

Development of an algorithm to convert linear belief function inputs to exponential conditional probability functions for multiple method applications

Sean Loughney

Department of Maritime and Mechanical Engineering, Liverpool John Moores University, Liverpool, UK

E-mail: s.loughney@ljmu.ac.uk

Jin Wang

Department of Maritime and Mechanical Engineering, Liverpool John Moores University, Liverpool, UK

E-mail: j.wang@ljmu.ac.uk

Evidential Reasoning (ER), based on the Dempster-Schafer theory of evidence, and Bayesian Networks (BN) are two distinct theories and methodologies for modelling and reasoning with data regarding propositions in uncertain domains. Both ER and BNs incorporate graphical representations and quantitative approaches of uncertainty. BNs are probability models consisting of a directed acyclic graph, which represents conditional independence assumptions in the joint probability distribution. Whereas ER graphically describes knowledge through an evaluation hierarchy and the relationships of the attributes based on Dempster-Schafer theory of belief functions. Therefore, this paper proposes an algorithm, which allows for the conversion of the linear input data of ER (belief degrees and relative weights) to the exponential data input of BNs (conditional probability tables (CPTs)). The algorithm is applied to a validated case study where the ER approach has been utilized for decision-making.

Keywords: Bayesian Networks, Evidential Reasoning, utility ranking, linear data distribution, exponential data distribution, data conversion.

1. Introduction

Evidential Reasoning and Bayesian Networks (BN) are two distinct theories and methodologies for modelling and reasoning with data regarding propositions in uncertain domains (Yang, 2001) (Fenton & Neil, 2013). Both Evidential Reasoning and BNs incorporate graphical representations and quantitative approaches of uncertainty. There are many similarities and differences between these two approaches. It has been argued that the belief function approach (Evidential Reasoning) is a generalisation of Bayesian probability theory calculations and it can be demonstrated that any Bayesian model of uncertainty is also a belief function model (Cobb & Shenoy, 2003). Thus, this research proposes the initial development of an algorithm, which transfers linear input data for Evidential Reasoning to the exponential Conditional Probability Functions of Bayesian Networks.

1.1 Evidential Reasoning

Numerous decision-making problems in management and engineering involve several attributes of both a qualitative and quantitative nature. It is the normal handling of qualitative attributes along with uncertain or incomplete information that causes complexity in multiple

attribute assessments. There has been an increase in the development of theoretically sound methods and tools which deal with Multiple Attribute Decision Analysis (MADA) problems in a coherent, rational, reliable, and repeatable manner (Yang, 2001) (Yang & Xu, 2002) (Chen, et al., 2013). There has been considerable research conducted on integrating techniques from Artificial Intelligence to Operational Research for handling uncertain information. From this line of research, the Evidential Reasoning (ER) approach was developed for MADA. This method of decision-making is based on an evaluation analysis model and the Dempster-Schafer (D-S) theory of evidence. More recently, the ER approach has been applied to decision-making problems in engineering, design and safety and risk assessment and supplier assessment. For example, motorcycle assessment, cargo ship design (Yang & Xu, 2002) and marine system safety analysis (Ren, et al., 2005) (Yang & Xu, 2002) (Chen, et al., 2013). This ER algorithm can be used to aggregate criteria in a multilevel structure, and a rational aggregation process needs to satisfy certain self-evident rules, commonly referred to as synthesis axioms (Yang & Xu, 2002) (Chen, et al., 2013).

Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference

Edited by Piero Baraldi, Francesco Di Maio and Enrico Zio

Copyright © ESREL2020-PSAM15 Organizers. Published by Research Publishing, Singapore.

ISBN/DOI: 978-981-14-8593-0

1.2 Bayesian Networks

BNs are a Directed Acyclic Graph (DAG) encoding Conditional Probability Distribution (CPD). There are two main components to BNs. The first is the graphical structure, which provides the qualitative part and the second is the probability distribution, which provides the quantitative part. The graphical structure is referred to as the DAG. The DAG contains a set of nodes each representing a random/chance variable and each node in the DAG has a number of possible states, which must apply at any one time. Probability distribution indicates the strength of the belief in how the states of parent nodes can affect the states of their child nodes. Either nodes can represent discrete random variables with a finite number of states. For root nodes, a marginal probability table is defined. Non-root nodes are assigned conditional probability tables (CPTs) (Neapolitan, 2004). When constructing a BN it is important to note that the number of permutations in the CPTs increases exponentially with the number of parent nodes and the number of states in the CPT. For example, if node A has 'X' parents with 'n' number of states, then there will be 'Xⁿ' permutations in the CPT or node A. Similarly, the total number of cells in a CPT is equal to the product of the possible number of states in the node and the number of states in the parent nodes. Conditional probabilities are essential to BNs. They can be expressed by statements such as "B occurs given that A has already occurred" and "given event A, the probability of event B is 'p'", which is denoted by $P(A|B)=p$. This specifically means that if event A occurs and everything else is unrelated to event B (except event A), then the probability of B is 'p'. Conditional probabilities are part of the joint probability of the intersection of A and B, $P(A \cap B)$.

1.4 ER with BN

There is very little literature in terms of a method for combining the ER and BN approaches in the manner suggested in this research. This is in terms of applying the same input data and using the two approaches as validation tools for each technique. However, there has been research whereby the results of a BN analysis have been applied to ER to produce further results. For example, Wang *et al.* (2013) developed quantitative accident analysis model by integrating Human Factors Analysis and Classification System (HFACS) with BN, which was utilized to present accident prevention measures. They also proposed the prevention measures in a cost-effectiveness manner through

Best-Fit method and ER approach. Similarly, the use of data from BN to ER has been around for a long time. Cheng and Druzdzel (2000) proposed an adaptive importance-sampling algorithm, AIS-BN, which showed promising convergence rates under extreme conditions and outperformed the existing sampling algorithms from BN models to ER. However, this differs from the research present here as this algorithm applied the same ER input data to BN.

1.3 Rationale

This research was developed as an additional problem from research relating to the development of a MADA methodology for site selection of floating offshore wind farms in the UK and Ireland. The MADA methodology was developed and tested utilising the Evidential Reasoning approach via a case study. This case study was a large offshore region off the northern coast of Scotland. This region was analysed in terms of its suitability for Floating Offshore Wind (FOW) farm implementation. The area itself is 170km East to West ($3^\circ - 6^\circ$ West) by 83km North to South ($58.75^\circ - 59.5^\circ$ North), and was divided into 450 grid squares, each approximately $5.5\text{km} \times 5.5\text{km}$. A grid reference was used to identify each site; A – AD from West to East and 1 – 15 north to South (see Fig. 1). The final result identified 45 sites that were suitable for floating offshore wind implementation, based upon an evaluation hierarchy. Initially this MADA methodology was validated via a benchmark testing with a series of axioms that test the consistent aggregation of data. However, it was felt that another technique could be utilised to provide further validation.

2. Initial Evidential Reasoning Analysis

Before the process of ranking each individual site in the area (see Fig.1) in terms of its

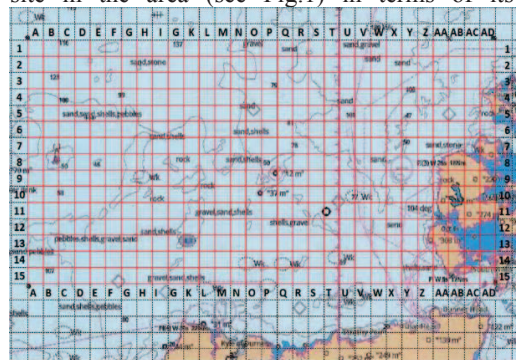


Fig. 1: Large area, north of Scotland for FOW site selection analysis
suitability for FOW implementation, the area

was first be evaluated against an initial set of criteria, specific to the area of the northern coast of Scotland, to determine unsuitable areas. This part of the analysis was mainly qualitative and identifies a range of criteria to initially exclude areas from later evaluation. Similarly, some criteria involve met-ocean data, where areas were excluded if they regularly experience extreme environments, i.e. consistently large waves, or high wind speeds. In order to apply the ER algorithm to the decision of the most suitable site for FOW implementation, a set of variables and a hierarchical structure of general and basic criteria were defined. The variables and hierarchical structure were based upon the initial evaluation criteria but apply a more intricate quantitative approach with an increased number of criteria. In this analysis, there are three general criteria outlined and sixteen basic criteria (see Fig. 2). Once the criteria have been established, a hierarchy must be determined in order to coherently develop a solution to the problem. This hierarchy groups certain criteria under one general criterion. This allows for a smaller number of criteria to be aggregated gradually to reduce the calculation complexity of the decision-making algorithm. Objective data was then collected for each of the 16 basic criteria. This data consists of a set of 5 belief degrees for each criteria for each of the 45 sites under evaluation, and a weight for each of the criteria. Each site is evaluated under each criteria given its suitability on a scale of *Poor*, *Indifferent*, *Average*, *Good* and *Excellent*. These beliefs are combined with the relative weights of

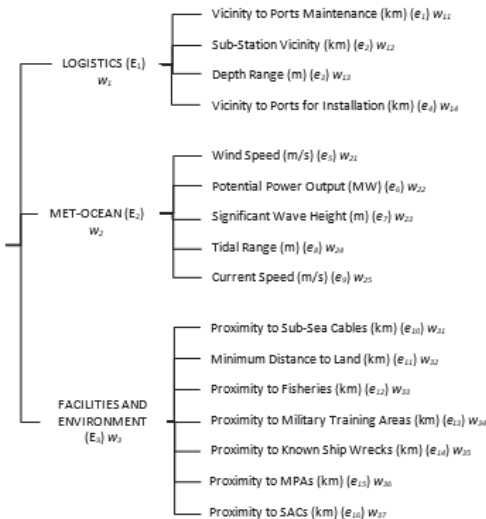


Fig. 2: Evaluation Hierarchy for the available sites for FOW off the northern coast of Scotland

the criteria and aggregated to produce a final set of beliefs for each site. An example of the belief degrees and weights is demonstrated in Table 1, which shows the data for the general criterion *Logistics* for site A15 (see Fig. 1 and Fig. 2). The weights are not shown as they are all assumed to be equal for this research in this paper. Forty-five sites were determined from the application of the initial exclusion criteria are ranked based upon their performance in the decision-making analysis. The overall belief degrees of each site were determined and ranked in terms of their suitability for FOW implementation. These utility intervals are ranked from greatest to smallest. The site with the greatest value is the most suited for FOW implementation. The utility ranking method is to be outlined here but for further information regarding the application of the ER algorithm can be found at the following references (Yang, 2001) (Yang & Xu, 2002) (Wang, et al., 1995) (Sadeghi, et al., 2018).

Table 1: An example of a generalised decision matrix for site selection assessment with basic criteria belief degrees

Basic Criteria	Beliefs Site A15	Evaluation Grade	Grading Scale
Vicinity to Ports Maintenance (km) (e_1)	0	Poor	$\leq 700\text{km}$
	0.6	Indifferent	$500\text{km} \geq x < 700\text{km}$
	0.2	Average	$300\text{km} \geq x < 500\text{km}$
	0.2	Good	$100\text{km} \geq x < 300\text{km}$
Sub-Station vicinity (km) (e_2)	0	Excellent	$>100\text{km}$
	0	Poor	$\geq 175\text{km}$
	0	Indifferent	$125\text{km} \geq x < 175\text{km}$
	0	Average	$75\text{km} \geq x < 125\text{km}$
Depth (m) (e_3)	1	Good	$25\text{km} \geq x < 75\text{km}$
	0	Excellent	$<25\text{km}$
	0	Poor	$<50\text{m}$
	0	Indifferent	$50\text{m} \leq x < 100\text{m}$
Vicinity to Ports for Installation (km) (e_4)	0.5	Average	$100\text{m} \leq x < 150\text{m}$
	0.5	Good	$150\text{m} \leq x < 250\text{m}$
	0	Excellent	$\geq 250\text{m}$
	0	Poor	≤ 700
Installation (km) (e_4)	0.67	Indifferent	$500 \geq x < 700\text{km}$
	0	Average	$300\text{km} \geq x < 500\text{km}$
	0	Good	$100\text{km} \geq x < 300\text{km}$
	0.33	Excellent	$>100\text{km}$

Suppose the utility of an evaluation grade, H_n , is denoted by $u(H_n)$. The utility of the evaluation grade must be determined beforehand, with $u(H_1)=0$ and $u(H_5)=1$ assuming there are five evaluation grades (Yang, 2001). If there a lack of information available then the values of $u(H_n)$ can be assumed to be equidistant, as shown by Eq. (1):

$$u(H_n) = \{u(H_1) = 0; u(H_2) = 0.25; u(H_3) = 0.5; u(H_4) = 0.75; u(H_5) = 1\} \quad (1)$$

The estimated utility for the general and basic criteria, $S(z(e_i))$, given the set of evaluation grades is given by Eq. (2) (Yang & Xu, 2002):

$$u(S(z(e_i))) = \sum_{n=1}^N u(H_n)\beta_n(e_i) \quad (2)$$

In Eq. (2) the term $\beta_n(e_i)$ determines the lower bound of the likelihood, that e_i is assessed to a grade H_n (Yang, 2001). For the comparison in this research, the performance results from ER for the sites under the criterion *Logistics* will be used. Table 2 outlines the ranking order and utility values of the 45 sites under this criterion.

Table 2: Utility assessment results and ranking order of the sites in northern Scotland for the general criterion *Logistics*

Site	Rank	Utility Value	Site	Rank	Utility Value
T7	1	0.566	E15	23	0.500
A15	2	0.564	F15	23	0.500
B15	3	0.556	E14	23	0.500
C15	3	0.556	F14	23	0.500
A14	3	0.556	D13	23	0.500
B14	3	0.556	E13	23	0.500
G7	7	0.542	F13	23	0.500
K5	7	0.542	D12	23	0.500
L5	7	0.542	E12	23	0.500
M5	10	0.534	F12	23	0.500
T6	11	0.519	F11	23	0.500
T3	11	0.519	G8	23	0.500
U3	11	0.519	T5	35	0.486
W2	11	0.519	R4	35	0.486
J5	15	0.515	S4	35	0.486
D15	16	0.515	T4	35	0.486
C14	16	0.515	U4	35	0.486
D14	16	0.515	S3	35	0.486
A13	16	0.515	T2	41	0.483
B13	16	0.515	U2	41	0.483
C13	16	0.515	W1	41	0.483
S5	22	0.511	X1	41	0.483
			R3	45	0.453

3. Case Study

It was proposed that the initial input data for the Evidential Reasoning belief function model could be adapted to the conditional probability functions within a Bayesian Network. The same beliefs and weights utilised in the aggregation assessment of the floating offshore wind sites is to be used as a test case to develop the conversion algorithm (Cobb & Shenoy, 2003). One section of the evaluation hierarchy is to be assessed to test the application of ER belief functions to BN conditional data functions. The criterion *Logistics* and its basic criteria was to be used as a test case for ranking all 45 sites via BNs.

3.1 Adaptation on the weighted sum algorithm

The development of the algorithm, to convert linear to exponential functions, has a base in the symmetric method of building large Conditional Probability Tables (CPT). This symmetric method works off the basis that there is a relative weight for each parent in the BN and that there is a probability of occurrence for the child node given that one parental event occurs. The symmetric method combines these two parameters to form the weighted sum algorithm. This algorithm applies a weight to each of the parent nodes and produces each individual conditional probability in the CPT by combining the compatible probabilities, given the state of the child node, and the relative weights of the parent nodes, as shown in Eq. (3).

$$P(E_N^l | e_1^n, e_2^n, \dots, e_i^n) = \sum_{j=1}^n w_j \cdot \beta_{n,i} \quad (3)$$

$$l = 0, 1, \dots, m \quad S_j = 1, 2, \dots, k_j$$

Where w_j represents the weight of the parent and $\beta_{n,i}$ is the probability of the child given the occurrence of one individual parent of weight w_j , and $E_N^l | e_1^n, e_2^n, \dots, e_i^n$ represents the final conditional probability of a given parental configuration. Eq. (3) demonstrates that the probability of event D , in state l , is given by the probability of the parents in the states j of the set of states' S . In this algorithm the set of states, S , is vital, as the set of states for each parent must be the same i.e.: "Yes" and "No". This conditional probability of D^l is given by the sum of the product of the relative weights of each parent and each linear compatible parental configuration. For further information on the symmetric method and the weighted sum algorithm see the following sources (Das, 2008), (Loughney & Wang, 2017) (Loughney *et al.* 2018). This algorithm was applied to the ER belief degrees and relative weights to complete the large CPTs generated by having a number of parent nodes with five states (belief degrees) into a child node with five states. Thus, the weighted sum algorithm was applied to complete this CPTS using the equal weights of 0.25 for each of the four basic criteria and the belief degrees for each criteria for each of the 45 sites (see Table 1 for an example of the data for one site, A15). Fig. 3 shows the BN for the *Logistics* criteria, where each site has an individual BN. Due to the vast number of iterations and calculations the results had to be normalised in order for the data to fulfil the requirements for a CPT. i.e. each column should sum to one. Table 3 shows an example of five parental configurations where the weighted sum algorithm is applied (top

section), then the data is filtered based upon which parents are affecting the probability (middle Section), the data is then normalised to a CPT (bottom Section). The issue here is that once the data is normalised the CPT is not representative of reality. This is because the input data produces identical results and an equal distribution regardless of whether belief degrees are higher or lower.

3.2. Algorithm modification

As data was not available to calculate the relative weights of the basic criteria (parents), an

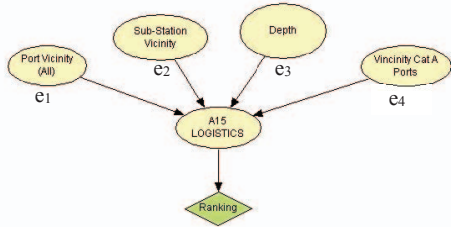


Fig. 3: BN for the criterion logistics and its basic criteria

alternative solution was devised. This solution was designed to combat the fact that the CPT was not representative of reality while still utilising the available data. Thus the weighted sum algorithm was modified to include an additional exponent. This modification is resented by the factor α_l . Eq. (4) outlines the algorithm for the symmetric method, where the conditional probability is calculated following the normalisation, as demonstrated by Eq. (5). The same notations for the criteria and the weights has been applied in the BN analysis as with the ER analysis. In Eq. (4) w_i ($i = 1 \dots n$) is the weight of the general criteria, w_{ij} ($j = 1 \dots n$) is the weight of the basic criteria, E_N is the notation for the general criteria and e_i is the notation for the basic criteria.

Table 3: Example of CPT generation utilising the weighted sum algorithm.

Parent Nodes	Parental Distribution				
e4	5	5	5	5	5
e3	2	2	2	2	2
e2	2	2	2	2	2
e1	1	2	3	4	5
States (grades)	Weighted Sum Algorithm application				
Poor (1)	0.08	0.23	0.13	0.13	0.08
Indifferent (2)	0.08	0.23	0.13	0.13	0.08
Average (3)	0.08	0.23	0.13	0.13	0.08
Good (4)	0.08	0.23	0.13	0.13	0.08
Excellent (5)	0.08	0.23	0.13	0.13	0.08

	Filtering of data relative to parental configuration				
Poor	0.08	0.00	0.00	0.00	0.00
Indifferent	0.08	0.23	0.13	0.13	0.08
Average	0.00	0.00	0.13	0.00	0.00
Good	0.00	0.00	0.00	0.13	0.00
Excellent	0.08	0.23	0.13	0.13	0.08
	Normalising and final CPT				
Poor	33.33%	0%	0%	0%	0%
Indifferent	33.33%	50%	33.33%	33.33%	50%
Average	0%	0%	33.33%	0%	0%
Good	0%	0%	0%	33.33%	0%
Excellent	33.33%	50%	33.33%	33.33%	50%

$$E_N^l | e_1^n, e_2^n, \dots, e_i^n = \frac{\left(\sum_{i=1}^n w_{ij} \cdot \beta_{n,i} \right)}{\alpha_l} \quad (4)$$

$$P(E_N^l | e_1^n, e_2^n, \dots, e_i^n) = \frac{E_N^l}{\sum_{m=1}^z (E_N^m)} \quad (5)$$

where,

$$\sum_{m=1}^z (E_N^m) = E_N^1 + E_N^2 + \dots + E_N^l \quad l = 1, \dots, Z$$

Eq. (4) and Eq. (5) demonstrate that the probability of event $E_N(P(E_N^l))$ represents the probability of the general attribute E_N in the state l , given various parental combinations, $e_1^n, e_2^n, \dots, e_i^n$, where n is some state of the parent e_i . This conditional probability of $P(E_N^l)$ is given by the sum of the product of the relative weights, w_{ij} , of each basic criteria (parent) and each belief degree given each evaluation grade, $\beta_{n,i}$. These values are then divided by a factor α_l . This factor is dependent on the state of the child and is calculated by the sum of the identical evaluation grades for each basic criterion, then normalised. Initially the sum of the identical evaluation grades is calculated, *i.e.* the belief degree of the grade *excellent* for each basic criteria are added together. This process is demonstrated by Eq. (6) by initially calculating A_l , then normalising the results of A_l to determine α_l .

$$A_l = \sum_{i=1}^z (\beta_{n,i}) \quad l = 1 \dots Z, i = 1 \dots n \quad (6)$$

where $\beta_{n,i}$ denotes the complete belief degree of an attribute given a certain evaluation grade. In Eq. (5), $z=n$ and denotes the belief degrees, in a given evaluation grade, that are to be summarized. For example, take the belief degrees for the basic attributes e_1, e_2, e_3 , and e_4 as shown in Fig. 3 and Table 4. If $l=5$, then $n=5$ as $n=l$, hence, the fifth evaluation grade, *excellent*, is to be assessed for each basic attribute where $i=1 \dots L$. This example can be expressed as follows:

$$A_5 = \sum_{\substack{l=1 \\ n=l \\ i=1}}^5 (\beta_{5,i}) = \beta_{5,1} + \beta_{5,2} + \beta_{5,3} + \beta_{5,4}$$

$$= 0 + 0 + 0 + 0.33$$

$$= 0.33 \text{ (excellent)}$$

Where, A_5 is the summation of all of the evaluation grades that are *excellent* across all basic criteria. Eq. (6) is applied to all permutations for all possible values of n and i . These values are then normalised to determine α_i . This is determined by Eq. (7)

$$\alpha_i = \frac{A_i}{\sum_{m=1}^n (A_m)} \quad (7)$$

where,

$$\sum_{m=1}^n (A_m) = A_1 + A_2 + \dots + A_n \quad m = 1, \dots, n$$

Therefore, the various values of α_i for site A15 for the CPT of the general attribute *Logistics* are outlined in Table 4. The algorithm demonstrated in Eq. (4) does not produce the final CPT. Applying Eq. (4) results in a series of values for every given state of the child given the occurrence of every possible relevant combination of state of the parent nodes. These values are filtered to only include values from the parental configuration for that state in the child (see the middle section of Table 3 and Table 5). These occurrence probabilities then need to be normalised in order for the individual columns in the CPTs to sum to 1. This is vital in order for the BN to be coherent and provide consistent results. This process is demonstrated in Table 5 through an example of a section of the CPT for the child node *logistics* (E_i) for Site A15. As the factor α_i has been outlined, an example of Eq. (4) is applied in this case for the state of the child node *logistics*, E_1 , given the parental distribution 1, 2, 2 and 5. In this calculation, only the values pertaining to states 1, 2 and 5 (E_{11} , E_{12} , and E_{15}) are applied, as states 3 and 4 are not part of the configuration. It should also be noted that the value for state 1 is zero as all the belief degrees for the state *Poor* are all zero thus the factor α_1 is also zero. These values are represented by the middle section of Table 5.

In this calculation, only the values pertaining to states 1, 2 and 5 (E_{11} , E_{12} , and E_{15}) are applied, as states 3 and 4 are not part of the configuration. It should also be noted that the value for state 1 is zero as all the belief degrees for the state *Poor* are all zero thus the factor α_1 is also zero. These values are represented by the middle section of Table 5. However, these values do not complete the CPT, as the individual columns must sum to one. Therefore, the occurrence probabilities are

normalised using Eq. (5). Table 5 outlines the steps in the process in a similar fashion to Table 3. The top section shows the application of the weighted sum algorithm with the factor α_i . The middle section then shows the filtering of the data for values that correspond to the configuration of the parental states. Finally, the bottom section shows the normalisation and the final values for the CPT.

Table 4: Example of the results generated from Eq. (6) and Eq. (7)

ei	Beliefs (n)	$\beta(n,i)$	Ai	α_i	α_i	
e1	0	$\beta(1,1)$	A1 (Poor)	0	α_1	0
	0.6	$\beta(2,1)$				
	0.2	$\beta(3,1)$	A2 (Indifferent)	1.27	α_2	0.32
	0.2	$\beta(4,1)$				
	0	$\beta(5,1)$	A3 (Average)	0.7	α_3	0.175
e2	0	$\beta(1,2)$	A4 (Good)	1.7	α_4	0.425
	0	$\beta(2,2)$				
	0	$\beta(3,2)$	A5 (Excellent)	0.33	α_5	0.08
	1	$\beta(4,2)$				
	0	$\beta(5,2)$	$\sum_{m=1}^n (A_m)$	4		
e3	0	$\beta(1,3)$	A6 (Poor)	0	α_6	0
	0	$\beta(2,3)$				
	0.5	$\beta(3,3)$	A7 (Indifferent)	0.317	α_7	0.083
	0.5	$\beta(4,3)$				
	0	$\beta(5,3)$	A8 (Average)	0.317	α_8	0.083
e4	0	$\beta(1,4)$	A9 (Poor)	0	α_9	0
	0.67	$\beta(2,4)$				
	0	$\beta(3,4)$	A10 (Indifferent)	0.317	α_{10}	0.083
	0	$\beta(4,4)$				
	0.33	$\beta(5,4)$	A11 (Excellent)	0.33	α_{11}	0.083

3.3. Discussion

It can be seen in Table 5 that the calculated CPTs are very much affected by this new factor, α_i . For example, take the parental combination of 1, 2, 2 and 5 and cross examine it in terms of the input data (belief degrees, the relative weights, and the factors, α_i). As the relative weights are equal, lets focus on the belief degrees and normalisation factors, where the beliefs and α_i values of the parental combination 1, 2, 2 and 5 are; (0, 0, 0, 0.33) and (0, 0.317, 0.175, 0.425, 0.083) respectively. Given the input data, it is safe to assume that the most likely outcome of the combination in the CPT would be state five (excellent) this is because the input belief degrees are all zero with the exception of excellent (0.33).

Table 5: Example of CPT generation utilising the weighted sum algorithm with the factor α_i .

Parent Nodes	Parental Distribution			
e4	5	5	5	5
e3	2	2	2	2
e2	2	2	2	2

e1	1	2	3	4	5
States (grades)	Weighted Sum Algorithm application				
Poor (1)	0.00	0.00	0.00	0.00	0.00
Indifferent (2)	0.26	0.74	0.42	0.42	0.26
Average (3)	0.48	1.33	0.76	0.76	0.48
Good (4)	0.20	0.55	0.31	0.31	0.20
Excellent (5)	1.00	2.80	1.60	1.60	1.00
	Filtering of data relative to parental configuration				
Poor	0.00	0.00	0.00	0.00	0.00
Indifferent	0.26	0.74	0.42	0.42	0.26
Average	0.00	0.00	0.76	0.00	0.00
Good	0.00	0.00	0.00	0.31	0.00
Excellent	1.00	2.80	1.60	1.60	1.00
	Normalising and final CPT				
Poor	0.00%	0%	0%	0%	0%
Indifferent	20.83%	21%	15.13%	18.03%	21%
Average	0%	0%	27.38%	0%	0%
Good	0%	0%	0%	13.44%	0%
Excellent	79.17%	79%	57.49%	68.53%	79%

However, when one also considers the effect of the normalising factor, α_1 , it can be seen that state 1 (*poor*) has zero influence, therefore the probability of the child being in the state poor given the combination of 1, 2, 2 and 5, is 0%. Similarly, state two (*indifferent*) has some influence, as its normalising factor here is 0.317, despite its belief being zero. Therefore, there should be some value of occurrence probability for the child node in state two, as demonstrated by the previous calculation. Finally, as the belief of state five for the parent node is 0.33 and the normalising factor is 0.083, this should have the greatest influence as the other two states (1 and 2) had belief degrees equal to zero. This hypothesis is reflected in the CPT where for the parental configuration, 1, 2, 2 and 5, the conditional probabilities of the child node are poor (0%), indifferent (20.83%) and excellent (79.17%). Here the states 3 and 4 (average and good) are automatically zero, as they do not feature in the parental configuration. All other CPTs for each general attribute and overall assessment for each site follow this hypothesis. This provides some verification that the CPTs and the subsequent BNs have been constructed to represent reality as efficiently as possible given the input data.

3.4. Comparison to ER

The key now is to compare the results of the ER analysis with the BN analysis, following the completion of 45 BNs using the outlined method (one BN for each site). Table 6 shows the differences in ranking position when the ranking order of both analyses of the sites are compared. The first column in Table 6 shows the ranking order of the sites following the ER analysis. The second column shows the difference in position when compared to the results of the BN analysis. A Positive number indicates an increase in

position; conversely, a negative number indicates a decrease in position.

Table 6: Comparison of the difference in ranking order of the sites from the ER and BN analyses

ER Rank	BN Difference	ER rank	BN Difference
T7	-7	F15	9
A15	1	E14	8
B15	-30	F14	7
C15	-31	D13	6
A14	-32	E13	5
B14	-33	F13	4
G7	-17	D12	-2
K5	-30	E12	3
L5	-31	F12	2
M5	-22	F11	1
T6	2	G8	0
T3	1	T5	32
U3	0	R4	31
W2	-1	S4	30
J5	-29	T4	29
D15	-10	U4	28
C14	-11	S3	-4
D14	-12	T2	1
A13	-13	U2	0
B13	-14	W1	-1
C13	-15	X1	-2
S5	20	R3	0
E15	10		

It can be seen from Table 6 that there is a considerable amount of fluctuation regarding the ranking of sites. It can be seen that 33.33% of the sites demonstrate a change in rank by $\leq \pm 3$ positions. Whereas 66.67% demonstrate a change in rank by $> \pm 3$ positions. Similarly, the position of the sites within the ER does not have a major effect on the ranking position of the BN analysis. Sites A15 at the top changes by only one position in the ranking, from second in ER to first in the BN. The sites ranked from ninth to twelfth also demonstrate very little change in position, as do the five lowest ranked sites. The largest decrease in position is demonstrated by site B14 that goes from sixth in the ER ranking to thirty-sixth in the BN ranking, a decrease in thirty positions. Similarly, the largest increase is demonstrated by site T5. This sites ranking goes from thirty-fifth in the ER ranking to third in the BN ranking, an increase of thirty-two places.

4. Conclusion & future work

This research set out to demonstrate if the linear data functions (ER) could be transferred and applied to exponential data functions (BN). Thus, a modified weighted sum algorithm was developed to satisfy this problem. The reasoning behind the development was fuelled by; i) potentially utilising BNs as a validation tool for ER and ii) the fact that ER can be regarded as a generalisation of Bayesian theory and a BN could be demonstrated as a belief function mode. The weighted sum algorithm is an accepted

method of compiling large CPTs and provided a solid basis for development as it incorporates occurrence probabilities and weights of criteria or parent nodes. Similarly, the ER approach utilises these two factors in its aggregation. The algorithm was modified and applied to a test case where ER had already been extensively applied. The results, as shown, were not completely consistent. However, there are some sections in the results, which may indicate that further work that can be done to optimise this method. This research focused on one specific section of the evaluation hierarchy in Fig. (2). Hence, the next step is to apply this to the complete ER aggregation for all of the criteria shown in Fig. (2) and all of the 45 sites. Secondly, this research assumed equal weighting for the criteria in the assessment. A further step would be to calculate these weights through objective or subjective analysis. It makes sense that varying weights would have an effect on the ranking outcomes. Finally, to conduct the analysis in the opposite direction. In other words, gather data under the assumption that the analysis is carried out through a BN first and utilise that data for the ER aggregation.

Acknowledgements

This research is supported by Liverpool John Moores University. Thanks are also given to the EU for its financial support under a European Commission funded project ARCWIND 2018 – 2020 (EAPA_344/2016)

References

Balaram, D., 2008. Generating conditional probabilities for Bayesian Networks: Easing the knowledge acquisition problem. Edinburgh : Command and Control Division, DSTO, Australia,

Cheng, J. & Druzdzal, M. J., 2000. AIS-BN: An Adaptive Importance Sampling Algorithm for Evidential Reasoning in Large Bayesian Networks. *Journal of Artificial Intelligence Research*, Volume 13, pp. 155-188.

Chen, S., Liu, J., Wang, H. & Augusto, 2013. An evidential reasoning based approach for decision making with partially ordered preference under uncertainty. Tianjin, Proceedings of the 2013 International Conference on Machine Learning and Cybernetics.

Cobb, B. R. & Shenoy, P. P., 2003. A Comparison of Bayesian and Belief Function Reasoning. *Information Systems Frontiers*, 5(4), pp. 345 - 358.

Fenton, N. & Neil, M., 2013. Risk Assessment and Decision Analysis with Bayesian Networks. 1st ed. s.l.:Taylor & Francis Group.

Loughney, S., & Wang, J., 2017. Bayesian network modelling of an offshore electrical generation system for applications within an asset integrity case for normally unattended offshore installations, *Proc IMechE Part M: J Engineering for the Maritime Environment*, 232(4), pp. 402-420.

Loughney, S., Wang, J., & Matellini, B., 2018. Utilising Bayesian Networks to demonstrate the potential consequences of a fuel gas release from an offshore gas driven turbine, *Proc IMechE Part M: J Engineering for the Maritime Environment*, online: December 7, 2018, pp. 1 - 21.

Matellini, D. B. et al., 2013. A study of human reaction during the initial stages of a dwelling fire using a bayesian network model. *Proc. IMechE Part O: Journal of Risk and Reliability*, 227(2), pp. 207 - 221.

Matellini, D. et al., 2013. Modelling Dwelling Fire Development and Occupancy Escape using Bayesian Network. *Reliability Engineering and System Safety*, Volume 114, pp. 75 - 91.

Ren, J. et al., 2005. an offshore safety assessment framework using fuzzy reasoning and evidential synthesis approaches. *Journal of Marine Engineering and Technology*, 4(1), pp. 3 - 16.

Sadeghi, A., Farhad, H., Moghaddam, A. M. & Qazizadeh, M. J., 2018. Identification of accident-prone sections in roadways with incomplete and uncertain inspection-based information: A distributed hazard index based on evidential reasoning approach. *Reliability Engineering & System Safety*, Volume 178, pp. 278 - 289.

Wang, F. Y., Xie, M., Chin, K. & Fu, X. J., 2013. Accident analysis model based on Bayesian Network and Evidential Reasoning approach. *Journal of Loss Prevention in the Process Industries*, Volume 26, pp. 10-21.

Wang, J., B., Y. J. & Sen, P., 1995. Safety analysis and synthesis using fuzzy sets and evidential reasoning. *Reliability Engineering & System Safety*, 47(2), pp. 103 - 118.

Yang, J., 2001. Rule and Utility based evidential reasoning approach for multiattribute decision analysis under uncertainties. *European Journal of Operational Research*, Issue 131, pp. 31 - 61.

Yang, J. & Xu, D., 2002. On the Evidential Reasoning Algorithm for Multiple Attribute Decision analysis Under Uncertainty. *IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans*, 32(3), pp. 289 - 304.