They showed that none of the models did uniformly well in all  
167 comparison criteria (sum of squared errors, accuracy, area under the receiving operating  
168 characteristic curve) for the considered case. They concluded that they cannot justify using one  
169 discretization method against others. Using a 4-node BBN from the domain of coastal erosion,  
170 Beuzen et al. (2018) extended the tests to other types of discretization methods, namely manual  
171 and supervised techniques. They showed, on their specific test case, that supervised methods

172 led to a BBN of the highest average predictive skill, followed by the one with manual  
 173 discretization. They also outlined the advantages of the different methods, namely that:

- 174 - Manual methods allow ensuring physical meaningful BBNs;
- 175 - Supervised methods can autonomously and optimally discretize variables and may be  
 176 preferred when predictive skill is a modelling priority;
- 177 - Unsupervised methods are computationally simple and versatile.

178 Depending on the objective, some specific discretization algorithms have also been developed;  
 179 for instance, Zwirgmaier and Straub (2016) developed specific methods to deal with rare  
 180 events in reliability analysis; Neil et al. (2007) proposed a dynamic discretization method to  
 181 perform inference in hybrid BBNs, i.e. both dealing with continuous and discrete variables.

## 182 **2.4 Combining scarce data and expert judgements**

183 When data are scarce, the parameter learning may be improved by incorporating additional  
 184 information provided by experts. A popular approach relies on the Maximum a Posteriori  
 185 (MAP) estimation using Dirichlet priors, which express experts' belief (e.g., Heckerman et al.,  
 186 1995) about  $\theta$  in the absence of data. Formally, the Dirichlet distribution for CPT column  $\theta_j$  is  
 187 expressed as follows:

188

$$189 \quad p(\theta_{ij}) = \frac{1}{Z_{ij}} \prod_{k=1}^{r_i} \theta_{ijk}^{(\alpha_{ijk}+1)-1} \quad (\text{Eq. 2})$$

190 with  $\sum_k \theta_{ijk} = 1$ ,  $\theta_{ijk} \geq 0$ ,  $Z_{ij}$  is a normalisation term  $\int_{-\infty}^{+\infty} \prod_{k=1}^{r_i} \theta_{ijk}^{(\alpha_{ijk}+1)-1} d\theta_{ijk} = 1$ , and  
 191  $\alpha_{ijk}$  is the parameter of the Dirichlet distribution, which can be intuitively interpreted as “how  
 192 many times the expert believes he/she will observe  $X_i=k$  in a sample of  $\alpha_{ij}$  instances drawn  
 193 independently at random from the distribution  $\theta_j$ ” (Zhou et al., 2014). On this basis, MAP relies  
 194 on the following equation:

195

$$196 \quad p(\theta|D) \propto P(D|\theta)P(\theta) \propto \prod_{ijk} \theta_{ijk}^{(\alpha_{ijk}+N_{ijk})-1} \quad (\text{Eq.3})$$

197 This equation results in the estimate of  $\theta_{ijk}$  as  $\frac{N_{ijk}+\alpha_{ijk}-1}{N_{ij}+\alpha_{ij}-1}$ , which combines information from  
 198 the data and from the experts' prior guess. In their computer experiments using twelve publicly



199 available BBNs (available at <http://www.bnlearn.com/bnrepository/>), Zhou et al. (2016a)  
200 showed that MAP achieves better performances than conventional MLE, which suffers from  
201 the absence of data in several state configurations in situations of limited sample size (typically  
202 50).

203 Expert-based information can take several forms, and the one that corresponds to qualitative  
204 constraints have given rise to several developments. Instead of directly providing the exact  
205 value of the entries of binary BBN (denoted  $P_{1,2}$ ), the expert may feel more conformable in  
206 providing an ordering like “ $P_1 > P_2$ ”, “ $P_1 \approx P_2$ ”, “ $P_1 > 0.80$ ”, etc. Zhou et al. (2016a) showed that  
207 incorporating such expert knowledge about the monotonic influences between nodes (translated  
208 into probability constraints) further outperformed MAP and MLE and was also robust to errors  
209 in labelling the monotonic influences.

210 Different methods have been developed to incorporate qualitative constraints, namely:

- 211 - Convex Optimization (Niculescu et al., 2006; Zhou et al., 2016a; de Campos and Ji,  
212 2008; Liao and Ji, 2009; Altendorf et al., 2005) is an extension of the MLE by  
213 incorporating constraints via penalty functions or by restricting parameter spaces;
- 214 - Constrained MAP approach has also been proposed by Yang et al. (2019) to learn BN  
215 parameters by incorporating convex constraints;
- 216 - Isotonic Regression (Feelders and van der Gaag, 2005; 2006) builds on qualitative  
217 information about the influences between the variables of a BBN. The most recent  
218 algorithm by Masegosa et al. (2016) also enables the analyst to learn the CPT  
219 parameters from incomplete data;
- 220 - Qualitative MAP (originally proposed by Chang and Wang (2010) and further improved  
221 by Guo et al. (2017)) constructs Dirichlet priors from Monte-Carlo random samples of  
222 the constrained parameter space, which are used by the MAP algorithm;
- 223 - Multinomial Parameter Learning with Constraints (Zhou et al., 2014; Hospedales et al.,  
224 2015) rely on auxiliary BBNs, which are hybrid BBNs, to infer the posterior distribution  
225 of BBN parameters.

## 226 **2.5 Discussion**

227 Following a pure statistical data-driven approach for populating the BBN conditional model  
228 requires a large amount of statistically significant data to cover all BBN relationships. To  
229 compensate the lack of data, a possible option is to complement the analysis with expert-based  
230 information. Sect. 2.4 shows that a broad range of different tools/methods are available to

231 incorporate expert-based information either in the form of qualitative influences or constraints,  
232 namely constraints that should be almost linear and convex (i.e. concave constraints like  
233  $P_1 \neq 0.5$  cannot be accounted for). The improvement of the learning accuracy of the parameters  
234 in BBNs from a small data set has been shown using each of the described methods compared  
235 to conventional methods; for instance Guo et al. (2017) compared MLE, constrained MLE,  
236 maximum entropy and constrained maximum entropy estimator, MAP and their qualitatively  
237 MAP estimator. Yang et al. (2019) showed the higher performance of their constrained MAP  
238 estimator compared to conventional parameter learning algorithms, MLE and MAP, and to  
239 constrained maximum likelihood algorithm. Yet, to the author's best knowledge, no extensive  
240 benchmark exercise covering all the afore-mentioned estimators (as well as their pros and cons)  
241 is available yet; practical recommendations on how to implement them and their limitations is  
242 currently lacking in the literature.

243 Among the possible limitations, the problem of under-fitting related to the use of prior  
244 distributions (that are common ingredients of most of the methods of Sect. 2.4) is seldom  
245 tackled. As described by Gao et al. (2019), imposing certain a priori knowledge on the CPT  
246 parameters might decrease the likelihood of the parameters, hence a reduction of the fitness  
247 between parameters and data. Azzimonti et al. (2019) proposed a hierarchical procedure to  
248 improve the widely-used approach based on Dirichlet priors. Gao et al. (2019) proposed a  
249 Minimax Fitness algorithm combined with an improved constrained maximum entropy method  
250 to overcome this problem. They also concluded that there is a need for further investigation to  
251 develop learning methods that does not require specification of prior strength.

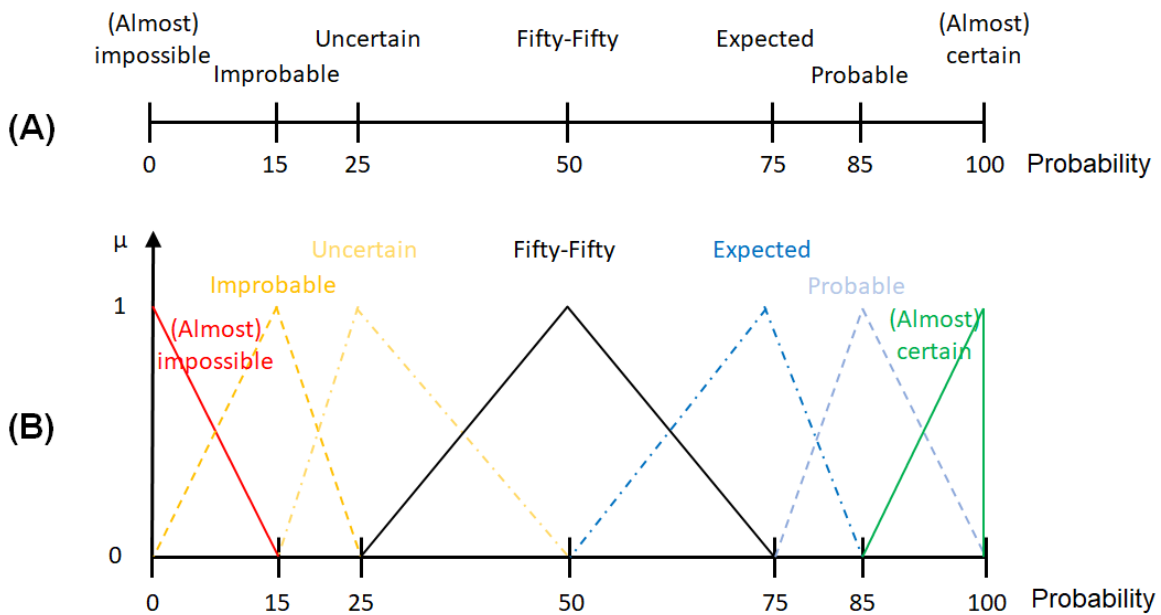
### 252 **3 Learning from experts**

253 In many situation, the primary source of information for learning the CPTs is not based on data,  
254 but on inputs from expert domain. For instance, for rare-event situations like reliability analysis,  
255 inputs from expert domain stem from questionnaires, interviews and panel discussions. Sect. 3  
256 focuses on the process of deriving information from experts that is named "elicitation". The  
257 issues and methods related to this task were analysed by review articles in different domains of  
258 application, namely shipping accidents by Zhang and Thai (2016), human reliability by  
259 Mkrtchyan et al. (2015) and more broadly regarding dependence in probabilistic modelling by  
260 Werner et al. (2017). The objective is to focus the elicitation on specific pieces of information  
261 to efficiently populate the CPTs by ensuring quality and consistency of the elicited result and  
262 minimizing the workload on the experts owing to the large number of CPT entries. Elicitation  
263 for CPT generally relies on three (possibly combined) main approaches through: (1) the

264 assessment of probabilities directly from an (or a panel of) expert (Sect. 3.1); (2) assumptions  
 265 on the causal structure either by simplifying the network structure or by simplifying the causal  
 266 dependence (Sect. 3.2); (3) filling-up methods (Sect. 3.3).

267 **3.1 Direct elicitation**

268 In a direct approach, experts are asked to give quantitative numbers (like frequencies or  
 269 confidence intervals) using methods like probability wheel, probability scale and gambling  
 270 analogy. An extensive discussion on the different types of biases are provided by Renooij  
 271 (2001), and more specifically in the domain of ecology by Kuhnert et al. (2010). Overall,  
 272 methods which map qualitative statements to numerical values like the probability scale (see  
 273 an example in Fig. 2(A)) is preferred for its simplicity, which improves the consistency (as  
 274 underlined by Wiegmann (2005), and as reported by Zhang and Thai (2016) for marine safety).  
 275 Probability wheel is criticized for not being appropriate for the elicitation of small or large  
 276 probabilities, and the gambling analogy is criticized for being too time-consuming.



277  
 278 Figure 2. (A) Example of probability scale used to assist expert elicitation of CPTs (adapted  
 279 from Knochenhauer et al. (2013)); (B) Translation of the probabilities qualified in (A) into  
 280 Fuzzy sets ( $\mu$  is the degree of membership).

281 [Figure 2 about here]

282  
 283 As an alternative, experts are preferably asked to give qualitative statements (like categorical  
 284 or relative measure). To support this indirect approach, tools from the domain of multicriteria

285 decision-making have been proposed. For instance, Chin et al. (2009) adapted the Analytical  
286 Hierarchy Process method for the task of probability elicitation and semi-automatic generation  
287 of the parameters of CPTs. The basic idea is to elicit paired comparisons about the relative  
288 likelihood of the possible events using predefined scores (equally possible, etc.) instead of  
289 directly asking the probability values. Yet, this procedure is at the expense of an increase in the  
290 number of comparisons as the number of conditional probabilities increases.

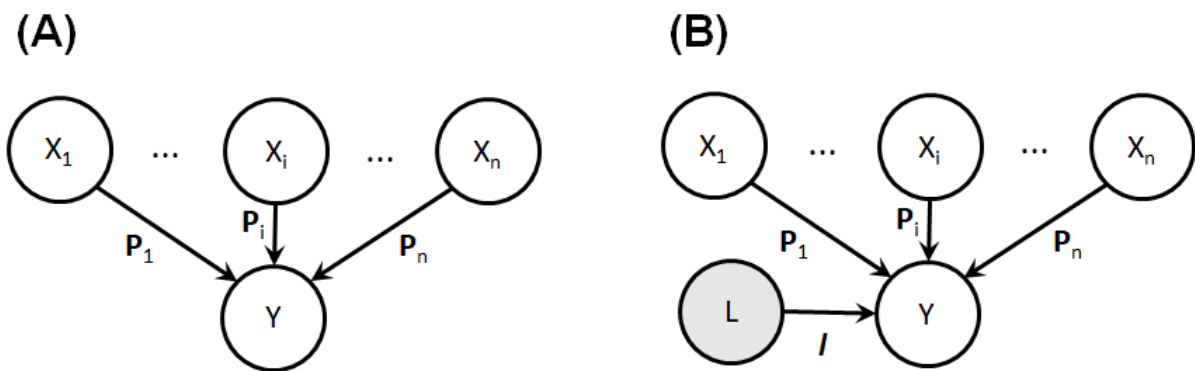
291 An alternative option proposes to directly process natural linguistic terms by mathematically  
292 modelling them using for instance a Fuzzy set (Zadeh, 1975). Let us consider the concept of  
293 membership function, which defines how each element  $x$  of the input space  $X$  (also named  
294 “universe of discourse”) is mapped to a degree of membership (denoted  $\mu$ ). Under the classical  
295 theory of Boolean logic, the membership function of a set  $A$  is simply defined as a binary  
296 function that takes the value  $\mu(x)=1$  if the element belongs to  $A$  and the value  $\mu(x)=0$ , otherwise.  
297 The Fuzzy set theory of Zadeh (1965) introduces the concept of a set without a crisp (i.e. clearly  
298 defined) boundary. Such a set can contain elements with only a gradual (partial) degree of  
299 membership ( $\mu$  is scaled between 0 and 1). The translation of the probability scale of Fig. 2(A)  
300 into Fuzzy sets is provided in Fig. 2(B). Some successful applications cover fault detection  
301 (D’Angelo et al., 2014), performance analysis of devices (Penz et al., 2012), safety risk analysis  
302 (Zhang et al., 2015), human reliability analysis (~~Peng-cheng et al., 2012~~; Li et al., 2012), and  
303 offshore risk (Ren et al., 2009). Two viewpoints exist in the literature on Fuzzy BBNs.  
304 Fuzziness can be incorporated in the variables (nodes) or on the probabilities. For instance, Ren  
305 et al. (2009) carried out studies using fuzzy probability calculations in BBNs (as illustrated in  
306 Fig. 2(B)). Conversely, Tang and Liu (2007) used fuzzy events (i.e. Fuzzy node states) in BBNs  
307 for a machinery fault diagnosis problem. İçen and Ersel (2019) incorporated both aspects with  
308 application in medicine.

### 309 **3.2 Making assumptions on the causal structure**

310 To reduce the elicitation burden, the number of CPT entries to be elicited should be kept  
311 “reasonable”. This can be performed by making assumptions regarding the causal structure.  
312 One option is by simplifying the structure through the introduction of “divorcing” nodes  
313 (Henderson et al., 2009). This involves aggregating a few of the nodes by adding a new node  
314 that summarizes them provided that the aggregations are logical and no interactions are lost in  
315 the procedure. Although this process adds nodes to the network, it reduces the combined size  
316 of CPTs in the network (Cain, 2001). Yet, divorcing might dilute the sensitivity of the final

317 node(s) to the input nodes and might increase the uncertainty propagated through the network  
 318 as underlined by Cain (2001).

319 A popular alternative aims at making some simplifications regarding the causal dependence  
 320 based on the logical Noisy-OR gate (Pearl, 1988). In their typical implementation, Noisy-OR  
 321 gates focus on binary BBN nodes and assume that the influence of the considered factor is  
 322 independent from the presence of the other factors. This means that the probability of the  
 323 outcome is the product of the probabilities of the outcome in presence of one factor at a time,  
 324 with all other factors being absent. Formally, let us consider a binary variable  $Y$  with two states  
 325 {False, True} and  $n$  binary parent variables  $X_{i=1,\dots,n}$ .



326

327 Figure 3. (A) Schematic representation of the Noisy-OR gate with  $P_{i=1,\dots,n}$  the link  
 328 probabilities; (B) Schematic representation of the Noisy-OR gate.

329

[Figure 3 about here]

330

331 The main principle of the Noisy-OR model is to define probabilities  $P_i$  (termed as link  
 332 probability, Fig. 3(A)), which are defined as the probability that  $Y$  is False given that  $X_i$  is False  
 333 and  $X_j$  is True for  $i \neq j$ . A Noisy-OR model is thus a disjunction “noisy” version of  $X_i$  (Pearl,  
 334 1988). This means that the distribution of  $Y$  conditional on  $X_1; X_2; \dots; X_n$  is  $P(Y =$   
 335  $F|X_1; \dots; X_n) = 1 - \prod_{i: X_i \in X_T} (1 - P_i)$  where  $X_T$  is the set of parent nodes whose states are True.  
 336 The Noisy-OR model enables the analyst to specify fewer CPT parameters; the number of  
 337 independent parameters being here reduced from  $2^n$  to  $2n$ . The extension of Noisy-OR gate to  
 338 multi-valued variables is the Noisy-MAX gate model (Diez, 1993; Henrion, 1989). If the parent  
 339 node  $X_i$  has  $n_{X_i}$  states, then the total number of parameters that have to be elicited using leaky  
 340 Noisy-MAX gate is  $N = \sum_{i=1}^n (n_{X_i} - 1) (n_Y - 1) + 1$  to be compared to the total number  
 341 without Noisy-MAX gate, namely  $N = (n_Y - 1) \cdot \prod_{i=1}^n n_{X_i}$ .

342 Different empirical studies have been conducted to investigate the performance of the leaky  
343 Noisy-OR approach. Several authors (Oniško et al., 2001; Anand and Downs, 2008; Bolt et al.,  
344 2010; among others) showed how this approach helped reducing the burden of elicitation in  
345 practical real-life applications without impacting too much the performance of the network.  
346 Besides, Zagorecki and Druzdzel (2012) explored to which extend the pattern of causal  
347 interaction induced by Noisy-OR(MAX) gates are common in real cases. Using three existing  
348 BBNs, they showed that the Noisy-MAX gate provides a good fit for as many as 50% of CPTs  
349 in two of these networks.

350 The Noisy-OR structure is based however on a strong assumption, i.e. that the node of interest  
351 is in the state False (considering the above illustrative case) with a probability equal to 1 if all  
352 its parent variables are in the state False. Yet, in many cases, it is often difficult to capture all  
353 the causes of the node of interest (e.g. for reliability purpose, it means to define all the failure  
354 modes of a component). To deal with this problem, Henrion (1989) proposed an extension  
355 called “leaky Noisy-OR” gate that includes a background probability that represents the  
356 influence of non-modelled causes as schematically depicted in Fig. 3(B). Zagorecki and  
357 Druzdzel (2004) proposed to elicit leaky and non-leaky Noisy-OR parameters as alternatives to  
358 conditional probabilities using statements like “What is the probability that  $Y$  is present when  
359  $X_1$  is present and all other causes of  $Y$  (including those not modelled explicitly) are absent?”.  
360 They showed that the leaky Noisy-OR parameter was assessed as the most accurate (in terms  
361 of Euclidean distance to empirical distribution).

362 The leaky Noisy-OR method was further extended by relaxing the necessity to define a crisp  
363 precise leaky probability value, i.e. by introducing uncertainty on this parameter. This type of  
364 uncertainty has been addressed within different uncertainty treatment settings (which are  
365 introduced in more details in Sect. 4). Antonucci (2011) developed an imprecise leaky Noisy-  
366 OR gate model with uncertainty on the link probabilities modelled by intervals within the  
367 formalism of credal networks (see Sect. 4.2). Alternatively, Fallet-Fidry et al. (2012) (further  
368 extended by Zhou et al. (2016b)) proposed an imprecise extensions of the Noisy-OR within the  
369 formalism of evidential networks (see Sect. 4.3). Finally, Dubois et al. (2017) developed a  
370 version of noisy logical gates within the theory of possibility (Dubois and Prade, 1988) using  
371 possibilistic causal networks (as presented by Benferhat et al. (2002)) with illustration on an  
372 example taken from human geography.

### 373 **3.3 Filling-up methods**

374 Alternative methods to Noisy-OR(MAX) gate are based on filling-up techniques. These  
375 methods are typically based on extracting information on the factor effects from known  
376 relationships (named anchor conditional probability distributions, denoted CPD) and  
377 extrapolating to the whole CPTs. Considering two BBNs (of respectively 3 and 4 nodes) for a  
378 human reliability problem, Mkrtchyan et al. (2015) tested five popular methods for CPT  
379 derivation considering nodes with multiple states, namely:

- 380 - Method 1: the functional interpolation method (Podofillini et al., 2014) approximate  
381 CPDs elicited at the anchor positions by functions described by parameters (e.g.,  
382 Normal functions); the parameters of the missing CPDs are then obtained by  
383 interpolating those corresponding to the anchor ones;
- 384 - Method 2: the Elicitation BBN method (Wisse et al., 2008) is based on piecewise linear  
385 functions interpolating among the elicited CPDs, and on state influencing factors and  
386 importance weights;
- 387 - Method 3: The Cain calculator (Cain, 2001) uses interpolation factors derived from  
388 CPDs to populate the missing relationships in CPTs;
- 389 - Method 4: The method presented by Røed et al. (2009) is also based on functional  
390 relationships between influencing factors and outcome nodes; the parameters of the  
391 function (exponential) are then determined based on the elicitation of selected CPDs;
- 392 - Method 5: the ranked node method by Fenton et al. (2007) (further improved by Laitila  
393 and Virtanen, 2016) is not based on interpolation of known CPDs. In this approach, all  
394 the nodes are defined on the interval [0–1]. For instance, let us consider a node with 5  
395 states, namely “very low”, “low”, “average”, “high”, and “very high”; each of the  
396 state is assigned to an interval width of 0.2; for instance, the value “low” is assigned to  
397 the interval [0.2–0.4]. To generate CPTs, the experts are asked to provide the weight  
398 parameters and to choose one algorithm (the mean average, the Minimum, the  
399 Maximum and the MixMinMax). Using this method, if there are  $m$  ranked nodes and  
400 each node has  $n$  states, the expert will only need  $m + 1$  parameter values, while it requires  
401  $n \times m + 1$  values for full elicitation.

402 Mkrtchyan et al. (2015) showed that:

- 403 - All methods allow representing the different importance of the various influencing  
404 factors;

- 405 - The representation of the interactions (combined effects of multiple factors) is  
406 problematic for methods eliciting information on the influence of factors taken one at a  
407 time (methods 2-4);
- 408 - Functional representation of the CPTs (methods 1, 5 and 4) can be traced more easily,  
409 because they allow an explicit representation of uncertainty in the factor relationships;
- 410 - But methods 4 and 5 have difficulties in representing the different degrees of uncertainty  
411 in the relationships;
- 412 - The method allowing the largest modelling flexibility is method 1 with respect to strong  
413 factor influences (single and multi-factor) and proper uncertainty characterization, but  
414 becomes too costly for large BBNs.

### 415 **3.4 Discussion**

416 The conclusions drawn by Mkrtchyan et al. (2015) serve as valuable recommendations  
417 regarding the use and applicability of the five most popular filling-up methods for reducing the  
418 expert burden of CPT elicitation. Despite the practical usefulness of this comparative exercise,  
419 it should be noted that they primarily focused, by construction, on the modelling aspects  
420 important for their application domain (here human reliability analysis), namely the  
421 representation of strong factor influences and interactions, and the characterization of different  
422 degrees of uncertainty in the relationships. Broader exercises are needed to cover a larger  
423 spectrum of methods (i.e. filling-up methods should be completed by Noisy-OR/MAX models,  
424 direct elicitation among others), of contexts (different network sizes, binary versus multivalued  
425 nodes, etc.), as well as of domains of application.

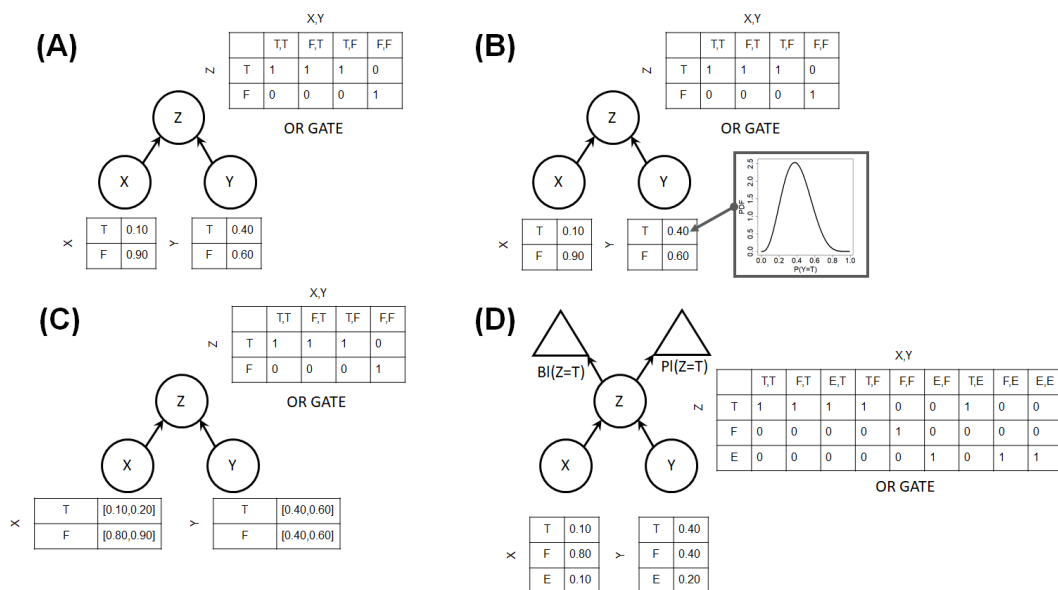
426 Despite the clear advantages of these methods for BBN engineering, they cannot be applied  
427 uncritically, because the probability values can only be considered approximations of the true  
428 probabilities and whatever the considered methods, they are all based on simplifications that  
429 may hamper the BBN performance. Initiatives like the one by Woudenberg and van der Gaag  
430 (2015) for the Noisy-OR model should be intensified. They identified the conditions under  
431 which ill-considered use of this method can result in large impact on output probabilities; in  
432 particular, when the yet-unobserved cause variables in the mechanism have relatively skewed  
433 probability distributions and/or the obtained parameter probabilities have small values. For this  
434 purpose, sensitivity methods as described in Sect. 5 can play an important role. Fenton et al.  
435 (2019) also dealt with the limitations of leaky Noisy-OR model for backward inference. When  
436 the binary node of interest Y of the example in Fig. 3 is observed to be in the state False, the  
437 normal “explaining away” behaviour fails, which means that after observing the state of any



438 parent the remaining parents become independent, and the results may not result in what BBN  
 439 practitioners expected. Fenton et al. (2019) described a simple extension of the model that  
 440 requires the elicitation of only one extra parameter that can solve this problem for a large  
 441 spectrum of cases in practice.

#### 442 4 Propagating the uncertainties

443 Whatever the methods used to populate the CPTs, residual uncertainties may still prevail. This  
 444 residual uncertainty should be reflected in BBN-based results. This means that the uncertainty  
 445 on CPT entries should be propagated in order to evaluate their consequences on the BBN  
 446 results. The propagation can either rely on probabilities (Sect. 4.1), or alternative mathematical  
 447 representation tools like intervals (Sect. 4.2) or a generalization of a probability distribution  
 448 (Sect. 4.3), i.e. within the theory of belief functions as introduced by Shafer (1976) and  
 449 Dempster (1967). Fig. 4 summarizes the main principles of the different approaches by using a  
 450 simple OR-gate model.



451  
 452 Figure 4. (A) Example of an OR gate model translated into a BBN with two binary parent nodes  
 453  $X$  and  $Y$  (with states corresponding to  $T=$ True or  $F=$ False). The truth table related to the OR  
 454 gate corresponds to the table next to the child node  $Z$ . Illustration of a probability-based  
 455 approach where uncertainties on CPT entries are represented by: (B) Beta probability  
 456 distributions (with an example here for node  $Y$ ); (C) Interval-valued probabilities (credal  
 457 network approach); (D) Mass probability tables; here the truth table includes the epistemic state  
 458  $E=\{T,F\}$ ; Two nodes were added to the network to calculate the belief and plausibility functions  
 459 (see Sect. 4.3 for more details).

460

[Figure. 4 about here]

#### 461 **4.1 Methods using probabilities**

462 The problem of uncertainty propagation for BBN has originally been addressed using  
463 probabilities. This approach assumes that the uncertainty on CPT entries follows a Beta  
464 probability distribution (or for more generic cases, a Dirichlet probability distribution), as  
465 schematically depicted in Fig. 4(B). Kleiter (1996) originally described a Monte-Carlo-based  
466 random simulation procedure to carry out the approximation of the spread of the probability  
467 distribution for the considered query. The method requires, however, a large number of random  
468 samples to accurately characterize the true variance. Van Allen et al. (2008) proposed an  
469 improved method by avoiding Monte Carlo sampling through the combination of bucket  
470 elimination (Dechter, 1998) with the “delta rule” that linearizes the relationship between the  
471 query probabilities and the corresponding Dirichlet conditional probabilities connecting the  
472 query variable to its parents and children. They further proved that the Beta approximation (for  
473 binary BBNs) is asymptotically valid. The conditions of the exact Beta distribution has  
474 extensively been investigated by Hooper (2008). This problem has further been formalized  
475 within the setting of subjective logic (Jøsang, 2001; 2016) as proposed by Kaplan and Ivasnoska  
476 (2018), who developed an efficient belief propagation for inference in a binary Bayesian  
477 network with a singly-connected graph. To introduce any type of probability distribution on  
478 CPTs, Fenton (2018) proposed to extend the BBN with continuous nodes corresponding to the  
479 uncertain prior probability distributions, but at the expense of a potentially large increase of  
480 computational time cost when the number of nodes and of CPTs increases.

#### 481 **4.2 Methods using interval-valued probabilities**

482 Instead of specifying a crisp single value of each CPT entry, the formal setting of credal  
483 network, denoted CN (Cozman, 2000; 2005), integrates BBNs with credal sets, i.e. set of  
484 probability measures. A CN can be viewed as the representation of a set of BBNs, which share  
485 the same graphical structure but are associated to different conditional probability parameters;  
486 the interest being to provide a richer representation of uncertainty. In Fig. 4(C), the uncertainty  
487 in the CPTs are presented by intervals.

488 Formally, given a variable  $X$ , we denote by  $\Omega$ , the possibility space of  $X$ ,  $x$  a generic element  
489 of  $\Omega$ ,  $P(X)$  the probability mass function for  $X$  and  $P(x)$  the probability of  $x$ . The credal set over  
490  $X$  is  $K(X)$ , which corresponds to a closed convex set of probability mass functions over  $X$ . For  
491 any  $x \in \Omega$ , the lower probability for  $x$  according to the credal set  $K(X)$  is  $\underline{P}(x) =$

492  $\min_{P(X) \in K(X)} P(x)$ . Similar expression can be given for the upper probability. Within Walley's  
 493 theory of imprecise probabilities (Walley, 1991), credal sets can then be represented as  
 494 polytopes, where each inner point has a valid probability mass, and can be obtained by  
 495 computing the convex hull of a finite number of probabilities, called vertices (Cozman, 2000).  
 496 A credal set for a random variable  $X_i$  is labelled  $K(X_i)$ , while the set comprising its extreme  
 497 points is denoted by  $\text{ext}[K(X_i)]$ .

498 A credal network CN (Cozman, 2000) over a set of random variables is thus a DAG where  
 499 dependencies among variables are defined by a set of conditional credal sets as  $K(X_i|pa(X_i))$ .  
 500 By analogy with BBN, it is possible to define a joint credal set as follows:

501

$$502 \quad K(X) = CH(P(X): P(x) = \prod_{i=1}^n P(x_i|pa(X_i)) \quad (\text{Eq. 4})$$

503 where  $P(X_i|pa(X_i)) \in \text{ext}[K(X_i|pa(X_i))]$ , CH is the convex hull operator, applied to the  
 504 probabilities computed for the combination of all the vertices of all the conditional credal sets.

505 In this setting, the task of inference aims at computing the probability bounds of the largest  
 506 extension that satisfies the Markov condition (i.e., independence of each node of its non-  
 507 descendant non-parents given its parents) under the assumption of strong independence  
 508 (Cozman, 2000). This results in the convex hull of the set containing all joint distributions that  
 509 factorize the overall joint probability of the network, where the conditional distributions  
 510  $P(X_i|pa(X_i)=\pi_k)$  are selected from the local sets  $K(X_i|pa(X_i)=\pi_k)$ . This task is a NP-hard (de  
 511 Campos and Cozman, 2005), for which a number of exact and approximate algorithms have  
 512 been proposed (Antonucci et al., 2015; Mauá et al. 2012; Ide and Cozman, 2008; Cano et al.,  
 513 2007), but only exact inference algorithms are suitable to polytree-shape binary networks.

514 Though CN allows quantifying and integrating the uncertainty on CPTs on the BBN inference  
 515 results, this increase in expressiveness comes at the expense of higher computational costs.  
 516 Some real case applications of CN exist in different domains (Table 1), but the number of them  
 517 remain limited (with comparison to BBN), despite the availability of some open-source  
 518 solutions like OpenCossan (Tolo et al., 2018), the linear programming algorithm<sup>1</sup> of Antonucci  
 519 et al. (2015), the GL2U-II algorithm<sup>2</sup> of Antonucci et al. (2010).

520

	<b>Domain of application</b>	<b>Method</b>	<b>Reference</b>
--	------------------------------	---------------	------------------

<sup>1</sup> <http://ipg.idsia.ch/software.php?id=135>

<sup>2</sup> <http://ipg.idsia.ch/software.php?id=142>

1	Hazard assessment of debris flows	CN	Antonucci et al. (2007)
2	Military identification	CN	Antonucci et al. (2009)
3	Reliability analysis of a fire-detector system	EN	Simon et al. (2008; 2009; 2017)
4	Threat assessment	EN	Benavoli et al. (2009)
5	Convoy detection	EN	Pollard et al. (2010)
6	Reliability analysis of oil filter plug linked to aero engines	EN	Yang et al. (2012)
7	Railway dysfunction	EN	Aguirre et al. (2013)
8	Food processing	Dynamic CN	Baudrit et al. (2016)
9	Cyber attack analysis	EN	Friedberg et al. (2017)
10	Vulnerability analysis of Nuclear Power Plant subject to external hazards	CN	Tolo et al. (2017)
11	Reliability analysis of a safety instrumentation system for a pressurized vessel	EN	Zhang et al. (2017)
12	Medical prognostic and diagnostic	EN and Fuzzy sets	Janghorbani and Moradi (2017)
13	Fault diagnosis for railway	BBN vs EN	Verbert et al. (2017)
14	Maritime accidents	CN	Zhang and Thai (2018)
15	Terrorist attack analysis on a chemical storage plant	EN vs CN	Misuri et al. (2018)
16	Landslides	CN	He et al. (2018)
17	Risk assessment of an oscillating water column	CN	Estrada-Lugo et al. (2018)
18	Reliability analysis of a feeding control system	EN	Mi et al. (2018)
19	Human reliability for Nuclear Power Plant safety analysis	EN	Deng and Jiang (2018)
20	Safety assessment of a truss	EN	Khakzad (2019)

521 Table 1. Case studies of evidential networks (EN) and Credal Network (CN)

522 [Table 1 about here]

523

### 524 4.3 Methods based on Dempster-Shafer Theory

525 An alternative setting for representing imprecision is the theory of belief functions, also called  
526 Dempster-Shafer Theory, denoted DST (Shafer, 1976, Dempster, 1967). Let  $X$  be a variable  
527 taking values in the frame of discernment  $\Theta$  composed of  $q$  mutually and exhaustive possible  
528 state of  $X$ . For instance, for a binary node, the frame of discernment is  $\Theta = \{\text{True}, \text{False}\}$ .

529 Formally, the theory introduces the concept of basic belief assignment (BBA) based on the  
 530 belief mass function  $m: 2^{\Theta} \rightarrow [0,1]$  and satisfies  $\sum_{A \subseteq \Theta} m(A) = 1$ , and  $m(\emptyset) = 0$  (which  
 531 assumes that at least one element of  $\Theta$  is true). Every  $A \in 2^{\Theta}$  such that  $m(A) > 0$  is called a focal  
 532 element.

533 In classical probabilities, a probability value can be assigned to the state True or False only. By  
 534 defining the belief mass function based on the powerset of the frame of discernment  $2^{\Theta}$  (which  
 535 corresponds in the binary example to  $\{\emptyset, \text{True}, \text{False}, \{\text{True}, \text{False}\}\}$ ) enables the analyst to  
 536 allocate a quantity supporting an additional state termed as epistemic state  $E = \{\text{True}, \text{False}\}$ .  
 537 Due to uncertainty, the analyst may not always be able to determine the amount of masses to  
 538 attribute to each state, and the variable  $X$  may then be in both states, True or False. This means  
 539 that the method allows characterizing uncertainty about the state of a given node.

540 From a mass function  $m$ , two measures can be defined (instead of one for the probabilistic case)  
 541 called the belief ( $Bl$ ) and plausibility ( $Pl$ ) measures. The latter are respectively defined, for any  
 542 event  $A$  as follows:

543

$$544 \quad Bl(A) = \sum_{E \subseteq A} m(E), \text{ and } Pl(A) = \sum_{E \cap A \neq \emptyset} m(E) \quad (\text{Eq. 5})$$

545 where  $Bl$  measures how much event  $A$  is implied by the information (it sums masses that must  
 546 be redistributed over elements of  $A$ ),  $Pl$  measures how much event  $A$  is consistent with the  
 547 information (it sums masses that could be redistributed over elements of  $A$ ). These two  
 548 measures can be associated to a (closed convex) set of bound probabilities  $\{P \mid \forall A \subseteq \Theta, (A) \leq(A) \leq Pl(A)\}$ . It is thus possible to associate an interval-valued probability to the event  $A$ , with  
 550 minimum and maximum probabilities provided by  $Bel$  and  $Pl$ , respectively. This makes the  
 551 formal link with CN. Conversely, it is also possible to reconstruct BBAs from  $Pl$  and  $Bel$   
 552 functions using a Möbius Transformation (Smets, 2002).

553 As an illustration, let us assume consider a binary node for which the expert only knows that  
 554 the probability of the event  $\{X = \text{True}\}$  is at least 0.8. The corresponding BBA is  $m(\{\text{True}\}) = 0.8$ ,  
 555  $m(E = \{\text{True}, \text{False}\}) = 0.2$ ,  $m(\{\text{False}\}) = 0$ . This means that  $Bl(\{\text{True}\}) = m(\{\text{True}\})$ , and  
 556  $Pl(\{\text{True}\}) = m(\{\text{True}\}) + m(E) = 0.8 + 0.2 = 1.0$ . This also means that  $Bl(\{\text{True}\}) = m(\{\text{True}\}) = 0.8$ ,  
 557 and  $Pl(\{\text{True}\}) = m(\{\text{True}\}) + m(E) = 0.8 + 0.2 = 1.0$ . Then  $0.8 \leq P(\{\text{True}\}) \leq 1.0$ .

558 The evidence theory is the basis of evidential networks (EN), which is a DAG propagating  
 559 belief masses. One of the first formulation by Xu and Smets (1996) is based on the Dempster's

560 rule for combining and reasoning with the belief masses. Yet, one major limitation is that the  
561 inference algorithms in this formulation are less effective than the one for traditional BBNs as  
562 underlined for instance by Khakzad (2019). In the domain of system reliability analysis, Simon  
563 et al. (2008) proposed an alternative by mapping logical gates (like OR or AND typically used  
564 for fault tree analysis), as EN with the hypothesis described by Guth (1991). Despite its  
565 similarity with BBN, relations in EN between variables are not probabilities, but belief masses.  
566 The truth table of gates are replaced by conditional mass tables for AND and OR gates (see an  
567 example in Fig. 4(D)). To compute belief and plausibility measures in EN, specific nodes (as  
568 proposed by Simon and Weber 2009) are introduced. Three types of nodes (as represented in  
569 Fig. 4(D)) are thus defined (Simon and Bicking, 2017), namely:

- 570 - Root nodes to which BBA are assigned, correspond to components;
- 571 - Non-root nodes correspond to logical gates that encode its resulting states {True, False,  
572 {True,False}} given the states of its parents;
- 573 - Evaluation nodes correspond to nodes that aim at providing estimates of the belief and  
574 plausibility measures of the system state.

575 In the formulation by Simon and Weber (2009), the inference computation is based on the  
576 Bayes theorem, which is extended to DST by specifying a mass of 1 on one of the focal elements  
577 of the frame of discernment for a specific evidence (hard evidence). Non-specific evidence (soft  
578 evidence) corresponds to a mass distribution on the focal elements of the frame of discernment.  
579 This means that probability updating in such EN can be based on BBN inference algorithms.

580 Misuri et al. (2018) compared CN and EN with illustration on a terrorist attack analysis on a  
581 chemical storage plant. They highlighted that:

- 582 - When used for uncertainty propagation, EN and CN give the same results;
- 583 - In terms of implementation, EN is simpler to use, because they can be built using  
584 existing codes for BBN, whereas CN requires specific codes;
- 585 - In terms of interpretation, Misuri et al. (2018) concluded that EN is more intuitive,  
586 because experts directly assign some weight to the epistemic state (e.g.  $E=\{\text{True, False}\}$   
587 for a binary node), whereas they have to specify interval-valued probabilities for CN,  
588 which can become tricky for multivalued nodes.

589 Khakzad (2019) further filled the gap between CN and EN by proposing some heuristic rules  
590 to determine prior belief masses based on imprecise probabilities. They further modified Simon  
591 and co-authors' EN formulation to both improve the propagation and updating of the belief

592 masses using BBNs. In order to deal with linguistic variables for the network node' states, the  
593 EN method can be combined either with Fuzzy sets (Zadeh, 1975) as applied by Janghorbani  
594 and Moradi (2017) for medical prognostic, or with a Naive Bayes classifier model as applied  
595 by Zhang et al. (2017) for safety analysis for nuclear power plant.

#### 596 **4.4 Discussion**

597 Different settings are available to help the BBN analyst to deal with the problem of uncertainty  
598 propagation. A natural question is the justification for using approaches that are alternative to  
599 classical probabilities. A first argument often highlights the epistemic nature of the CPT  
600 uncertainties. Contrary to aleatory uncertainty (also referred to as randomness), which  
601 represents the variability of the physical environment or engineered system under study,  
602 epistemic uncertainty mainly stems from the incomplete/imprecise nature of available  
603 information (e.g. Hoffman and Hammonds, 1994). While tools from the probabilistic setting  
604 can appropriately handle aleatory uncertainties, it is the second type, which raises several  
605 problems in practice. In our situation, probability distribution cannot be inferred from  
606 data/observations, and should therefore be assumed; the procedure described in Sect. 4.1 is  
607 mainly based on the assumption that the uncertainties on the CPTs are described by a Beta (or  
608 for more generic cases, a Dirichlet probability distribution) probability distribution. Yet, this  
609 assumption may influence the final results of the BBN-based analysis (see. e.g., Ditlevsen, 1994  
610 for an extensive discussion in reliability analysis); Relying only on probabilities masks this  
611 problem and might induce an appearance of more refined knowledge with respect to the existing  
612 uncertainty than is really present (Klir, 1989; 1994). Sect. 4.2 and 4.3 describe alternative non-  
613 probabilistic frameworks to represent uncertainty in situations characterized by limited  
614 available pieces of information, which are mainly restricted to expert judgements. Both  
615 approaches allow improving the expressiveness with respect to uncertainty representation (as  
616 shown by the few tens of application studies using these techniques, see Table 1), in particular  
617 by enabling the BBN analyst to translate his/her uncertainty on the node states or his/her  
618 imprecision on the CPT parameters by avoiding the need for specifying a probability model.

619 Yet, extra-probabilistic approaches (whatever the considered methods, CN, EN or networks  
620 combined with linguistic variables or based on alternative uncertainty theories like possibility  
621 theory, see Dubois et al. (2017)) might come at the expense of higher level of sophistication  
622 and of complexity of the inference algorithms (and potentially higher computational costs). The  
623 danger is to add more confusion than insights as discussed by Aven and Zio (2011) with the  
624 viewpoint of decision making for risk management. The question of selecting the most

625 appropriate approaches for representing and characterizing the risk and uncertainties (in  
626 particular with application on BBNs) still remains open (see e.g., an extensive discussion by  
627 Flage et al. 2014).

## 628 **5 Characterizing the uncertainties**

629 Methods presented in Sect. 4 allows evaluating the impacts of CPT uncertainties on the BBN  
630 results. But, this tells nothing about the respective contribution of the different CPT entries on  
631 the total uncertainty, i.e. the influence of the different uncertainties. This is the purpose of  
632 sensitivity analysis (SA), which can be used, in the construction phase of the BBN model, to  
633 study how the output of a model varies with variation of the CPT parameters. Subsequently,  
634 the results from SA can be used as a basis for parameter tuning, as well as for studying the  
635 robustness of the model output to changes in the parameters (Coupé and van der Gaag, 2002;  
636 Laskey et al., 1995).

### 637 **5.1 Description of the methods**

638 For discrete BBNs, a widespread SA method relies on the use of sensitivity functions (Coupé  
639 and van der Gaag, 2002; Castillo et al., 1997), which describe how the considered output  
640 probability varies as one CPT entry value is changed. An example of application in the domain  
641 of marine safety is provided by Hänninen and Kujala (2012). Formally, consider the conditional  
642 probability  $P(Z=k|e)$ , where  $e$  denotes the available evidence, and a CPT entry  $x=P(X=i|\pi)$  where  
643  $i$  is a value of a variable  $X$  and  $\pi$  is a combination of values for the parents of  $X$ . The sensitivity  
644 function then corresponds to a quotient of two functions that are linear in  $x$  of the following  
645 form:

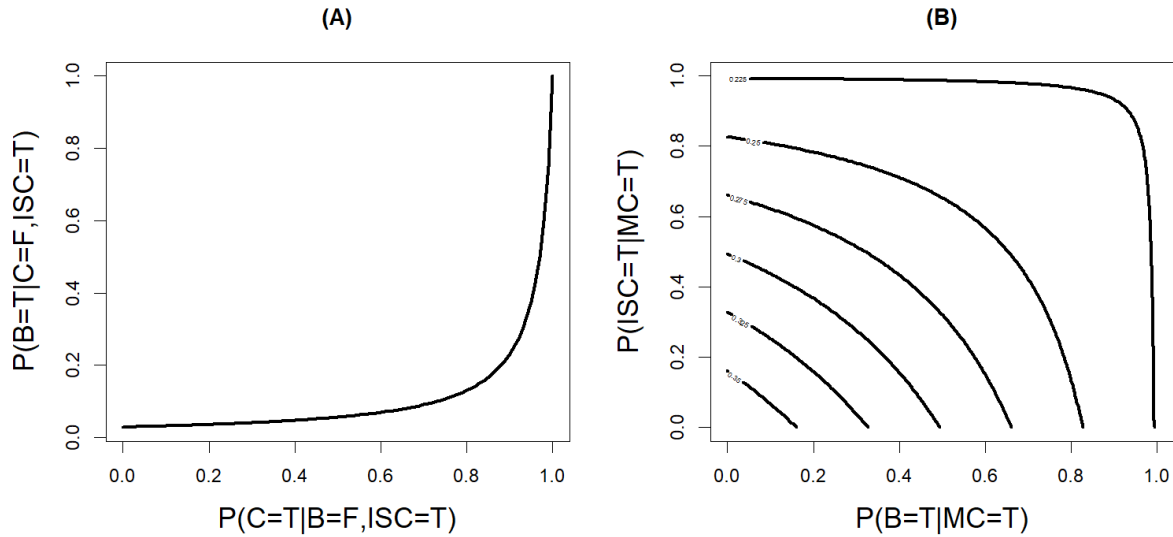
646

$$647 \quad f(x) = \frac{c_1x+c_2}{c_3x+c_4} \quad (\text{Eq. 6})$$

648 where the constants  $c_i$  are built from the values of the network's non-varied parameters. The  
649 numerator of Eq. 6 expresses the joint probability  $P(Z=k|e)$  as a function of  $x$ , and its  
650 denominator describes  $P(e)$  in terms of  $x$ . Using the example described in Fig. 1, we focus on  
651 the probability of having brain tumor given absence of coma but increased level of serum  
652 calcium, i.e.  $P(B = T|C = F, ISC = T)$ . Van der Gaag et al. (2013) estimated the sensitivity of  
653 this probability of interest to the probability of having coma given absence of brain tumor but  
654 increased level of serum calcium,  $x = P(C = T|B = F, ISC = T)$ . The sensitivity function



655 was established as  $\frac{-0.03}{x-1.03}$  (as depicted in Fig. 5(A)). This type of function shows that the  
 656 probability of interest steeply increases when  $x$  exceeds 0.80, i.e. above the original  
 657 parametrization given by Cooper (1984).



658  
 659 Figure 5. (A) One-way sensitivity function for the BBN described in Fig. 1; (B) Two-way  
 660 sensitivity function for the BBN described in Fig. 1 considering the probability  $P(C=T)$  as the  
 661 targeted probability.

662 [Figure 5 about here]

663  
 664 Two-way sensitivity functions can be expressed in a similar form as a quotient of two bi-linear  
 665 functions. Consider the sensitivity function of  $f(x, y)$  that expresses  $P(Z=k|e)$  as a function of  
 666 the parameter probabilities  $x = P(X = i|\pi_X)$  and  $y = P(Y = j|\pi_D)$ , where  $i$  and  $j$  are values of  
 667 the variables  $X$  and  $Y$ , and  $\pi_X$  and  $\pi_Y$  are combinations of values for the parents of  $X$  and of  $Y$ .  
 668 The function holds as follows:

669  
 670 
$$f(x, y) = \frac{c_1x.y+c_2x+c_3y+c_4}{c_5x.y+c_6x+c_7y+c_8} \quad (\text{Eq. 7})$$

671 where the constants  $c_i$  are estimated from the values of the network's non-varied parameters.  
 672 In the tumor BBN example (Fig. 1), we focus on the probability of having a cancer  $P(C=T)$  and  
 673 its sensitivity to the simultaneous variation of the conditional probabilities  $x=P(B=T|MC=T)$   
 674 and  $y=P(ISC=T|MC=T)$ , which was established by van der Gaag et al. (2013) as  $0.374 +$

675  $0.15 \cdot x \cdot y - 0.15 \cdot x - 0.15 \cdot y$  (Fig. 5(B)). This type of function shows that despite the large  
676 variation of  $x$  and  $y$  (from 0 to 1), the probability of interest varies over a moderate range of  
677 values of only ~15%.

678 A complementary approach for SA involves the study of the Chan–Darwiche (CD) distance  
679 (Chan and Darwiche, 2002; 2005), which is a measure for bounding probabilistic belief change.  
680 It is complementary in the sense that it gives insight in the effect of parameter changes on the  
681 global joint distribution, rather than on a specific (posterior) output probability of interest (as  
682 sensitivity functions do). In practices, the CD distance can be used to identify parameter  
683 changes, which lead the closer distance between the original and the varied BBN distributions  
684 (Chan and Darwiche, 2005). It should however be noted that the choice of the type of distance  
685 is rather arbitrary as outlined by Renooij (2014) and other distances like the KL-divergence  
686 (Kullback and Leibler, 1951) or the  $\phi$ -divergences (Ali and Silvey, 1966) could also be of  
687 interest.

688 Recent studies have focused on the properties of the SA methods. Renooij (2014) thoroughly  
689 investigated the different schemes for varying a probability from a (conditional) distribution,  
690 while co-varying the remaining probabilities from the same distribution; the proportional co-  
691 variation scheme being the most popular one. Leonelli et al. (2017) further formalized the SA  
692 problem for discrete BBNs within the generic setting of multilinear models. They developed a  
693 unifying approach to sensitivity methods via the interpolating polynomial representation of  
694 discrete statistical models in the context of “BBNs single full CPT analyses”, i.e. where one  
695 parameter from each CPT of one vertex of a BBN given each configuration of its parents is  
696 varied. This approach based on multilinear probabilistic models enabled them to address the  
697 problem of multi-way SA (with dimension  $\geq 2$ ). Furthermore, they proved the optimality of  
698 proportional covariation by showing that the CD distance is minimized when parameters are  
699 proportionally co-varied.

## 700 **5.2 Discussion**

701 Since the nineties, the BBN community has seen the developments of SA methods that are  
702 specifically dedicated to their respective needs regarding the BBN use and application. Though  
703 simple and efficient to implement, the approach based on sensitivity functions (combined with  
704 CD-distance analysis) remains local, because one parameter values are varied, while the other  
705 ones are kept constant. Multi-way SA methods have been proposed, but can rapidly become  
706 intractable. Interestingly, outside the BBN community, the problem of SA is commonly  
707 addressed with alternative tools; variance-based global SA techniques being the most popular

708 one (Iooss and Lemaitre, 2015). Such techniques were adapted by Li and Mahadevan (2018) to  
 709 bridge the gap between both communities. Their approach has the advantage to be global i.e.  
 710 in the sense that all CPT parameters' values are changed all together. Besides, this approach  
 711 can be applicable to any types of BBN (discrete, hybrid or continuous), i.e. it is model-free.

## 712 **6 Concluding remarks**

### 713 **6.1 Summary**

714 The current survey has investigated how to deal with uncertainties related to the specification  
 715 of CPTs for discrete BBNs. Three questions were addressed, namely: (1) how to constrain the  
 716 uncertainties related to CPT derivation; (2) how to integrate these uncertainties in the BBN-  
 717 based analysis; (3) how to test the robustness of the BBN-based results to these uncertainties.  
 718 Table 2 provides a summary of the main methods/approaches (together with their advantages  
 719 and limits) to answer these questions.

720

Question	Approach	Advantages	Limits	Section
1	Learning CPT by combining data and expert prior knowledge via MAP estimation	<ul style="list-style-type: none"> <li>- It improves the MLE-based fitting when the number of data is limited.</li> </ul>	<ul style="list-style-type: none"> <li>- The representation of expert belief is restricted to the use of Dirichlet priors;</li> <li>- There is a possible problem of “under-fitting” in sparse situations.</li> </ul>	Sect. 2.4-2.5
1	Learning CPT by combining data and qualitative constraints	<ul style="list-style-type: none"> <li>- The accuracy of the MLE/MAP-based fitting is largely improved when data are scarce;</li> <li>- The experts may feel more conformable in providing ordering than precise CPT values.</li> </ul>	<ul style="list-style-type: none"> <li>- Many new estimators are available, but many lack practical recommendations;</li> <li>- There is a possible problem of “under-fitting” depending on the chosen priors.</li> </ul>	Sect. 2.4-2.5
1	Direct elicitation using qualitative statements	<ul style="list-style-type: none"> <li>- The experts may feel more conformable in providing qualitative statements than quantitative estimates.</li> </ul>	<ul style="list-style-type: none"> <li>- Mathematical modelling of linguistic terms may lead to information loss or increased computation burden.</li> </ul>	Sect. 3.1
1	Use of “divorcing” nodes	<ul style="list-style-type: none"> <li>- The number of nodes is decreased through aggregation of nodes.</li> </ul>	<ul style="list-style-type: none"> <li>- Care should be paid to avoid the loss of interactions in the procedure;</li> <li>- It may dilute the sensitivity of the final node(s) to the input nodes;</li> </ul>	Sect. 3.2

			<ul style="list-style-type: none"> <li>- It might increase the uncertainty propagated through the BBN.</li> </ul>	
1	Simplification of the causal structure using logical gates (e.g. Noisy-OR gate)	<ul style="list-style-type: none"> <li>- The number of nodes to be elicited is largely decreased (e.g. from <math>2^n</math> to <math>2n</math> for a binary node with <math>n</math> parents).</li> </ul>	<ul style="list-style-type: none"> <li>- The assumptions on the causal relationships might not always be valid in real life applications;</li> <li>- The simplifications may hamper the BBN performance.</li> </ul>	Sect. 3.2
1	Extracting information on the factor effects from known relationships and extrapolating them	<ul style="list-style-type: none"> <li>- A large variety of different “filling-up” methods exist to relieve the elicitation burden;</li> <li>- Some feedbacks on real case applications exist (e.g. for human reliability analysis).</li> </ul>	<ul style="list-style-type: none"> <li>- Simplifications are introduced and the derived probabilities can only be considered approximations of the true probabilities.</li> </ul>	Sect. 3.3
2	Uncertainty propagation using probabilities	<ul style="list-style-type: none"> <li>- The degree of confidence in the BBN-based results can be quantified.</li> </ul>	<ul style="list-style-type: none"> <li>- The uncertainty representation is restricted to the use of Beta/Dirichlet probability distributions;</li> <li>- It can become computationally intensive.</li> </ul>	Sect. 4.2
2	Uncertainty propagation using intervals with credal networks (CN)	<ul style="list-style-type: none"> <li>- It avoids selecting a probability model to represent the uncertainty;</li> <li>- The experts may feel more comfortable in assigning intervals than probabilities.</li> </ul>	<ul style="list-style-type: none"> <li>- The specification of interval-valued probabilities can become tricky for multivalued nodes;</li> <li>- It needs specific sophisticated inference algorithms and software solutions (with potential high computational costs).</li> </ul>	Sect. 4.3
2	Uncertainty propagation within the Dempster-Shafer Theory by using evidential networks (EN)	<ul style="list-style-type: none"> <li>- The expressiveness is improved like for CN;</li> <li>- EN is more intuitive than CN, because experts directly assign some weight to the epistemic state;</li> <li>- It can be implemented with existing BBN softwares.</li> </ul>	<ul style="list-style-type: none"> <li>- The translation of interval-valued probabilities within this setting can become difficult for multivalued nodes;</li> <li>- The inference algorithms for combining joint/disjoint belief masses are not so effective as those based on probability theory.</li> </ul>	Sect. 4.4
3	Sensitivity Analysis (SA) using sensitivity functions	<ul style="list-style-type: none"> <li>- The theory is well-established;</li> <li>- It is simple to implement;</li> </ul>	<ul style="list-style-type: none"> <li>- It focuses on the influence of one (or multiple) CPT parameters while the</li> </ul>	Sect. 5.1

		<ul style="list-style-type: none"> <li>- The graphical representation is straightforward to interpret.</li> </ul>	<ul style="list-style-type: none"> <li>other ones are kept constant;</li> <li>- It requires specific co-variations schemes;</li> <li>- Multi-way SA can rapidly become intractable.</li> </ul>	
3	Sensitivity Analysis using Chan–Darwiche distance	<ul style="list-style-type: none"> <li>- It complements the sensitivity functions by giving insight in the effect of parameter changes on the global joint distribution.</li> </ul>	<ul style="list-style-type: none"> <li>- It presents the same disadvantages than the sensitivity functions;</li> <li>- The choice of the distance can be rather arbitrary.</li> </ul>	Sect. 5.1

721 Table 2. Summary of the advantages and limits of the main approaches

722 [Table 2 about here]

723 Considering the first question, we described methods for deriving CPT entries from different  
724 sources of information (observations, prior knowledge, expert-based information, etc.).  
725 Traditional estimators like MLE and MAP (or new ones) were proposed to make the best use  
726 of the data available even in scarce situations when completed by qualitative constraints like  
727 knowledge about the monotonic influences between nodes. For rare-event situations like  
728 reliability analysis, the main source of information relies on inputs from expert domain using  
729 different elicitation techniques; the main challenge being the minimization of the workload on  
730 the experts owing to the large number of CPT entries while preserving the quality and  
731 consistency of the elicited result. Elicitation for CPTs generally relies on three (possibly  
732 combined) main approaches: (1) through the assessment of probabilities directly from an expert  
733 or a panel of experts; (2) through a simplification of the causal structure using the popular  
734 Noisy-OR(MAX) model (and its improved versions like the leaky one); (3) through filling-up  
735 methods, which have in particular been thoroughly benchmarked on test cases in the domain of  
736 human reliability analysis.

737 The second question can be addressed using different approaches, either using probabilities, or  
738 imprecise probabilities either using interval-valued probabilities within the setting of credal  
739 networks or within the Dempster-Shafer theory within the setting of evidential networks.  
740 Though the latter approach enables an increase in expressiveness with respect to uncertainty  
741 representation (as shown by the few tens of application studies using these techniques, see Table  
742 1), this might come however at the expense of higher complexity of the inference algorithms  
743 (and higher computational costs). Finally, the third question is investigated by methods  
744 specifically developed for sensitivity analysis of BBN; in particular through the use of one- or  
745 multi- way sensitivity functions.

## 746 **6.2 Discussion and open questions**

747 BBN is now viewed as a suitable tool for overcoming data gaps, estimating uncertainties, and  
748 visualizing complex causal relationships. Despite its clear advantages, it cannot be applied  
749 uncritically, and addressing the question of uncertainties in its construction, and more  
750 specifically in the CPT derivation, should become standard practices to increase the confidence  
751 in its use. The analysis of the literature (Table 2) shows that any analyst is now equipped with  
752 a handful of different tools/methods to address the question of uncertainties. This contributes  
753 to the minimization of the concern of Neil et al. (2000): “In the literature much more attention  
754 is given to the algorithmic properties of BBNs than to the method of actually building them in  
755 practice”.

756 Yet, the developments of these techniques is only one part of the problem, and effort should be  
757 intensified to bring them to an operative state. To this purpose, the implementation of these  
758 techniques within commonly-used BBN software packages (like Beuzen and Simmons 2019  
759 for the widely-used Netica software, Norsys Software Corp., 2006 or Tolo et al. (2018) who  
760 proposed an open-source software package OpenCossan) should be strengthened. Second many  
761 methods lack practical recommendations. Therefore, more benchmark / comparative exercises  
762 are needed to cover broader situations and to serve as best practices for selecting the most  
763 appropriate tools depending the characteristics of the considered situation. For instance, the  
764 filling-up methods benchmarked by Mkrtchyan et al. (2015) should be completed by Noisy-  
765 OR/MAX models, direct elicitation among others, and applied in different contexts (different  
766 network sizes, binary versus multivalued nodes, etc.), as well as domains of application.  
767 Similarly, there is a need for comparing the pros and cons of using alternative frameworks for  
768 uncertainty representation and propagation in BBNs, i.e. comparing approaches using  
769 probabilities, or interval-valued or Dempster-Shafer structures or possibility distributions or  
770 Fuzzy sets, for instance by following the initiatives conducted for probabilistic risk analysis  
771 (e.g., Pedroni et al., 2013, Loschetter et al., 2016).

772 The current work has focused on CPT derivation for discrete BBN development. The second  
773 key ingredient of BBNs is the DAG specification, whose learning from data has been  
774 investigated in numerous studies (e.g., Heinze-Deml et al. (2018), Scutari et al. (2018), Beretta  
775 et al. (2018), etc.). To address the whole spectrum of uncertainties in BBN building, studies  
776 both covering DAG and CPT learning would be beneficial. To integrate both sources of  
777 uncertainty, possible lines of future research may either focus on the improvement of existing  
778 algorithm like the structural expectation-maximization algorithm (Benjumbeda et al., 2019) to  
779 simultaneously learn the structure and parameters of a BN from incomplete data, or on the

780 combination/aggregation of multiple BBNs, each of them being based on a different set of  
781 assumptions either regarding structure or CPT parametrisation (Kim and Cho, 2017; Feng et  
782 al., 2014).

783 Finally, it should be underlined that BBN modelling is a rapidly advancing field (see e.g.,  
784 Marcot and Penman, 2019) that covers new applications and features (like the incorporation of  
785 the time and space dimension, the improvements in the treatment of discrete and continuous  
786 variables, its links with artificial intelligence, among others). The research on the uncertainty  
787 treatment for these new developments is active (see e.g., recent advances for sensitivity analysis  
788 of a wide array of graphical models by Leonelli (2019)), and the scope of the current work  
789 should be broadened in the future to include them.

790

## 791 **Acknowledgements**

792 This study has been carried out within the NARSIS project, which has received funding from  
793 the European Union's H2020-Euratom Programme under grant agreement N° 755439.

794

795 **References**

- 796 Aguirre, F., Sallak, M., Vanderhaegen, F., Berdjag, D., 2013. An evidential network approach  
797 to support uncertain multiviewpoint abductive reasoning. *Information Sciences*, 253, 110-125.
- 798 Ali, S.M., Silvey, S.D., 1966. A general class of coefficients of divergence of one distribution  
799 from another *J. R. Stat. Soc. Ser. B*, 28, 131–142.
- 800 Altendorf, E.E., Restificar, A.C., Dietterich, T.G., 2005. Learning from sparse data by  
801 exploiting monotonicity constraints. In: *Proceeding of the 21st Conf. Uncertainty in Artificial*  
802 *Intelligence*, pp. 18-26.
- 803 Anand, V., Downs, S. M., 2008. Probabilistic asthma case finding: a noisy or reformulation.  
804 In: *Annual AMIA Symposium*, pp. 6–10.
- 805 Antonucci, A., 2011. The Imprecise Noisy-OR Gate. In: *Proceedings of the 14th International*  
806 *Conference on Information Fusion*, 709-715, Chicago.
- 807 Antonucci, A., Brühlmann, R., Piatti, A., Zaffalon, M., 2009. Credal networks for military  
808 identification problems. *International Journal of Approximate Reasoning*, 50(4), 666-679.
- 809 Antonucci, A., de Campos, C.P., Huber, D., Zaffalon, M., 2015. Approximate credal network  
810 updating by linear programming with applications to decision making. *International Journal of*  
811 *Approximate Reasoning*, 58, 25-38.
- 812 Antonucci, A., Salvetti, A., Zaffalon, M., 2007. Credal networks for hazard assessment of  
813 debris flows. In: *Advanced Methods for Decision Making and Risk Management in*  
814 *Sustainability Science*, pp. 237-256.
- 815 Antonucci, A., Sun, Y., de Campos, C. P., Zaffalon, M., 2010. Generalized loopy 2U: a new  
816 algorithm for approximate inference in credal networks. *International Journal of Approximate*  
817 *Reasoning*, 51(5), 474-484.
- 818 Aven, T., Zio, E., 2011. Some considerations on the treatment of uncertainties in risk  
819 assessment for practical decision-making. *Reliability Engineering and System Safety*, 96, 64–  
820 74.
- 821 Azzimonti, L., Corani, G., Zaffalon, M., 2019. Hierarchical estimation of parameters in  
822 Bayesian networks. *Computational Statistics and Data Analysis*, 137, 67–91.
- 823 Baudrit, C., Destercke, S., Wuillemin, P.H., 2016. Unifying parameter learning and modelling  
824 complex systems with epistemic uncertainty using probability interval. *Information Sciences*,  
825 367, 630-647.
- 826 Bauer, E., Koller, D., Singer, Y., 1997. Update rules for parameter estimation in Bayesian  
827 networks. In: *Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence*, pp.  
828 3–13.
- 829 Benavoli, A., Ristic, B., Farina, A., Oxenham, M., Chisci, L., 2009. An application of evidential  
830 networks to threat assessment. *IEEE Trans. Aerospace Electron Syst.*, 45(2), 620–639.
- 831 Benferhat, S., Dubois, D., Garcia, L., Prade, H., 2002. On the transformation between  
832 possibilistic logic bases and possibilistic causal networks. *International Journal of Approximate*  
833 *Reasoning*, 29(2), 135–173.



- 834 Benjumbeda, M., Luengo-Sanchez, S., Larrañaga, P., Bielza, C. 2019. Tractable learning of  
835 Bayesian networks from partially observed data. *Pattern Recognition*, 91, 190-199.
- 836 Beretta, S., Castelli, M., Gonçalves, I., Henriques, R., Ramazzotti, D., 2018. Learning the  
837 structure of Bayesian Networks: A quantitative assessment of the effect of different algorithmic  
838 schemes. *Complexity*, 2018, Article ID 1591878, 12 pages.
- 839 Beuzen, T., Simmons, J., 2019. A variable selection package driving Netica with Python.  
840 *Environmental modelling & software*, 115, 1-5.
- 841 Brooker, P., 2011. Experts, Bayesian belief networks, rare events and aviation risk estimates.  
842 *Safety Science*, 49(8), 1142–1155.
- 843 Beuzen, T., Marshall, L., Splinter, K.D. 2018. A comparison of methods for discretizing  
844 continuous variables in Bayesian Networks. *Environmental modelling & software*, 108, 61-66.
- 845 Bolt, J., van der Gaag, L.C., 2010. An empirical study of the use of the noisy-OR model in a  
846 real-life Bayesian network. In: *International Conference on Information Processing and  
847 Management of Uncertainty in Knowledge-Based Systems*. Springer, Berlin, Heidelberg, pp.  
848 11-20.
- 849 Cain, J., 2001. Planning Improvements in Natural Resources Management. In: *Guidelines for  
850 Using Bayesian Networks to Support the Planning and Management of Development  
851 Programmes in the Water Sector and Beyond*. Centre for Ecology and Hydrology, Wallingford,  
852 UK.
- 853 de Campos, C.P., Cozman, F.G., 2005. The inferential complexity of Bayesian and credal  
854 networks. In: *Proceedings of the Nineteenth International Joint Conference on Artificial  
855 Intelligence*, pp. 1313–1318.
- 856 de Campos, C.P., Ji, Q., 2008. Improving Bayesian network parameter learning using  
857 constraints, In: *Proceedings of the 19th International Joint Conference on Artificial  
858 Intelligence*, Edinburgh, pp. 1–4.
- 859 Cano, A., Gómez, M., Moral, S., Abellán, J., 2007. Hill-climbing and branch-and-bound  
860 algorithms for exact and approximate inference in credal networks. *International Journal of  
861 Approximate Reasoning*, 44(3), 261-280.
- 862 Castillo, E., Gutiérrez, J.M., Hadi, A.S., 1997. Sensitivity analysis in discrete Bayesian  
863 networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*,  
864 27(4), 412-423.
- 865 Chan, H., Darwiche, A., 2002. When do numbers really matter? *J. Artif. Intell. Res.*, 17, 265-  
866 287.
- 867 Chan, H., Darwiche, A., 2005. A distance measure for bounding probabilistic belief change.  
868 *International Journal of Approximate Reasoning*, 38, 149-174.
- 869 Chang, R., Wang, W., 2010. Novel algorithm for Bayesian network parameter learning with  
870 informative prior constraints. In: *Proceedings of the International Joint Conference on Neural  
871 Networks*, 2010, pp.1–8.
- 872 Chen, S.H., Pollino, C.A., 2012. Good practice in Bayesian network modelling. *Environmental  
873 modelling & software*, 37, 134–45.

874 Chin, K.S., Tang, D.W., Yang, J.B., Wong, S.Y., Wang, H., 2009. Assessing new product  
875 development project risk by Bayesian network with a systematic probability generation  
876 methodology. *Expert Systems with Applications*, 36(6), 9879-9890.

877 Chojnacki, E., Plumecocq, W., Audouin, L., 2019. An expert system based on a Bayesian  
878 network for fire safety analysis in nuclear area. *Fire Safety Journal*, 105, 28-40.

879 Cooper, G.F., 1984. NESTOR: a Computer-based Medical Diagnostic Aid that Integrates  
880 Causal and Probabilistic Knowledge. Report HPP-84-48, Stanford University.

881 Coupé, V.M.H., van der Gaag, L.C., 2002. Properties of sensitivity analysis of Bayesian belief  
882 networks *Ann. Math. Artif. Intell.*, 36, 323-356.

883 Cowell, R.G., 1999. Parameter learning from incomplete data using maximum entropy I:  
884 principles, *Statistical Research Report*, 21.

885 Cozman, F.G.. 2000. Credal networks. *Artificial Intelligence*, 120, 199–233.

886 Cozman, F.G., 2005. Graphical models for imprecise probabilities. *International Journal of*  
887 *Approximate Reasoning* 39(2–3), 167–184.

888 D’Angelo, M.F., Palhares, R.M., Cosme, L.B., Aguiar, L.A., Fonseca, F.S., Caminhas, W.M.,  
889 2014. Fault detection in dynamic systems by a fuzzy/Bayesian network formulation. *Appl. Soft*  
890 *Comput.*, 21, 647–653.

891 Dempster, A.P., 1967. Upper and lower probabilities induced by a multivalued mapping. *The*  
892 *annals of mathematical statistics*, 325–339.

893 Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data  
894 via the EM algorithm. *The Royal Statistical Society Series*, B39, 1–38.

895 Dechter, R., 1998. Bucket elimination: A unifying framework for probabilistic inference. In:  
896 *Learning and Inference in Graphical Models*, Springer, Dordrecht, pp. 75-104.

897 Deng, X., Jiang, W., 2018. Dependence assessment in human reliability analysis using an  
898 evidential network approach extended by belief rules and uncertainty measures. *Annals of*  
899 *Nuclear Energy*, 117, 183-193.

900 Diez, F.J., 1993. Parameter adjustment in Bayes networks. The generalized noisy or-gate. In  
901 *Proceedings of the ninth international conference on uncertainty*. In: *artificial intelligence*,  
902 *Morgan Kaufmann Publishers Inc., USA*, pp. 99–105

903 Ditlevsen, O., 1994. Distribution arbitrariness in structural reliability. In: *Proceedings of*  
904 *ICOSSAR’93. Structural Safety & Reliability*, Rotterdam, The Netherlands, pp. 1241-1247.

905 Drury, B., Valverde-Rebaza, J., Moura, M. F., de Andrade Lopes, A., 2017. A survey of the  
906 applications of Bayesian networks in agriculture. *Engineering Applications of Artificial*  
907 *Intelligence*, 65, 29-42.

908 Druzdzel, M.J., van der Gaag, L., 2000. Building Probabilistic Networks: “where do the  
909 numbers come from?” *IEEE Trans. Knowl. Data Eng.*, 12(4), 481–486.

910 Dubois, D. Prade, H., 1988. *Possibility theory*, New York: Plenum.

911 Dubois, D., Fusco, G., Prade, H., Tettamanzi, A.G., 2017. Uncertain logical gates in  
912 possibilistic networks: Theory and application to human geography. *International Journal of*  
913 *Approximate Reasoning*, 82, 101-118.

914 Estrada-Lugo, H.D., Patelli, E., de Angelis, M., Raj, D.D., 2018. Bayesian networks with  
915 imprecise datasets: application to oscillating water column. In: *Safety and Reliability–Safe*  
916 *Societies in a Changing World*. CRC Press, pp. 2611-2618.

917 Fallet-Fidry, G., Weber, P., Simon, C., Iung, B., Duval, C., 2012. Evidential network-based  
918 extension of Leaky Noisy-OR structure for supporting risks analyses. In: IFAC Proceedings  
919 45(20), 672-677.

920 Feelders, A., van der Gaag, L.C., 2006. Learning Bayesian network parameters under order  
921 constraints, *International Journal of Approximate Reasoning*, 42(1–2), 37–53.

922 Feelders, A., van der Gaag, L.C., 2005. Learning Bayesian network parameters with prior  
923 knowledge about context-specific qualitative influences. In: *Proceedings of the 31<sup>st</sup> Conference*  
924 *Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, pp. 193–200.

925 Feng, G., Zhang, J.D., Liao, S.S., 2014. A novel method for combining Bayesian networks,  
926 theoretical analysis, and its applications. *Pattern Recognition*, 47(5), 2057-2069.

927 Fenton, N.E., 2018. Handling Uncertain Priors in Basic Bayesian Reasoning.  
928 doi10.13140/RG.2.2.16066.89280

929 Fenton, N.E., Noguchi, T., Neil, M., 2019. An extension to the noisy OR function to resolve  
930 the ‘explaining away’ deficiency for practical Bayesian network problems. *IEEE Transactions*  
931 *on Knowledge and Data Engineering*, in press.

932 Fenton, N.E., Neil, M., Caballero, J.G., 2007. Using ranked nodes to model qualitative  
933 judgments in Bayesian networks *IEEE Trans. Knowl. Data Eng.*, 19(10), 1420-1432.

934 Flage, R., Aven, T., Zio, E., Baraldi, P., 2014. Concerns, challenges, and directions of  
935 development for the issue of representing uncertainty in risk assessment. *Risk Analysis*, 34(7),  
936 1196-1207.

937 Friedberg, I., Hong, X., McLaughlin, K., Smith, P., Miller, P.C. 2017. Evidential network  
938 modeling for cyber-physical system state inference. *IEEE Access*, 5, 17149-17164.

939 Gao, X.G., Guo, Z.G., Ren, H., Yang, Y., Chen, D.Q., He, C.C., 2019. Learning Bayesian  
940 network parameters via minimax algorithm. *International Journal of Approximate Reasoning*,  
941 108, 62-75.

942 Gehl, P., D’Ayala, D., 2016. Development of Bayesian Networks for the multi-hazard fragility  
943 assessment of bridge systems. *Structural Safety*, 60, 37-46.

944 Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distribution and the Bayesian  
945 restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721–  
946 741.

947 Guo, Z.G., Gao, X. G., Ren, H., Yang, Y., Di, R.H., Chen, D.Q., 2017. Learning Bayesian  
948 network parameters from small data sets: A further constrained qualitatively maximum a  
949 posteriori method. *International Journal of Approximate Reasoning*, 91, 22-35.

950 Guth, M., 1991. A probability foundation for vagueness and imprecision in fault tree analysis.  
951 *IEEE Trans. Reliab.*, 40(5), 563–570.

952 Hänninen, M., 2014. Bayesian networks for maritime traffic accident prevention: benefits and  
953 challenges. *Accident Analysis & Prevention*, 73, 305-312.

954 Hänninen, M., Kujala, P., 2012. Influences of variables on ship collision probability in a  
955 Bayesian belief network model. *Reliability Engineering & System Safety*, 102, 27-40.

- 956 Hänninen, M., Banda, O. A. V., Kujala, P., 2014. Bayesian network model of maritime safety  
957 management. *Expert Systems with Applications*, 41(17), 7837-7846.
- 958 He, L., Topa Gomes, A., Broggi, M., Beer, M., 2018. Risk analysis of infinite slope failure  
959 using advanced Bayesian networks. In: 8th International Workshop on Reliable Engineering  
960 Computing, "Computing with Confidence", 16–18 July 2018, University of Liverpool,  
961 Liverpool, UK
- 962 Heckerman, D., Geiger, D., Chickering, D.M., 1995. Learning Bayesian networks: the  
963 combination of knowledge and statistical data. *Machine Learning*, 20, 197–243.
- 964 Henrion, M., 1989. Some practical issues in constructing belief networks, *Uncertainty in*  
965 *Artificial Intelligence 3*, Elsevier Sciences Publisher, North Holland.
- 966 Henderson, C., Pollino, C.A., Hart, B.T., 2009. Appendix A: the review - workshop discussion  
967 document. In: Hart, B.T., Pollino, C.A. (Eds.), *Bayesian Modelling for Risk-based*  
968 *Environmental Water Allocation*, Waterlines Report Series No. 14, Canberra, pp. 33-83.
- 969 Heinze-Deml, C., Maathuis, M.H., Meinshausen, N., 2018. Causal structure learning. *Annual*  
970 *Review of Statistics and Its Application*, 5, 371-391.
- 971 Hooper, P.M., 2008. Exact distribution theory for belief net responses. *Bayesian Analysis*, 3(3),  
972 615-624.
- 973 Hoffman, F.O., Hammonds, J.S., 1994. Propagation of uncertainty in risk assessments: the need  
974 to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability.  
975 *Risk analysis*, 14(5), 707-712.
- 976 Hospedales, T., Zhou, Y., Fenton, N., Neil, M., 2015. Probabilistic graphical models pa-  
977 rameter learning with transferred prior and constraints. In: *Proceedings of the 31st Conference*  
978 *on Uncertainty in Artificial Intelligence*, AUAI Press, pp. 972-981.
- 979 Ide, J.S., Cozman, F.G., 2008. Approximate algorithms for credal networks with binary  
980 variables. *International Journal of Approximate Reasoning*, 48(1), 275-296.
- 981 İçen, D., Ersel, D., 2019. A new approach for probability calculation of fuzzy events in  
982 Bayesian Networks. *International Journal of Approximate Reasoning*, 108, 76-88.
- 983 Iooss, B., Lemaître, P., 2015. A review on global sensitivity analysis methods. In: *Uncertainty*  
984 *management in simulation-optimization of complex systems*. Springer, Boston, MA, pp. 101-  
985 122.
- 986 Jaeger, M., 2006. The AI&M procedure for learning from incomplete data. In: *Proceedings of*  
987 *the 22nd Conference on Uncertainty in Artificial Intelligence*, pp. 225–232.
- 988 Jäger, W.S., Christie, E.K., Hanea, A.M., den Heijer, C., Spencer, T., 2018. A Bayesian network  
989 approach for coastal risk analysis and decision making. *Coastal Engineering*, 134, 48-61.
- 990 Janghorbani, A., Moradi, M.H., 2017. Fuzzy evidential network and its application as medical  
991 prognosis and diagnosis models. *Journal of biomedical informatics*, 72, 96-107.
- 992 Jensen, F., 2001. *Bayesian Networks and Decision Graphs*, Springer, New York.
- 993 Jøsang, A., 2001. A logic for uncertain probabilities, *Int. J. Uncertain. Fuzziness Knowl. Based*  
994 *Syst.* 9(3), 279–311.

- 995 Jøsang, A., 2016. *Subjective Logic: A Formalism for Reasoning Under Uncertainty*, Springer,  
996 Heidelberg.
- 997 Kaplan, L., Ivanovska, M., 2018. Efficient belief propagation in second-order Bayesian  
998 networks for singly-connected graphs. *International Journal of Approximate Reasoning*, 93,  
999 132-152.
- 1000 Khakzad, N., 2019. System safety assessment under epistemic uncertainty: Using imprecise  
1001 probabilities in Bayesian network. *Safety Science*, 116, 149-160.
- 1002 Kim, K.J., Cho, S.B., 2017. Ensemble bayesian networks evolved with speciation for high-  
1003 performance prediction in data mining. *Soft Computing*, 21(4), 1065-1080.
- 1004 Kleiter, G.D., 1996. Propagating imprecise probabilities in Bayesian networks. *Artificial  
1005 Intelligence*, 88, 143–161.
- 1006 Klir, G.J., 1989. Is there more to uncertainty than some probability theorists might have us  
1007 believe?. *International Journal of General System*, 15(4), 347–378.
- 1008 Klir, G.J., 1994. On the alleged superiority of probabilistic representation of uncertainty. *IEEE  
1009 Transactions on Fuzzy Systems*, 2(1), 27–31.
- 1010 Knochenhauer, M., Swaling, V.H., Dedda, F.D., Hansson, F., Sjökvist, S., Sunnegaerd, K.  
1011 2013. Using Bayesian Belief Network (BBN) modelling for rapid source term prediction. Final  
1012 report (No. NKS--293). Nordisk Kernesikkerhedsforskning, Roskilde, Denmark.
- 1013 Kuhnert, P.M., Martin, T.G., Griffiths, S.P., 2010. A guide to eliciting and using expert  
1014 knowledge in Bayesian ecological models. *Ecol. Lett.*, 13, 900–914.
- 1015 Kullback, S., Leibler, R.A., 1951. On information and sufficiency. *The annals of mathematical  
1016 statistics*, 22(1), 79-86.
- 1017 Kwag, S., Gupta, A., 2017. Probabilistic risk assessment framework for structural systems  
1018 under multiple hazards using Bayesian statistics. *Nuclear Engineering and Design*, 315, 20-34.
- 1019 Laitila, P., Virtanen, K., 2016. Improving construction of conditional probability tables for  
1020 ranked nodes in Bayesian networks. *IEEE Transactions on Knowledge and Data Engineering*,  
1021 28(7), 1691-1705.
- 1022 Langseth, H., Portinale, L., 2007. Bayesian networks in reliability. *Reliability Engineering and  
1023 System Safety*, 92(1), 92–108.
- 1024 Laskey, K.B., 1995. Sensitivity analysis for probability assessments in Bayesian networks *IEEE  
1025 Trans. Syst. Man Cybern.*, 25(6), 901-909.
- 1026 Leonelli, M., 2019. Sensitivity analysis beyond linearity. *International Journal of Approximate  
1027 Reasoning*, in press.
- 1028 Leonelli, M., Goergen, C., Smith, J.Q., 2017. Sensitivity analysis in multilinear probabilistic  
1029 models. *Information Sciences*, 411, 84–97.
- 1030 Li, P.C., Chen, G.H., Dai, L.C., Zhang, L., 2012. A fuzzy Bayesian network approach to  
1031 improve the quantification of organizational influences in HRA frameworks. *Safety science*,  
1032 50(7), 1569-1583.

- 1033 Li, C., Mahadevan, S., 2018. Sensitivity analysis of a Bayesian network. *ASCE-ASME Journal*  
 1034 *of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, 4(1),  
 1035 011003.
- 1036 Liao, W., Ji, Q., 2009. Learning Bayesian network parameters under incomplete data with  
 1037 domain knowledge, *Pattern Recognition*, 42(11), 3046–3056.
- 1038 Loschetter, A., Rohmer, J., de Lary, L., Manceau, J., 2016. Dealing with uncertainty in risk  
 1039 assessments in early stages of a co2 geological storage project: comparison of pure-probabilistic  
 1040 and fuzzy-probabilistic frameworks. *Stochastic Environmental Research and Risk Assessment*,  
 1041 30(3), 813-829.
- 1042 Malagrino, L.S., Roman, N.T., Monteiro, A.M., 2018. Forecasting stock market index daily  
 1043 direction: a Bayesian network approach. *Expert Systems with Applications*, 105, 11-22.
- 1044 Marcot, B.G., Steventon, J.D., Sutherland, G.D., McCann, R.K., 2006. Guidelines for  
 1045 developing and updating Bayesian belief networks applied to ecological modeling and  
 1046 conservation. *Canadian Journal of Forest Research (Revue Canadienne De Recherche*  
 1047 *Forestiere)*, 36, 3063-3074.
- 1048 Marcot, B.G., Penman, T.D., 2019. Advances in Bayesian network modelling: Integration of  
 1049 modelling technologies. *Environmental modelling & software*, 111, 386-393.
- 1050 Masegosa, A.R., Feelders, A.J., van der Gaag, L.C., 2016. Learning from incomplete data in  
 1051 Bayesian networks with qualitative influences. *International Journal of Approximate*  
 1052 *Reasoning*, 69, 18-34.
- 1053 Mauá, D.D., de Campos, C.P., Zaffalon, M., 2012. Updating credal networks is approximable  
 1054 in polynomial time. *International Journal of Approximate Reasoning*, 53(8), 1183-1199.
- 1055 Mi, J., Li, Y. F., Peng, W., Huang, H.Z., 2018. Reliability analysis of complex multi-state  
 1056 system with common cause failure based on evidential networks. *Reliability Engineering &*  
 1057 *System Safety*, 174, 71-81.
- 1058 Milns, I., Beale, C.M., Smith, V.A., 2010. Revealing ecological networks using Bayesian  
 1059 network inference algorithms. *Ecology*, 91(7), 1892-1899.
- 1060 Misuri, A., Khakzad, N., Reniers, G., Cozzani, V., 2018. Tackling uncertainty in security  
 1061 assessment of critical infrastructures: Dempster-Shafer Theory vs. Credal Sets Theory. *Safety*  
 1062 *science*, 107, 62-76.
- 1063 Mkrtchyan, L., Podofillini, L., Dang, V.N., 2015. Bayesian belief networks for human  
 1064 reliability analysis: A review of applications and gaps. *Reliability engineering & system safety*,  
 1065 139, 1-16.
- 1066 Musharraf, M., Bradbury-Squires, D., Khan, F., Veitch, B., MacKinnon, S., Imtiaz, S., 2014.  
 1067 A virtual experimental technique for data collection for a Bayesian network approach to human  
 1068 reliability analysis. *Reliability Engineering & System Safety*, 132, 1-8.
- 1069 Neil, M., Taylor, M., Marquez, D., 2007. Inference in hybrid Bayesian networks using dynamic  
 1070 discretization. *Statistics and Computing*, 17(3), 219-233.
- 1071 Neil, M., Fenton, N., Nielson, L., 2000. Building large-scale Bayesian networks. *The*  
 1072 *Knowledge Engineering Review*, 15(3), 257-284.

- 1073 Norsys Software Corp., 2006. Sensitivity to Findings Netica Software Documentation available  
1074 at [www.norsys.com](http://www.norsys.com)
- 1075 Niculescu, R.S., Mitchell, T.M., Rao, R.B., 2006. Bayesian network learning with parameter  
1076 constraints, *J. Mach. Learn. Res.*, 7(3), 1357–1383.
- 1077 Nojavan, F., Qian, S.S., Stow, C.A., 2017. Comparative analysis of discretization methods in  
1078 Bayesian networks. *Environmental modelling & software*, 87, 64-71.
- 1079 Oniśko, A., Druzdzel, M. J., Wasyluk, H., 2001. Learning Bayesian network parameters from  
1080 small data sets: Application of Noisy-OR gates. *International Journal of Approximate*  
1081 *Reasoning*, 27(2), 165-182.
- 1082 Pearl, J., 1988. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, San  
1083 Francisco, CA
- 1084 Pedroni, N., Zio, E., Ferrario, E., Pisanisi, A., Couplet, M., 2013. Hierarchical propagation of  
1085 probabilistic and non-probabilistic uncertainty in the parameters of a risk model. *Computers &*  
1086 *Structures*, 126, 199–213.
- 1087 Penz, C.A., Flesch, C.A., Nassar, S.M., Flesch, R.C.C., De Oliveira, M.A., 2012. Fuzzy –  
1088 Bayesian network for refrigeration compressor performance prediction and test time reduction,  
1089 *Expert Systems with Applications*, 39, 4268–4273.
- 1090 Podofillini, L, Mkrtychyan, L., Dang, V.N., 2014. Aggregating expert-elicited error probabilities  
1091 to build HRA models. In: *Proceedings of ESREL 2014*. Sept. 14–18, Wroclaw, Poland; 2014,  
1092 pp. 1083–1091.
- 1093 Pollard, E., Rombaut, M., Pannetier, B., 2010. Bayesian networks vs. evidential networks: An  
1094 application to convoy detection. In: *International Conference on Information Processing and*  
1095 *Management of Uncertainty in Knowledge-Based Systems*. Springer, Berlin, Heidelberg, pp.  
1096 31-39.
- 1097 Ramoni, M., Sebastiani, P., 2001. Robust learning with missing data. *Machine Learning*, 45(2),  
1098 147–170.
- 1099 Ren, J., Jenkinson, I., Wang, J., Xu, D., Yang, J., 2009. An offshore risk analysis method using  
1100 fuzzy Bayesian network. *J. Offshore Mech. Arct. Eng.*, 131, 041101.
- 1101 Renooij, S., 2001. Probability elicitation for belief networks: issues to consider. *Knowled. Eng.*  
1102 *Rev.*, 16, 255–269.
- 1103 Renooij, S., 2014. Co-variation for sensitivity analysis in bayesian networks: properties,  
1104 consequences and alternatives. *International Journal of Approximate Reasoning*, 55, 1022-  
1105 1042.
- 1106 Røed, W., Mosleh, A., Vinnem, J.E., Aven, T., 2009. On the use of the hybrid causal logic  
1107 method in offshore risk analysis. *Reliability Engineering and System Safety*, 94(2), 445-455.
- 1108 Scutari, M., Howell, P., Balding, D. J., Mackay, I., 2014. Multiple quantitative trait analysis  
1109 using Bayesian networks. *Genetics*, 198(1), 129-137.

- 1110 Scutari, M., Graafland, C.E., Gutierrez, J.M., 2018. Who Learns Better Bayesian Network  
1111 Structures:Constraint-Based, Score-based or Hybrid Algorithms? In: Proceedings of Machine  
1112 Learning Research, 72, pp. 1-12.
- 1113 Seixas, F.L., Zadrozny, B., Laks, J., Conci, A., Saade, D.C.M., 2014. A Bayesian network  
1114 decision model for supporting the diagnosis of dementia, Alzheimer's disease and mild  
1115 cognitive impairment. *Comput. Biol. Med.*, 51, 140–158.
- 1116 Shafer, G., 1976. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton,  
1117 NJ.
- 1118 Simon, C., Bicking, F., 2017. Hybrid computation of uncertainty in reliability analysis with p-  
1119 box and evidential networks. *Reliability Engineering & System Safety*, 167, 629-638.
- 1120 Simon, C., Weber, P., 2009. Evidential networks for reliability analysis and performance  
1121 evaluation of systems with imprecise knowledge. *IEEE Transactions on Reliability*, 58(1), 69–  
1122 87.
- 1123 Simon, C., Weber, P., Evsukoff, A., 2008. Bayesian networks inference algorithm to implement  
1124 Dempster Shafer theory in reliability analysis. *Reliability Engineering and System Safety*,  
1125 93(7), 950–963.
- 1126 Smets, P., 2002. The application of the matrix calculus to belief functions. *International Journal*  
1127 *of Approximate Reasoning*, 31, 1–30.
- 1128 Tang, H., Liu, S., 2007. Basic theory of fuzzy Bayesian networks and its application in  
1129 machinery fault diagnosis. In: *Proceedings of Fourth International Conference on Fuzzy*  
1130 *Systems and Knowledge Discovery*, 24–27 Aug. 2007, Haikou, Hainan, China, pp.132–137.
- 1131 Thiesson, B., 1995 Accelerated quantification of Bayesian networks with incomplete data, In:  
1132 *Proceedings of the First International Conference on Knowledge Discovery and Data Mining*,  
1133 pp. 306–311.
- 1134 Tolo, S., Patelli, E., Beer, M., 2017. Robust vulnerability analysis of nuclear facilities subject  
1135 to external hazards. *Stochastic Environmental Research and Risk Assessment*, 31(10), 2733-  
1136 2756.
- 1137 Tolo, S., Patelli, E., Beer, M., 2018. An open toolbox for the reduction, inference computation  
1138 and sensitivity analysis of Credal Networks. *Advances in Engineering Software*, 115, 126-148.
- 1139 Van Allen, T., Singh, A., Greiner, R., Hooper, P.M., 2008. Quantifying the uncertainty of a belief  
1140 net response: Bayesian error-bars for belief net inference. *Artificial Intelligence*, 172, 483–513.
- 1141 Van der Gaag, L.C., Kuijper, R., Van Geffen, Y.M., Vermeulen, J.L., 2013. Towards  
1142 uncertainty analysis of Bayesian Networks. In *25th Benelux Conference on Artificial*  
1143 *Intelligence*, Delft, The Netherlands.
- 1144 Verbert, K., Babuška, R., De Schutter, B., 2017. Bayesian and Dempster–Shafer reasoning for  
1145 knowledge-based fault diagnosis—A comparative study. *Engineering Applications of Artificial*  
1146 *Intelligence*, 60, 136-150.
- 1147 Walley, P. 1991. *Statistical reasoning with imprecise probabilities*. Chapman and Hall, London.



- 1148 Weber, P., Medina-Oliva, G., Simon, C., Iung, B., 2012. Overview on Bayesian networks  
 1149 applications for dependability, risk analysis and maintenance areas. *Engineering Applications*  
 1150 *of Artificial Intelligence*, 25(4), 671-682.
- 1151 Werner, C., Bedford, T., Cooke, R.M., Hanea, A.M., Morales-Napoles, O., 2017. Expert  
 1152 judgement for dependence in probabilistic modelling: a systematic literature review and future  
 1153 research directions. *European Journal of Operational Research*, 258(3), 801-819.
- 1154 Wiegmann, D.A., 2005. Developing a methodology for eliciting subjective probability  
 1155 estimates during expert evaluations of safety interventions: application for bayesian belief  
 1156 networks. Technical report from Aviation Human Factors Division AH FD-05-13/NASA-05-  
 1157 4. <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20050238445.pdf>
- 1158 Wisse, B.W., van Gosliga, S.P., van Elst, N.P., Barros, A.I., 2008. Relieving the elicitation  
 1159 burden of bayesian belief networks. In: *Proceedings of the sixth Bayesian modelling*  
 1160 *applications workshop on UAI*. Helsinki, Finland.
- 1161 Woudenbergh, S.P., van der Gaag, L.C., 2015. Propagation effects of model-calculated  
 1162 probability values in Bayesian networks. *International Journal of Approximate Reasoning*, 61,  
 1163 1-15.
- 1164 Xu, H., Smets, P., 1996. Reasoning in evidential networks with conditional belief functions.  
 1165 *International Journal of Approximate Reasoning*, 14(2-3), 155-185.
- 1166 Yang, Y., Gao, X., Guo, Z., Chen, D., 2019. Learning Bayesian networks using the constrained  
 1167 maximum a posteriori probability method. *Pattern Recognition*, 91, 123-134.
- 1168 Yang, J., Huang, H. Z., Liu, Y., Li, Y.F., 2012. Evidential networks for fault tree analysis with  
 1169 imprecise knowledge. *Int. J. Turbo Jet Engines*, 29(2), 111-122.
- 1170 Zadeh, L.A., 1965. Fuzzy sets. *Information and control*, 8(3), 338–353.
- 1171 Zadeh, L.A., 1975. The concept of a linguistic variable and its application to approximate  
 1172 reasoning. *Information sciences*, 8(3), 199–249.
- 1173 Zagorecki, A., Druzdzel, M.J., 2004. An empirical study of probability elicitation under noisy-  
 1174 or assumption. In: *Flairs conference*, pp. 880–886.
- 1175 Zagorecki, A., Druzdzel, M.J., 2012. Knowledge engineering for Bayesian networks: How  
 1176 common are noisy-MAX distributions in practice?. *IEEE Transactions on Systems, Man, and*  
 1177 *Cybernetics: Systems*, 43(1), 186-195.
- 1178 Zhang, X., Mahadevan, S., Deng, X., 2017. Reliability analysis with linguistic data: An  
 1179 evidential network approach. *Reliability Engineering & System Safety*, 162, 111-121.
- 1180 Zhang, G., Thai, V.V., 2018. Addressing the epistemic uncertainty in maritime accidents  
 1181 modelling using Bayesian network with interval probabilities. *Safety science*, 102, 211-225.
- 1182 Zhang, L., Wu, X., Qin, Y., Skibniewski, M.J., Liu, W., 2015. Towards a fuzzy Bayesian  
 1183 network based approach for safety risk analysis of tunnel-induced pipeline damage, *Risk Anal.*  
 1184 36(2), 278–301.

- 1185 Zhou, Y., Fenton, N., Neil, M., 2014. Bayesian network approach to multinomial param- eter  
1186 learning using data and expert judgments. *International Journal of Approximate Reasoning*,  
1187 55(5), 1252–1268.
- 1188 Zhou, Y., Fenton, N., Zhu, C., 2016a. An empirical study of Bayesian network parameter  
1189 learning with monotonic influence constraints. *Decis. Support Syst.*, 87(C), 69–79 .
- 1190 Zhou, K., Martin, A., Pan, Q., 2016b. The belief noisy-or model applied to network reliability  
1191 analysis. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*,  
1192 24(6), 937-960.
- 1193 Zhang, G., Thai, V.V., 2016. Expert elicitation and Bayesian Network modeling for shipping  
1194 accidents: A literature review. *Safety science*, 87, 53-62.
- 1195 Zwirglmaier, K., Straub, D., 2016. A discretization procedure for rare events in Bayesian  
1196 networks. *Reliability Engineering & System Safety*, 153, 96-109.