Safety Science 157 (2023) 105915

Contents lists available at ScienceDirect

Safety Science

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

Distortion models for estimating human error probabilities

Pablo-Ramsés Alonso-Martín [a], Ignacio Montes [b], Enrique Miranda [b,*]

[a] University of Warwick, Department of Statistics, United Kingdom
[b] University of Oviedo, Department of Statistics and Operations Research, Spain

ARTICLE INFO

MSC:
90B25
90B15
90C31
93B35

Keywords:
Human reliability analysis
Human error probability
Bayesian network
Credal network
Distortion models

ABSTRACT

Human Reliability Analysis aims at identifying, quantifying and proposing solutions to human factors causing hazardous consequences. Quantifying the influence of the human factors gives rise to human error probabilities, whose estimation is a cumbersome problem. Since these human factors are usually related to other organisational or technological factors, it has been proposed to apply probabilistic graphical models, such as Bayesian or credal networks. However, these can be problematic when conditional probabilities on missing data are involved. While the solutions proposed so far combine frequentist and subjective approaches and are in general not robust to small modifications in the dataset, in this paper we propose an alternative based on distortion models, which are a type of imprecise probabilities. We perform a comparative analysis, showing that our proposal is consistent with the previous studies while giving rise to robust estimations.

1. Introduction

Major accidents in industry may have catastrophic consequences, and for this reason it is of the utmost importance to have techniques that allow measuring, and then reduce, the associated risks. In this respect, *human reliability analysis* (HRA) collects the different qualitative and quantitative methods that aim to analyse the human errors involved in these accidents. The former seek to identify the factors involved in the human errors and the latter measure the extent of these errors. By means of the quantitative approach, the computation of *human error probabilities* (HEP) is made.

There are many HRA techniques that have been successfully employed in the estimation of HEP and the associated risks; we may consider for instance the SLIM method (Kirwan, 1994; Noroozi et al., 2013; Svenson, 1989) in which the preference for a set of options is quantified based on expert judgment; HEART (Ward et al., 2013; Williams, 1986), that modifies the estimations of HEP taking into account the Error Promoting Conditions (EPC); or THERP (Humphreys, 1995; Swain and Guttmann, 1983), that combines a dataset of error probabilities suitably modified by the assessor using the Performance Shaping Factors (PSF). We refer to (Kirwan, 1996; Kirwan et al., 1997; Kirwan, 1997) for a validation of some of these techniques.

One of the main contributions in HRA is the *Swiss Cheese Model* developed by Reason (1990), that shows that major accidents are usually due to a combination of several errors, both human caused and not, instead of a single one. Based on this, several HRA methods take into account the variety of factors and the interaction between them; this is for instance the case of ATHEANA (Cooper et al., 1996), CAHR (Sträter, 2000) or CREAM (Hollnagel, 1998).

One of the main challenges when performing a rigorous HEP analysis with the above methods is that there may be some degree of subjectivity in the models because of the role the assessor plays; in addition, the use of several experts may entail having to aggregate preferences and resolve disagreements. For this reason, it has also been advocated the use of methods that depend only on the available data. However, it is also difficult to find a complete dataset with information about major accidents with a unified taxonomy. After some earlier studies in Bellamy et al. (2007), Moura et al. (2015, 2016) created the *Multi-attribute Technological Accidents Dataset*, MATA-D, with information about 238 major accidents that occurred in different industries from 1953 to 2013, using reliable sources such as governments, regulators or insurance companies. Following the CREAM taxonomy (Hollnagel, 1998), they analysed the presence of 53 different factors in each accident, splitting them into organisational, technological and person-related factors.

The MATA-D dataset was statistically analysed in Moura et al. (2016), showing that in the vast majority of accidents there was a combination of the three families of errors. Moreover, Morais et al. (2019a,b) built a Bayesian network gathering the conditional dependence between the factors. While Bayesian networks have been advocated in the HRA context for some time (see, among many others, Groth and Mosleh (2011), Islam et al. (2018, 2020), Mkrtchyan et al. (2015),

* Correspondence to: Facultad de Ciencias. Federico García Lorca, 18. 33007 Oviedo, Spain
E-mail addresses: Pablo.Alonso-Martin@warwick.ac.uk (P.-R. Alonso-Martín), imontes@uniovi.es (I. Montes), mirandaenrique@uniovi.es (E. Miranda).

https://doi.org/10.1016/j.ssci.2022.105915
Received 3 January 2022; Received in revised form 14 June 2022; Accepted 22 August 2022
Available online 9 September 2022

Mu et al. (2015)) due to their efficient representation of the probabilistic information and the dependencies between the nodes, Morais et al. (2019a,b) noticed that specifying the conditional probabilities can be quite arbitrary when unobserved events enter the picture. They propose to tackle this issue by considering instead a *credal network*, that allow for sets of probability measures instead of a single one to be assessed on each node. In this manner, they consider a fully conservative (vacuous) model when conditioning must be done on an unobserved event, and proceed with the standard algorithms for credal networks to obtain a set of estimations for the HEP. The use of credal networks is in line with other authors that advocated the use of tools from the imprecise probability theory in reliability engineering, such as Zhang et al. (2018), who proposed the use of a credal network approach in the analysis of maritime accidents, or (Coolen, 1997; Coolen and Newby, 1994), who proposed the use of the imprecise Dirichlet model in reliability analysis.

While we agree with Morais et al. that a credal network approach is the path to follow in situations of imprecise or missing information, we think that the vacuous approach they consider has a couple of issues that may be overcome using tools from imprecise probability theory. On the one hand, the results obtained are not necessarily robust: by just modifying one observation from the dataset we can eliminate all zero probabilities from the conditioning events, and this would lead to much more precise estimations of the HEP. On the other hand, the approach alternates between a purely frequentist non-robust approach (when there is some observation about the conditional parent nodes in the credal network) with an overly cautious subjective approach (when there is not) in the estimation of the probabilities.

We believe that these two issues can be addressed more efficiently by considering a robust approach based on *distortion or neighbourhood models* (Destercke et al., 2022; Montes et al., 2020a,b), which are sets of probabilities built around a probability measure in terms of a distortion parameter and a distance function. They include as particular cases some of the usual models employed in robust statistics (Huber, 1981) such as the linear vacuous model (Walley, 1991) or the Kolmogorov model, among others, and they are connected to the mathematical *theory of imprecise probabilities* (Walley, 1991). Moreover, in some cases they allow to overcome the issues related to conditioning on events of zero probability.

Credal networks have been very scarcely employed in the context of HRA, with the few exceptions discussed earlier; this paper contributes to illustrate their usefulness in this framework and introduces the novelty of using distortion models to give a robust model of the uncertainty in each node.

The paper is organised as follows: after recalling the basics of Bayesian and credal networks (Section 2) and distortion models (Section 3), we explain in Section 4 the use of credal networks by Morais et al. in HRA, and our approach based on some types of distortion models. Next, we perform a comparison of the estimations obtained between the two approaches in Section 5. Our final comments are given in Section 6.

2. Bayesian and credal networks

2.1. Bayesian networks

One efficient graphical representation of the uncertainty associated with a complex experiment is by means of *Bayesian networks* (Pearl, 1988), which are probabilistic graphical models encoding the dependencies between the different variables as well as the associated (conditional) probability distributions.

Formally, a Bayesian network is a directed acyclic graph where the nodes correspond to the variables $\{X_1, \dots, X_n\}$, and the edges represent the dependencies between them. An edge $X_i \to X_j$ between two nodes X_i, X_j $(i \neq j)$ means that there is conditional dependence between the parent (X_i) and the child (X_j). The probabilistic information is thus represented by means of the conditional distribution of each node, given its parents, as well as the marginal distributions on the nodes without parents. From this information, it is possible to derive the joint distribution using the law of total probability and the assumptions of conditional independence.

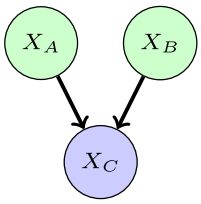

Fig. 1. Example of the relationship between events.

Fig. 1 shows an example of a Bayesian network with two parents (X_A and X_B, in green) and one child (X_C, in blue), the three of them binary variables indicating the occurrence, or not, or the events A, B and C, respectively. The edges $X_A \to X_C$ and $X_B \to X_C$ show a conditional dependence of X_C on the values of the parents X_A and X_B, i.e., the occurrence of C depends on the occurrence of the events A and B. However, since there is no arc between X_A and X_B, events A and B are assumed to be statistically independent.

In this example, we must specify the probability of occurrence of A, $P(X_A = A)$, and that of B, $P(X_B = B)$. Next, we must specify the conditional probability of occurrence of C given the occurrence or not of the parents, A and B. This would allow us for instance to obtain

$$P(X_A = A, X_B = B, X_C = C) =$$
$$P(X_C = C | X_A = A, X_B = B) P(X_B = B) P(X_A = A),$$

using the independence between A and B.

2.2. Credal networks

While Bayesian networks provide an efficient tool for managing the probabilistic information associated with an experiment, there are situations where this information may be imprecisely or ill-specified, due for instance to the existence of missing data or conflicting sources of information. This has given rise to a number of models, usually referred to under the common term *imprecise probabilities* (Augustin et al., 2014).

The question naturally arises of whether it is possible to extend the ideas behind Bayesian networks to be able to deal also with these scenarios of imprecise information. This has produced the model called *credal networks*, from the seminal work by Lamata and Moral (1990) as well as the works in Cano et al. (1993), Cozman (2000), Fagiuoli and Zaffalon (1998); we refer to Mauá and Cozman (2020) for a recent survey on the topic.

A credal network is a generalisation of a Bayesian network where, instead of considering (conditional) probability measures, the uncertainty is represented using a closed and convex set of probability measures, or *credal set* (Levi, 1980). This credal set contains the probability measures which are compatible with the available information about the probability distribution of the random variable X. A credal set, usually denoted by $\mathcal{M}(X)$, may be equivalently represented by means of its lower and upper envelopes[1] $\underline{P}$ and $\overline{P}$, given by:

$$\begin{aligned} \underline{P}(A) &= \min\{P(A) \mid P \in \mathcal{M}(X)\}, \\ \overline{P}(A) &= \max\{P(A) \mid P \in \mathcal{M}(X)\} \end{aligned} \tag{1}$$

for each event A. These two functions are conjugate, in the sense that $\underline{P}(A) + \overline{P}(A^c) = 1$, and therefore it suffices to work with any of the two.

[1] These are called *coherent* lower and upper probabilities in the terminology of (Walley, 1981).

The possibility of conditioning on sets of (lower) probability zero leads to a number of different updating procedures. Out of these, the most informative one is referred to as the *regular extension* (de Campos et al., 1990; Fagin and Halpern, 1991), and is applicable when the conditioning set has positive upper probability. In that case, we can update the credal set by considering:

$$\mathcal{M}(X \mid B) = \{P(\cdot \mid B) \mid P \in \mathcal{M}(X), P(B) > 0\}.$$

That is, we apply Bayes' rule to all the probability measures in the credal set that give strictly positive probability to the conditioning event.

When $\underline{P}(B) > 0$, the regular extension coincides with what Peter Walley called the *natural extension* in Walley (1991). However, if $0 = \underline{P}(B) < \overline{P}(B)$, the natural extension will be *vacuous*: $\mathcal{M}(X|B)$ will be the set of all probability measures on B, while the regular extension will produce in general a more informative model. We refer to Miranda (2009), Miranda and Montes (2015) for a detailed study of the connection between the regular and natural extensions.

Several algorithms have been proposed over the last 30 years for dealing with credal networks. See for example (Antonucci et al., 2013; Cano et al., 2004; de Campos and Cozman, 2007; Mauá et al., 2012b) as well as (Cabañas et al., 2016; De Bock et al., 2014; Mauá et al., 2012a) for algorithms in the context of decision making.

3. Credal networks with distortion models

Credal networks allow accounting for imprecision in a Bayesian network using closed and convex sets of (conditional) probability measures. However, working general credal sets may be involved and computationally expensive: for instance, some inferences with credal sets require determining their set of extreme points, which may have an infinite number of elements or, even when it is finite, may be difficult to determine. For this reason, in this paper we propose to consider particular cases of credal sets that will be easier to handle, and that are associated with *distortion models* (Montes et al., 2020a,b).

Consider a probability measure P_0, a distorting function d and a distortion factor[2] δ. These elements allow us to consider the credal set of those probability measures differing at most δ from P_0:

$$B_d^\delta(P_0) = \{P \text{ probability measure} \mid d(P, P_0) \leq \delta\}.$$

Taking lower and upper envelopes on events, this set determines a lower and upper probability $\underline{P}_d, \overline{P}_d$ (see Eq. (1)); more generally, for any function $f : \mathcal{X} \to \mathbb{R}$ it allows us to determine lower and upper expectation operators[3]:

$$\underline{P}_d(f) = \min\{P(f) \mid P \in B_d^\delta(P_0)\},$$
$$\overline{P}_d(f) = \max\{P(f) \mid P \in B_d^\delta(P_0)\},$$

where $P(f)$ is understood as the expectation of f with respect to P. Whenever d is convex and continuous (Montes et al., 2020a, Prop.1), it is possible to obtain $B_d^\delta(P_0)$ as:

$$B_d^\delta(P_0) = \{P \text{ probability measure } \mid P(f) \geq \underline{P}_d(f)\ \forall f\}.$$

Distortion models appear naturally in many different scenarios. Under a frequentist approach, we may estimate P_0 from the available data and let δ be related to the proportion of noisy data or the distance to the model up to which we want to be robust; within decision making, an expert may elicit its (subjective) probability measure P_0 and δ may represent her credibility; and from a behavioural point of view, we can sometimes interpret $\underline{P}_d, \overline{P}_d$ as supremum buying and infimum selling prices for gambles. This has led to the proposal of many different distortion models, such as the pari mutuel (Montes et al., 2019; Pelessoni et al., 2010; Walley, 1991), the constant odds ratio (Benavoli and Zaffalon, 2013; Walley, 1991), the linear vacuous (Huber, 1981), the total variation (Herron et al., 1997) or the distortion models based on the Kolmogorov or L_1 distances (Montes et al., 2020b).

The vast amount of distortion models renders important the existence of comparison criteria that allow selecting the most appropriate one in each scenario. We may consider the following desirable properties:

(a) That the distortion model is determined by its values $\underline{P}_d(\{x\})$, $\overline{P}_d(\{x\})$ on singletons, and therefore that it is computationally simple.[4]
(b) That it satisfies $\overline{P}(B) > 0$ for any event B, allowing to avoid the problem of conditioning on sets of probability zero. In this respect, we are assuming throughout that there is logical independence between the factors, meaning that any combination of them is assumed to be possible. In the finitary context of this paper, it makes sense then that an imprecise model gives zero lower probability to an event that has not been observed but also a strictly possible upper probability. Note also that if the estimation of the model has been done with a large dataset, the size of this dataset can be taken into account in the estimation of this upper probability, by means of the distortion parameter δ.
(c) That the conditional model belongs to the same family of distortion models, i.e., the model is closed under conditioning.
(d) In that case, that we avoid the phenomenon of *dilation* (Seidenfeld and Wasserman, 1993), meaning that the distortion parameter of the conditional model is not greater than that of the unconditional one.
(e) That the set of probability measures $B_d^\delta(P_0)$ is simple, in that it has a small number of extreme points.
(f) That the lower probability $\underline{P}$ satisfies the property of ∞-monotonicity, that allows us to interpret it in terms of multi-valued mappings (Dempster, 1967).

Finally, it is also worth analysing the amount of imprecision caused by the distortion model, in terms of the comparison between the credal sets $B_d^\delta(P_0)$ for different distorting functions d in terms of set inclusion, once δ and P_0 are fixed.

Table 1 establishes such a comparison, based on the work carried out in Montes et al. (2020a,b).

Taking these results into account, in this work we have decided to work with the *linear vacuous model* and the *total variation model*. Given a probability measure P_0 and a distortion factor δ, the linear vacuous model is defined as the coherent lower probability given by:

$$\underline{P}_{LV}(A) = \begin{cases} (1-\delta)P_0(A) & \text{if } A \neq \mathcal{X}. \\ 1 & \text{if } A = \mathcal{X}. \end{cases}$$

Its conjugate upper probability is given by:

$$\overline{P}_{LV}(A) = 1 - \underline{P}_{LV}(A^c) = \begin{cases} (1-\delta)P_0(A) + \delta & \text{if } A \neq \emptyset. \\ 0 & \text{if } A = \emptyset. \end{cases}$$

Its associated credal set is given by:

$$\mathcal{M}(\underline{P}_{LV}) = \{(1-\delta)P_0 + \delta P \mid P \text{ probability measure}\}.$$

This allows us to give a robust interpretation of the linear vacuous model: we consider that with probability $1-\delta$ the "true" probability measure is P_0, and with probability δ any other probability measure is possible.

[2] While some distortion models are defined by transforming directly a probability measure into a lower probability, it can be checked (Montes et al., 2020a) that they can be embedded into the above, arguably more intuitive, formalism.

[3] These are called lower and upper *previsions* in the imprecise probability literature (Walley, 1991).

[4] In the language of imprecise probabilities, this means that the model is a *probability interval* (de Campos et al., 1994).

Table 1
Comparison of the distortion models. (PMM: Pari-mutuel model; LV: Linear-vacuous; COR: constant odds ratio; TV: total variation: K: Kolmogorov; L_1: neighbourhood model associated with the L_1 distance). N.A.=Not applicable; 1-Worst, 5-Best.

Criterion	Distortion model					
	PMM	LV	COR	TV	K	L_1
Determined by values on singletons?	YES	YES	NO	NO	NO	NO
$\overline{P}(B) > 0\ \forall B \neq \emptyset$?	NO	YES	NO	YES	YES	YES
The conditional model belongs to the same family of distortion models?	YES	YES	YES	YES	NO	NO
Avoids dilation?	NO	NO	YES	NO	N.A.	N.A.
Order in terms of fewer extreme points	3	5	1	4	3	Open problem
$\underline{P}$ ∞-monotone?	NO	YES	NO	NO	YES	NO
Order of imprecision for P_0, δ fixed	4	4	5	3	2	4

On the other hand, the total variation model is defined as the conjugate coherent lower and upper probabilities:

$$\underline{P}_{TV}(A) = \max\{P_0(A) - \delta, 0\},$$
$$\overline{P}_{TV}(A) = \min\{P_0(A) + \delta, 1\} \quad \forall A \subseteq \mathcal{X}.$$

This model has a robust interpretation: the credal set $\mathcal{M}(\underline{P}_{TV})$ is formed by those probability measures at a TV-distance of at most δ from P_0.

As we can see from Table 1, these distortion models possess a number of interesting properties. Among them, they both give strictly positive upper probability to any event B, allowing to compute the conditional models by means of regular extension. For the linear vacuous, the updated model is given by:

$$\underline{P}_{LV}(A \mid B) = \begin{cases} (1 - \delta_{LV})P_{0|B}(A), & \text{if } A \subset B, \\ 1, & \text{if } A = B, \end{cases} \tag{2}$$

where $P_{0|B}(A) = P_0(A \mid B)$ and

$$\delta_{LV} = \frac{\delta}{(1-\delta)P_0(B) + \delta} = \frac{\delta}{\overline{P}_{LV}(B)}, \tag{3}$$

for any $A \subseteq B$ whenever $P_0(B) > 0$, while $\underline{P}_{LV}(A) = 0$ for any $A \subset B$ when $P_0(B) = 0$. This means that the conditional model $\underline{P}_{LV}(\cdot \mid B)$ also belongs to the family of linear vacuous models.

For the total variation model, the updated model is given by:

$$\underline{P}_{TV}(A \mid B) = P_{0|B}(A) - \frac{\delta}{P_0(B)} = P_{0|B}(A) - \delta_{TV}, \tag{4}$$

for any $A \subseteq B$ whenever $P_0(B) > 0$, while $\underline{P}_{TV}(A) = 0$ for any $A \subset B$ when $P_0(B) = 0$. This means that $\underline{P}_{TV}(\cdot \mid B)$ is again a total variation model when $P_0(B) > 0$, being vacuous otherwise.

Even if both the linear vacuous and the total variation models are preserved under conditioning, on the downside they endure the phenomenon of dilation: in the conditional model the distortion parameter δ is greater than in the unconditional one, and therefore the updated model is more imprecise than the initial one for both the linear vacuous and total variation models.

4. Estimation of Human Error Probabilities

In this section we summarise the work carried out in Morais et al. (2019a,b), Moura et al. (2016) to estimate human error probabilities. We begin (Section 4.1) by recalling the MATA-D dataset built in Moura et al. (2016) as well as the procedures carried out in Morais et al. (2018) using Bayesian networks (Section 4.2) and in Morais et al. (2019a,b) using credal networks (Section 4.3). Finally, we introduce our proposal in Section 4.4: the use of distortion models in order to overcome some of the shortcomings of the papers above.

4.1. MATA-D dataset

The lack of complete, unified and rigorous datasets containing information about major accidents is an obstacle that hinders the accuracy of a human reliability analysis. To address this issue, Moura et al. (2016) created the *Multi-Attribute Technological Accidents Dataset* (MATA-D in what follows), containing information about 238 major accidents in industry (refine, oil and gas, ...) reported from reliable sources such as governments or regulators.[5] For each of the 238 catastrophes, 53 factors that could have had an influence were analysed. These were split into three main categories following the CREAM taxonomy (Hollnagel, 1998): *man*, *organisational* and *technological*, each of them with a number of subcategories (see Figures 2–4 in Moura et al. (2016)).

A descriptive analysis of the MATA-D dataset can be found in Moura et al. (2016), from which a number of observations stand out. First of all, the vast majority of accidents are due to the combination of factors from different categories: for instance, in merely 0.84% of the accidents only errors from the factors included in the *man* group were involved, whereas in 47.48% of the accidents there was a combination of factors from the man, organisational and technological groups. This is in line with the Swiss Cheese Model mentioned in Section 1. Secondly, at least one error from the *man* category appeared in 57.14% of the cases, for 82.35% and 95.38% in the case of *technological* or *organisational* issues, meaning that the former, while frequent, are less common than latter two categories. And finally, design failures were detected in 157 of the 238 accidents, and in 72.80% of the accidents where at least one factor from the *man* group was involved.

On the other hand, a deeper statistical analysis of the MATA-D dataset, based on clustering and data mining procedures, was carried out in Moura et al. (2017a,b). It split the accidents into four clusters, taking into account the factors involved in each of them. Recently, (Morais et al., 2022b) applied different machine learning tools to the MATA-D dataset to identify human error and to find interactions between the factors.

4.2. Bayesian network approach

Following (Mkrtchyan et al., 2015), Morais et al. constructed a Bayesian network summarising the interactions between the different factors (Morais et al., 2018). For this aim, each factor is represented by a node, and the arcs represent the significant conditional dependences between them, based on the analysis performed in Moura et al. (2017a,b). A simplified version of the Bayesian network including 15 of the 53 factors was considered in Morais et al. (2019b). It is represented in Fig. 2, where the colour of each node is related to the type of factor. The meaning of these 15 factors is summarised Table 2.[6]

The estimation of the probability distribution in each node is made by means of the MATA-D dataset. In the case of nodes without parents, the unconditional probability distribution is estimated using the relative frequency. For instance, since *design failure* occurred in 157 accidents out of 238, the estimation of its probability of occurrence is 0.6597. For those nodes with predecessors in the graph, the conditional

[5] This dataset is freely available at http://datacat.liverpool.ac.uk/1018/.

[6] While we are reproducing here the 15 factors in Morais et al. (2019b), the full list of possible PSF can be found in Morais et al. (2022b) or in Moura et al. (2015).

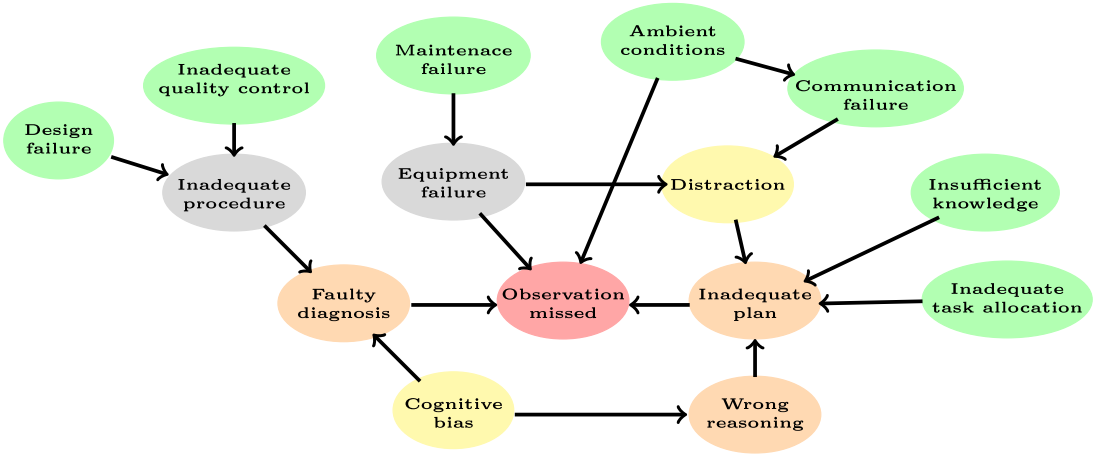

Fig. 2. Bayesian network representing the connection between the considered factors.
Source: From (Morais et al., 2019b, Figure 3).

Table 2
Classification of the factors considered in the Bayesian network in Fig. 2.

Organisational factors	Technological factors	Person-related factors	Human Errors
Design failure	Inadequate procedure	Distraction	Faulty diagnosis
Inadequate quality control	Equipment failure	Cognitive bias	Inadequate plan
Maintenance failure			Wrong reasoning
Ambient conditions			Observation missed
Communication failure			
Insufficient knowledge			
Inadequate task allocation			

probabilities are estimated using also the conditional relative frequencies. For example, *equipment failure* has only one parent, *maintenance failure*, and using the dataset it is estimated that the probability of *equipment failure* when there is *maintenance failure* is 0.675, while the probability of equipment failure when there is not maintenance failure is 0.484.

This approach has in our view a couple of shortcomings. The first one is already mentioned in Morais et al. (2019b): if for a given node the relative frequency of a combination of values of its parents is zero (that is, if that combination has not been observed in the dataset), then it is not possible to obtain any conditional relative frequency. This is for instance the case with the factor *inadequate plan*, that has four parents; there is no accident in the MATA-D dataset combining the occurrence of *distraction* and *wrong reasoning* but without errors in *insufficient knowledge* and *inadequate task allocation*; we refer to Table 5 later on where the full conditional probability table for the factor *inadequate plan* is shown.

The second issue is the lack of robustness of the estimated probabilities. Even if the MATA-D dataset contains information about 238 major accidents, the probabilities obtained from the dataset are estimated from the relative frequencies. Since these involve rare events, changing the absolute frequency in one unit may have a significant impact in the subsequent estimations. This is particularly relevant when an estimated probability changes from zero to a strictly positive number, because in the latter case we will be able to compute conditional relative frequencies.

4.3. Credal network approach

The first of these issues was tackled in Morais et al. (2019a,b, 2022a) by using a credal network. Specifically, they used the relative frequencies of the MATA-D dataset to estimate the probability distributions in the nodes without parents and the conditional probability distributions in those cases where there were observations of the combination of values for the parents. For those cases where there were no observations, they used the [0,1] interval as the estimation of the conditional probability. This procedure leads then to a network where in each node we have possibly a set of (conditional) probability measures, that is, a credal network.

Even if we agree with the benefits of using a credal network-based approach, we believe that this approach has a number of shortcomings: first of all, it does not address the lack of robustness, in the sense that the modification of one observation may have a significant impact in the estimations, particularly because the conditional probability measure may change from being the [0,1] interval to an exact relative frequency[7]; secondly, the approach combines a very conservative approach in some nodes and a precise approach in others, while it would be better to have a unified principle; and finally, from a technical point of view the use of the vacuous model [0,1] is equivalent to applying what is called *natural extension* in the imprecise probability literature, while other, more informative options, can also be applied.

4.4. Our proposal: distortion model approach

Our proposal in this paper is to use distortion models to overcome the issues discussed above. We follow a two-step approach:

(i) First of all, we consider a distortion of the precise unconditional probability measure P_0 that is estimated from the MATA-D dataset. Taking into account the discussion in Section 3, we shall use the linear vacuous and total variation models with some fixed distortion parameter δ. These will give rise to a credal set $\mathcal{M}$, or equivalently to a lower and an upper probability $\underline{P}, \overline{P}$.

[7] Note that this problem may also be overcome by considering an imprecise model that takes into account the amount of data used in the estimation in each node, as proposed recently by Morais et al. (2021) using *c-boxes* (Ferson, 2020).

Table 3
Full conditional probability tables for the factors *Wrong Reasoning* (left) and *Faulty Diagnosis* (right) for the direct estimation from the dataset (frequentist), the approach in (Morais et al., 2019a,b, 2022a) and the LV and TV approaches with distortion parameter $\delta = 0.001$.

Parent	Cognitive bias		True	False
Wrong Reasoning	True	Frequentist	0.294	0.0995
		Morais et al.	0.294	0.0995
		LV ($\delta = 0.001$)	[0.2899,0.3039]	[0.0994,0.1005]
		TV ($\delta = 0.001$)	[0.2800,0.3080]	[0.0984,0.1006]

Parents	Inadequate Plan		True	True	False	False
	Cognitive Bias		True	False	True	False
Faulty Diagnosis	True	Frequentist	0.5	0.155	0.444	0.065
		Morais et al.	0.5	0.155	0.444	0.065
		LV ($\delta = 0.001$)	[0.4839,0.5161]	[0.1546,0.1571]	[0.4330,0.4578]	[0.0649,0.0668]
		TV ($\delta = 0.001$)	[0.4703,0.5298]	[0.1525,0.1575]	[0.4180,0.4708]	[0.0631,0.0669]

Table 4
Estimation of the HEP with the linear vacuous model for different values of δ (FD: Faulty Diagnosis, WR: Wrong Reasoning, OM: Observation Missed, IP: Inadequate Plan (IP)).

Estimations with the linear vacuous model				
δ	FD	WR	OM	IP
0.00001	[0.13005,0.1301]	[0.11338,0.11341]	[0.15546,0.15562]	[0.10344,0.10361]
0.00005	[0.13,0.13023]	[0.11337,0.11347]	[0.15541,0.15618]	[0.10336,0.10423]
0.0001	[0.12993,0.1304]	[0.11335,0.11357]	[0.15537,0.1569]	[0.10323,0.105]
0.00015	[0.12986,0.13057]	[0.11333,0.11366]	[0.15531,0.15765]	[0.10319,0.10579]
0.0002	[0.1298,0.13074]	[0.11331,0.11375]	[0.15526,0.1583]	[0.10302,0.10657]
0.0003	[0.12966,0.13108]	[0.11327,0.11392]	[0.15518,0.1597]	[0.10274,0.10825]
0.0004	[0.12953,0.13142]	[0.11322,0.1141]	[0.15485,0.16139]	[0.10254,0.10985]
0.0005	[0.12939,0.13176]	[0.11318,0.11428]	[0.15471,0.16279]	[0.10261,0.11127]
0.0006	[0.12926,0.1321]	[0.11314,0.11446]	[0.1547,0.16381]	[0.10243,0.11279]
0.0007	[0.12912,0.13245]	[0.1131,0.11464]	[0.15455,0.16504]	[0.10219,0.11482]
0.0008	[0.12899,0.13279]	[0.11306,0.11482]	[0.1542,0.16676]	[0.102,0.1159]
0.0009	[0.12885,0.13313]	[0.11302,0.115]	[0.15414,0.16825]	[0.10196,0.11743]
0.001	[0.12872,0.13347]	[0.11298,0.11518]	[0.15406,0.16874]	[0.10171,0.11932]
Result in (Morais et al., 2019b)	**0.13**	**0.113**	**[0.155,0.168]**	**[0.103,0.109]**

(ii) Secondly, the conditional model in each node with predecessors in the network shall be obtained by means of regular extension (Eq. (2) for the linear vacuous and (4) for the total variation model).

Note then that in step (i) we are adding imprecision to our model, while in step (ii) we are making it in some cases more precise: as we said before, when $P_0(B) = 0$ Morais et al. consider the vacuous conditional model on B (that is, the set of values for $P_0(A|B)$ is the $[0, 1]$ interval for any proper subset A of B); however, since by construction the linear vacuous model satisfies $\overline{P}_{LV}(B) > 0$ for any event $B \neq \emptyset$, we can apply regular extension, that will lead in general to an interval $\left[\underline{P}_{LV}(A|B), \overline{P}_{LV}(A|B)\right]$ that is strictly included in $[0, 1]$. The same comment applies to the total variation model, which also satisfies $\overline{P}_{TV}(B) > 0$ for any $B \neq \emptyset$. Whenever $P_0(B) > 0$, the updated total variation model also gives rise to a proper subinterval of [0,1].

This approach makes the estimations of the probabilities involved in the model more robust, and the extent of this robustness can be measured in terms of the parameter δ. It allows us moreover to avoid the presence of zero probabilities, that, under the frequentist approach considered in Morais et al. (2019a,b) appear as soon as a combination of factors was not observed in any of the 238 accidents: in our method, for any such event we will obtain an interval of probabilities $[0, \delta]$ that we can update by means of regular extension.

5. Comparative analysis

Using the MATA-D dataset, the Bayesian network represented in Fig. 2 and taking the linear vacuous and total variation models, we have performed a comparative analysis with the approach in Morais et al. (2019b). For the numerical computations we have used the Open Cossan Software (Patelli et al., 2018) with the credal network toolbox (Tolo et al., 2018). To illustrate the comparison, we consider two cases, where the approach by Morais et al. gives precise and imprecise estimations, respectively.

5.1. Faulty diagnosis and wrong reasoning

We start considering the factors faulty diagnosis (FD) and wrong reasoning (WR). Their respective full conditional probability tables are given in Table 3 for the frequentist estimation and the estimations using the approach by Morais et al. and LV approach with $\delta = 0.001$. As it can be seen in Fig. 2, the factor WR only has one parent: *cognitive bias*. In the MATA-D dataset, there are both observations where this factor is present and absent, meaning that it is possible to derive the conditional probabilities from the unconditional model using Bayes' rule. This means that in this case the approach in Morais et al. (2019b) gives a precise estimation of the presence of the factor WR in the accident. The same happens for the factor FD: even if this factor has several parents, there is enough information in the MATA-D dataset for applying Bayes' rule. These values are reported in the last row in Tables 4 and 6

Tables 4 and 6 also give the interval of lower and upper probabilities for the probability of error for FD and WR for different values of the distortion factor δ. Looking at these values, we observe that the imprecision obtained using the linear vacuous model is rather small: the difference between the upper and lower probabilities is smaller than 0.0025 (FD) and 0.0011 (WR) for $\delta \leq 0.005$. The imprecision increases when considering the total variation model: it is approximately twice the imprecision of the linear vacuous model approach.

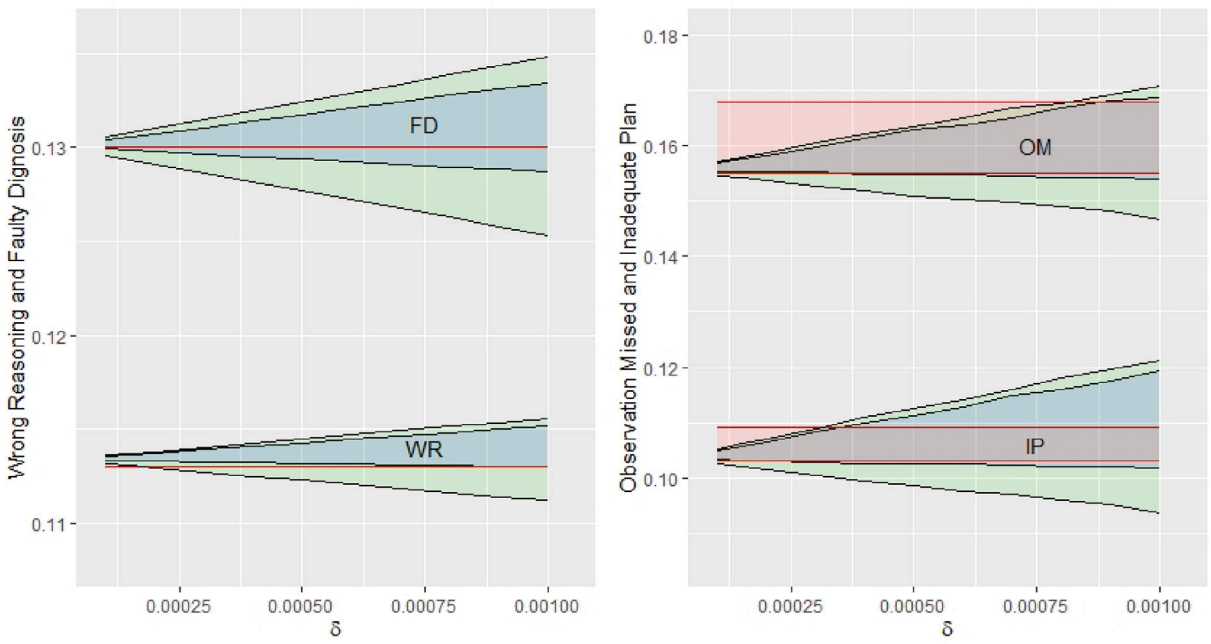

Fig. 3. Graphical representation of the estimated HEP (Faulty Diagnosis (FD), Wrong Reasoning (WR), Observation Missed (OM) and Inadequate Plan (IP)) using the approach from (Morais et al., 2019a,b) (in red), the linear vacuous model (in blue) and the total variation model (in green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5.2. Observation missed and inadequate procedure

We now consider two other factors: observation missed (OM) and inadequate plan (IP). The full conditional probability tables are given in Table 5 for the frequentists estimation and the estimations using the approach by Morais et al. and LV and TV approaches with $\delta = 0.001$. Both factors have four parents and, in contrast with WR and FD, here it is not always possible to apply Bayes' rule to estimate the conditional probabilities because some combinations of values of the parents have not been observed in the dataset. In those cases, the approach in Morais et al. (2019b) uses the vacuous model, and the combination of all the values produces the lower and upper probabilities in the last row of Tables 4 and 6.

These tables also show our estimation for the error probabilities of these factors for different values of the distortion factor δ. Note that, while these intervals have always a non-empty intersection, they neither include nor are included in general in the one obtained in the approach by Morais et al.

The graphical representation of the results obtained with the linear vacuous and total variation models can be seen in Fig. 3.

5.3. Discussion

The results shown in Tables 4 and 6 for the linear vacuous and total variation models, respectively, or the graphical representation in Fig. 3, suggest that the distortion model based approach is an interesting and robust alternative for the estimation of HEP. There are a number of reasons supporting our claim.

First of all, the estimations obtained are consistent with those in Morais et al. (2019b). It can be seen that the precise estimations (for FD or WR) are either included in the intervals determined by the lower and upper probabilities or, for very small values of δ, the precise estimation from Morais et al. (2019b) is very close to the lower bound of the interval probability. For the interval estimations (OM or IP), the interval probabilities obtained in Morais et al. (2019b) and the ones obtained using the linear vacuous or the total variation model are quite similar, and in particular they are never disjoint; we observe also that there is not in general a relation of inclusion between the intervals.

Secondly, as we can see from the detailed studies of the distortion models in Destercke et al. (2022), Montes et al. (2020a,b), the estimations obtained with the total variation model are uniformly more imprecise than those obtained with the total variation model. This means that for obtaining a similar imprecision, the total variation approach requires a smaller distortion parameter.

Thirdly, regarding the interval estimations (OM and IP), it should be mentioned that the lower bounds obtained using the linear vacuous model and the approach in Morais et al. (2019b) are quite similar for all the values considered for δ. Nevertheless, when considering the upper bound, the linear vacuous approach gives a tighter estimation than the approach in Morais et al. (2019b) for values of $\delta \leq 0.0005$. There is an intuitive reason behind this fact: in the approach of Morais et al. when there is not enough information to apply Bayes' rule, they assign an interval probability of [0,1]. If we compare it with the one determined by our approach, usually the value 0 will not be very distant from the lower bound given by the linear vacuous model. However, the upper probability 1 will be substantially larger than the upper probability determined by the conditional linear vacuous model, and this is what eventually leads to a too large upper estimation of the HEP for OM and IP. As we explained in previous sections, the vacuous model (the [0,1] interval) adds too much imprecision in the model, and this is a problem that can be easily overcome with the distortion based approach.

In spite of these positive comments regarding the distortion based approach, a natural criticism would be related to the appropriate election of the distortion parameter δ. Of course, choosing the adequate distortion parameter is the crucial point of this approach.

Indeed, the adequate choice of the distortion parameter has also been analysed in a number of other applications of distortion models (see for instance (Antonini et al., 2020) for the linear vacuous or (Langer, 2017) for the total variation model). While we refer to Montes et al. (2020a,b) for a general discussion of the interpretation of δ in the case of the linear vacuous and total variation models, in the specific context of HRA we believe that the election of the parameter should be made by an expert taking into account different facets: (i) we are estimating small probabilities, hence the amount of imprecision we add to the model (directly related to δ) should be "small". Note that any $\delta > 0$ allows to give strictly positive upper

Table 5
Full conditional probability tables for the factors *Observation Missed* (above) and *Inadequate Plan* (below) for the direct estimation from the dataset (frequentist), the approach in (Morais et al., 2019a,b, 2022a) and the LV and TV approaches with distortion parameter $\delta = 0.001$. Note: T, F and ? denote *true*, *false* and *unknown*, respectively.

Parents	Equipment Failure		T	T	T	T	T	T	T	T	F	F	F	F	F	F	F	F
	Ambient Conditions		T	T	T	T	F	F	F	F	T	T	T	T	F	F	F	F
	Faulty Diagnosis		T	T	F	F	T	T	F	F	T	T	F	F	F	T	T	F
	Inadequate Plan		T	F	T	F	T	F	T	F	T	F	T	F	T	F	T	F
Observation Missed	True	Frequentist	?	?	0	0.125	0	0.333	0.125	0.1	?	?	?	0.125	0.25	0.385	0.375	0.149
		Morais et al.	[0,1]	[0,1]	0	0.125	0	0.333	0.125	0.1	[0,1]	[0,1]	[0,1]	0.125	0.25	0.385	0.375	0.149
		LV ($\delta = 0.001$)	[0,0.2241]	[0,0.0334]	[0,1]	[0.1213,0.1510]	[0,1]	[0.3267,0.3466]	[0.1213,0.1510]	[0.0998,0.1021]	[0,0.2756]	[0,0.0411]	[0,1]	[0.1213,0.1510]	[0.2351,0.2946]	[0.3776,0.3959]	[0.3638,0.3936]	[0.1482,0.1514]
		TV ($\delta = 0.001$)	[0,0.2241]	[0,0.0334]	[0,1]	[0.0953,0.1547]	[0,1]	[0.3036,0.3631]	[0.0953,0.1547]	[0.0976,0.1024]	[0,0.2756]	[0,0.0411]	[0,1]	[0.0953,0.1547]	[0.1905,0.3095]	[0.3663,0.4029]	[0.3453,0.4047]	[0.1454,0.1519]
Parents	Wrong Reasoning		T	T	T	T	T	T	T	T	F	F	F	F	F	F	F	F
	Inadequate Task Allocation		T	T	T	T	F	F	F	F	T	T	T	T	F	F	F	F
	Insufficient Knowledge		T	T	F	F	T	T	F	F	T	T	F	F	T	T	F	F
	Distraction		T	F	T	F	T	F	T	F	T	F	T	F	T	F	T	F
Inadequate Plan	True	Frequentist	0.5	0	?	0.2	?	0.333	?	0	0.667	0.116	0.2	0.072	0	0.188	0	0.056
		Morais et al.	0.5	0	[0,1]	0.2	[0,1]	0.333	[0,1]	0	0.667	0.116	0.2	0.072	0	0.188	0	0.056
		LV ($\delta = 0.001$)	[0.4405,0.5595]	[0,0.0149]	[0,0.3830]	[0.1902,0.2393]	[0,0.3830]	[0.3069,0.3862]	[0,0.5764]	[0,0.2380]	[0.6138,0.6931]	[0.1156,0.1212]	[0.1905,0.2381]	[0.0722,0.0757]	[0,0.2380]	[0.1847,0.1996]	[0,0.0793]	[0.0561,0.0595]
		TV ($\delta = 0.001$)	[0.3810,0.6190]	[0,0.0149]	[0,1]	[0.1524,0.2476]	[0,0.3830]	[0.2540,0.4127]	[0,0.5764]	[0,0.2380]	[0.5873,0.7460]	[0.1107,0.1218]	[0.1524,0.2476]	[0.0690,0.0759]	[0,0.2380]	[0.1726,0.2024]	[0,0.0793]	[0.0530,0.0597]

Table 6
Estimation of the HEP with the total variation model for different values of δ (FD: Faulty Diagnosis, WR: Wrong Reasoning, OM: Observation Missed, IP: Inadequate Plan (IP)).

Estimations with the total variation model				
δ	FD	WR	OM	IP
0.00001	[0.13002,0.13011]	[0.11337,0.11341]	[0.1554,0.15564]	[0.10336,0.10364]
0.0005	[0.12983,0.1303]	[0.11328,0.1135]	[0.15508,0.15628]	[0.10298,0.10435]
0.0001	[0.12959,0.13054]	[0.11317,0.11361]	[0.15464,0.15713]	[0.10254,0.10521]
0.00015	[0.12935,0.13078]	[0.11306,0.11372]	[0.15418,0.15782]	[0.10191,0.10619]
0.0002	[0.12912,0.13191]	[0.11295,0.11383]	[0.15368,0.15865]	[0.10157,0.10699]
0.0003	[0.12865,0.13149]	[0.11273,0.11405]	[0.15276,0.16061]	[0.10055,0.10874]
0.0004	[0.12817,0.13196]	[0.11251,0.11427]	[0.15193,0.16204]	[0.099333,0.11092]
0.0005	[0.1277,0.13244]	[0.11229,0.11449]	[0.15097,0.16352]	[0.098504,0.11249]
0.0006	[0.12723,0.13292]	[0.11208,0.11471]	[0.15033,0.1651]	[0.097586,0.11414]
0.0007	[0.12676,0.1334]	[0.11186,0.11493]	[0.1497,0.16677]	[0.096944,0.11597]
0.0008	[0.12629,0.13388]	[0.11164,0.11515]	[0.14908,0.16756]	[0.095972,0.11806]
0.0009	[0.12582,0.13435]	[0.11142,0.11537]	[0.14825,0.16918]	[0.095247,0.1197]
0.001	[0.12535,0.13483]	[0.11121,0.11559]	[0.14667,0.1709]	[0.09371,0.12108]
Result in (Morais et al., 2019b)	**0.13**	**0.113**	**[0.155,0.168]**	**[0.103,0.109]**

probability to any combination of factor, capturing the assumption that all the combinations are possible, even if some of them may be rather improbable; (ii) the problem with the approach in Morais et al. appears when there is not enough information to apply Bayes' rule. This means that there are events whose estimation in the sample is 0 (out of 238). If such an event appears in the next observation, we would obtain an estimation of 1/239, hence it seems natural to take a parameter δ smaller than that value (which gives $\delta \leq 0.0042$); in fact, this idea of taking the sample size into account when considering the amount of imprecision that is entered into the model is also present in the recent work by Morais et al. (2021); (iii) we should take into account that, even if both the linear vacuous and total variation models are preserved under conditioning, they both suffer from the problem of dilation. This means that the distortion parameter increases any time that we update the model. Hence, the parameter δ should be small enough such that after updating the model a number of times, the distortion parameter is still small enough. In order to control the dilation, a strategy could be to fix the amount of imprecision in the updated models, and from this imprecision derive, using Eqs. (3) or (4), the largest δ that assures that after conditioning we will always obtain an updated parameter smaller than the fixed amount of imprecision.

All these comments led us to perform our analysis with different values of δ varying from 0.0001 to 0.001. In fact, we have already argued that for distortion factors $\delta \leq 0.0005$, the results from the linear vacuous model are quite consistent with those in Morais et al. (2019b), or even less imprecise. This goes in line with our previous comments: choosing a distortion factor smaller than the inverse of the sample size allows to obtain satisfactory results. If in addition the number of times that the model must be updated is large, the expert may decrease further to account for the dilation, as discussed earlier.

6. Conclusions

Our results show that the use of a distortion model can be used to overcome the issues caused by conditioning on sets of probability zero in the Bayesian network, while at the same time providing a robust interpretation of the model. Moreover, the estimations are comparable to those obtained by Morais et al. by means of a mixture of precise and vacuous models, when the conditioning events have positive and zero probability, respectively.

One of the fundamental ideas of the distortion model approach proposed in this paper is that it allows to encompass the idea that any combination of factors is possible, by giving it a positive upper probability, even if this can be very small, considering the sample size, the fact that it may have been so far unobserved (see for instance Tables 3 and 5) and the value of the distortion parameter chosen.

However, this is not to say that our approach is without shortcomings. While the linear vacuous and total variation models have many interesting properties, they also suffer from the phenomenon of dilation, that means that the conditional models also belong to the same family but are associated with a greater distortion parameter. It would be interesting then to consider some alternative that is not affected by this problem. Unfortunately, the constant odds ratio model, that is the only dilation free in our comparative, does not guarantee in general that all conditioning events have positive upper probability, which is necessary if we want to apply the procedure of regular extension to obtain the conditional models.

In addition, it would be interesting to deepen into our approach so as to give some further guidelines about the choice of the distortion parameter δ and its relationship with the imprecision in the estimation of the probabilities of the different events. This would also allow us to deepen in the comparison with the results of Morais et al.

As future lines of research, it would also be interesting to compare our estimations with the model recently proposed by Morais et al. (2021), where the transition between precise and imprecise conditional probabilities in the nodes is made more gradual by introducing confidence boxes. In this respect, one feature of our approach is that distortion models can be used directly when the credal network involves non-binary variables, and some of their properties, such as the small number of extreme points of the associated credal set, are more advantageous in that case; presumably, in such a case confidence boxes may have to be replaced by other models such as p-boxes. In addition, it would be interesting to apply the distortion based approach to the whole network with the 53 factors considered in Morais et al. (2022b).

Finally, our approach advocates the estimation of the HEP from the available data and the only input of the assessor is to introduce the cautious parameter δ. As argued before, this parameter should be smaller than the inverse of the sample size and may also take into account the number of times the model is updated. Nevertheless, there are of course many interesting features in the models mentioned in the introduction that may be interesting to incorporate in our model and that should help to improve the estimations. A deeper analysis of this matter is our main open task for the future.

CRediT authorship contribution statement

Pablo-Ramsés Alonso-Martín: Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – review & editing, Visualization. **Ignacio Montes:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualization. **Enrique Miranda:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Acknowledgements

The research reported in this paper has been supported by project PGC2018-098623-B-I00 from the Spanish Ministry of Science and Innovation. A preliminary version was presented at the ISIPTA'2021 conference. The helpful suggestions and remarks from the participants are gratefully acknowledged. We would like to thank the reviewers for their useful remarks.

References

Antonini, P., Petturiti, D., Vantaggi, B., 2020. Dynamic portfolio selection under ambiguity in the ϵ-contaminated binomial model. In: Proceedings of the 18th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems. IPMU2020, pp. 210–223.

Antonucci, A., de Campos, C., Huber, D., Zaffalon, M., 2013. Approximating credal networks inferences by linear programming. In: Proceedings of the 12th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, Vol. 7958. ECSQARU, pp. 13–25.

Augustin, T., Coolen, F., de Cooman, G., Troffaes, M. (Eds.), 2014. Introduction to Imprecise Probabilities. In: Wiley Series in Probability and Statistics, Wiley.

Bellamy, L., Ale, B., Geyer, T., Goossens, L., Hale, A., Oh, J., Mud, M., Bloemhof, A., Papazoglou, I., Whiston, J., 2007. Storybuilder-A tool for the analysis of accident reports. Reliab. Eng. Syst. Satefy 92, 735–744.

Benavoli, A., Zaffalon, M., 2013. Density-ratio robustness in dynamic state estimation. Mech. Syst. Signal Process. 37, 54–75.

Cabañas, R., Antonucci, A., Cano, A., Gómez-Olmedo, M., 2016. Evaluating interval-valued influence diagrams. Int. J. Approx. Reason. 80, 393–411.

Cano, J., Delgado, M., Moral, S., 1993. An axiomatic framework for propagating uncertainty in directed acyclic networks. Int. J. Approx. Reason. 8, 253–280.

Cano, A., Gómez, M., Moral, S., Abellán, J., 2004. Hill-climbing and branch-and-bound algorithms for exact and approximate inference in credal networks. Int. J. Approx. Reason. 44, 261–280.

Coolen, F., 1997. An imprecise Dirichlet model for Bayesian analysis of failure data including right-censored observations. Reliab. Eng. Syst. Saf. 56, 61–68.

Coolen, F., Newby, M., 1994. Bayesian reliability analysis with imprecise prior probabilities. Reliab. Eng. Syst. Saf. 43, 75–85.

Cooper, S., Ramey-Smith, A., Wreathall, J., Parry, G., 1996. A Technique for Human Error Analysis-Technical Basis and Methodology Description. Technical Report US Nuclear Regulatory Commision Library, Washington D.C.

Cozman, F., 2000. Credal networks. Artif. Intell. 120, 199–233.

De Bock, J., de Campos, C., Antonucci, A., 2014. Global sensitivity analysis for map inference in graphical models. In: Advances in Neural Information Processing Systems, vol. 27, MIT Press, pp. 2690–2698.

de Campos, C., Cozman, F., 2007. Inference in credal networks through integer programming. In: G. de Cooman, J. Vejnarová, Zaffalon, M. (Eds.), ISIPTA '07 – Proceedings of the Fifth International Symposium on Imprecise Probability: Theories and Applications. Action M Agency, Prague (Czech Republic), pp. 145–154.

de Campos, L.M., Huete, J.F., Moral, S., 1994. Probability intervals: A tool for uncertain reasoning. Int. J. Uncertain., Fuzziness Knowl.-Based Syst. 2, 167–196.

de Campos, L., Lamata, M., Moral, S., 1990. The concept of conditional fuzzy measures. Int. J. Intell. Syst. 5, 237–246.

Dempster, A.P., 1967. Upper and lower probabilities induced by a multivalued mapping. Ann. Math. Stat. 38, 325–339.

Destercke, S., Montes, I., Miranda, E., 2022. Processing distortion models: A comparative study. Int. J. Approx. Reason. 145, 91–120.

Fagin, R., Halpern, J., 1991. A new approach to updating beliefs. In: Bonissone, P., Henrion, M., Kanal, L., Lemmer, J. (Eds.), In: Uncertainty in Artificial Intelligence, Vol. 6, North-Holland, Amsterdam, pp. 347–374.

Fagiuoli, E., Zaffalon, M., 1998. 2U: An exact interval propagation algorithm for polytrees with binary variables. Artif. Intell. 106, 77–107.

Ferson, S., 2020. Computing with confidence. https://sites.google.com/site/confidenceboxes/.

Groth, K., Mosleh, A., 2011. Development and use of a bayesian network to estimate human error probability. In: Proceedings of the 2011 International Topical Meeting on Probabilistic Safety Assessment and Analysis. PSA 2011.

Herron, T., Seidenfeld, T., Wasserman, L., 1997. Divisive conditioning: Further results on dilation. Philos. Sci. 64, 411–444.

Hollnagel, E., 1998. Cognitive Reliability and Error Analysis Method (CREAM). Elsevier.

Huber, P., 1981. Robust Statistics. Wiley, New York.

Humphreys, P., 1995. Human Reliability Assessor's Guide: A Report by the Human Factors in Reliability Group Volume SRDA-11. AEA Technology.

Islam, R., Anantharaman, M., Khan, F., Abbassi, R., Garaniya, V., 2020. A hybrid human reliability assessment technique for themaintenance operations of marine and offshore systems. Process Saf. Prog. 39.

Islam, R., Khan, F., Abbassi, R., Garaniya, V., 2018. Human error probability assessment during maintenance activities of marine systems. Saf. Health Work 9, 42–52.

Kirwan, B., 1994. A Guide to Particular Human Reliability Assessment. Taylor and Francis.

Kirwan, B., 1996. The validation of three human reliability quantification techniques –THERP, HEART and JHEDI: Part 1 - technique descriptions and validation issues. Appl. Ergon. 27, 359–373.

Kirwan, B., 1997. The validation of three human reliability quantification techniques –THERP, HEART and JHEDI: Part 3- practical aspects of the usage of the techniques. Appl. Ergon. 28, 27–39.

Kirwan, B., Kennedy, R., Taylor-Adams, S., Lambert, B., 1997. The validation of three human reliability quantification techniques – THERP, HEART and JHEDI: Part 2 - results and validation exercise. Appl. Ergon. 28, 17–25.

Lamata, M., Moral, S., 1990. Dependence graphs: Upper and lower probabilities. Syst. Anal. Comput. Sci. 113–122.

Langer, A., 2017. Automated parameter selection for total variation minimization in image restoration. J. Math. Imaging Vis. 57, 239–268.

Levi, I., 1980. the Enterprise of Knowledge. MIT Press, Cambridge.

Mauá, D., Cozman, F., 2020. Thirty years of credal networks: Specification, algorithms and complexity. Int. J. Approx. Reason. 126, 133–157.

Mauá, D., de Campos, C., Zaffalon, M., 2012a. The complexity of approximately solving influence diagrams. In: Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence. pp. 604–613.

Mauá, D., de Campos, C., Zaffalon, M., 2012b. Updating credal networks is approximable in polynomial time. Int. J. Approx. Reason. 53, 1183–1199.

Miranda, E., 2009. Updating coherent lower previsions on finite spaces. Fuzzy Sets Syst. 160, 1286–1307.

Miranda, E., Montes, I., 2015. Coherent updating of non-additive measures. Int. J. Approx. Reason. 56, 159–177.

Mkrtchyan, L., Podofillini, K., Dang, V., 2015. Bayesian belief network for human reliability analysis: a review of applications and gaps. Reliab. Eng. Syst. Saf. 139, 1–16.

Montes, I., Miranda, E., Destercke, S., 2019. Pari-mutuel probabilities as an uncertainty model. Inf. Sci. 481, 550–573.

Montes, I., Miranda, E., Destercke, S., 2020a. Unifying neighbourhood and distortion models: Part I- new results on old models. Int. J. General Syst. 49, 602–635.

Montes, I., Miranda, E., Destercke, S., 2020b. Unifying neighbourhood and distortion models: Part II- new models and synthesis. Int. J. General Syst. 49, 636–674.

Morais, C., Estrada-Lugo, H., Tolo, S., Jacques, T., Moura, R., Beer, M., Patelli, E., 2022a. Robust data-driven human reliability analysis using credal networks. Reliab. Eng. Syst. Satefy 218.

Morais, C., Ferson, S., Moura, R., Tolo, S., Beer, M., Patelli, E., 2021. Handling the uncertanty with confidence in human reliability analysis. In: Proceedings of the 31st European Safety and Reliability Conference. Research Publishing, Singapore, pp. 3312–3318.

Morais, C., Moura, R., Beer, M., Patelli, E., 2018. Human reliability analysis - accounting for human actions and external factors through the project life cycle. In: Safety and Reliability - Safe Societies in A Changing World. CRC Press, p. 10.

Morais, C., Moura, R., Beer, M., Patelli, E., 2019a. Analysis and estimation of human errors from major accident investigation reports. ASME J. Risk Uncertain. Part B 6, 011014.

Morais, C., Tolo, S., Moura, R., Beer, M., Patelli, E., 2019b. Tackling the lack of data for human error probability with credal network. In: Proceedings of the ESREL.

Morais, C., Yung, K., Johnson, K., Moura, R., Beer, M., Patelli, E., 2022b. Identification of human errors and influencing factors: A machine learning approach. Saf. Sci. 146, 105528.

Moura, R., Beer, M., Patelli, E., Lewis, J., 2015. Human error analysis: Review of past accidents and implications for improving robustness of system design. In: Proceedings of the 24th European Safety and Reliability Conference. Taylor and Francis group, pp. 1037–1046.

Moura, R., Beer, M., Patelli, E., Lewis, J., 2017a. Learning from major accidents: Graphical representation and analysis of multi-attribute events to enhance risk communications. Saf. Sci. 99, 58–70.

Moura, R., Beer, M., Patelli, E., Lewis, J., Knoll, F., 2016. Learning from major accidents to improve system design. Saf. Sci. 84, 37–45.

Moura, R., Beer, M., Patelli, E., Lewis, J., Knoll, F., 2017b. Learning from accidents: Interactions between hyuman factors, technology and organisations as a central element to validate risk studies. Saf. Sci. 99, 196–214.

Mu, L., Xiao, B.P., Xue, W.K., Yuan, Z., 2015. The prediction of human error probability based on bayesian networks in the process of task. In: Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management. pp. 145–149.

Noroozi, A., Khazkad, N., Khan, F., MacKinnon, S., Abbassi, R., 2013. The role of human error in risk analysis: Application to pre-and post-maintenance procedures of process facilities. Reliab. Eng. Syst. Saf. 119, 251–258.

Patelli, E., Tolo, S., George-Williams, H., Sadeghi, J., Rocchetta, R., de Angelis, M., Broggi, M., 2018. Opencossan 2.0: An efficient computation toolbox for risk, reliability and resilence analysis. In: Proceedings of the Joint ICVRAM ISUMA UNCERTAINTIES Conference.

Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, San Mateo, CA.

Pelessoni, R., Vicig, P., Zaffalon, M., 2010. Inference and risk measurement with the pari-mutuel model. Int. J. Approx. Reason. 51, 1145–1158.

Reason, J., 1990. Human Error. Cambridge University Press.
Seidenfeld, T., Wasserman, L., 1993. Dilation for sets of probabilities. Ann. Stat. 21, 1139–1154.
Sträter, O., 2000. Evaluation of Human Reliability on the Basis of Operational Experience. Technical Report GRS, Cologne.
Svenson, O., 1989. On expert judgements in safety analyses in the process industries. Reliab. Eng. Syst. Saf. 25, 219–256.
Swain, A., Guttmann, H., 1983. A Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications. USNRC, NUREGICR-1278, Washington DC 20555.
Tolo, S., Patelli, E., Beer, M., 2018. An open toolbox for the reduction, inference computation and sensitivity analysis of credal networks. Adv. Eng. Software 115, 126–148.
Walley, P., 1981. Coherent Lower (and Upper) Probabilities. Statistics Research Report 22, University of Warwick, Coventry.
Walley, P., 1991. Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London.
Ward, J., Teng, Y.C., Horberry, T., Clarkson, P.J., 2013. Contemporary Ergonomics and Human Factors. Chapter Healthcare Human Reliability Analysis-By HEART. Taylor and Francis, pp. 277–288.
Williams, J.C., 1986. Heart- A proposed method for assessing and reducing human error. In: Proceedings of the Ninth Advances on Reliability Technology Symposium.
Zhang, G., Thai, V., Yuen, K., Loh, H., Zhou, Q., 2018. Addressing the epistemic uncertainty in maritime accidents modelling using Bayesian network with interval probabilities. Saf. Sci. 102, 211–225.