

Mapping Conditional Scenarios for Knowledge Structuring in (Tail) Dependence Elicitation

Christoph Werner^a, Tim Bedford^a and John Quigley^a

^aUniversity of Strathclyde, Department of Management Science, Glasgow, G4 0QU, UK

ARTICLE HISTORY

Compiled December 3, 2019

ABSTRACT

In decision and risk analysis, probabilistic modelling of uncertainties provides essential information for decision-makers. As uncertainties are typically not isolated and simplifying assumptions (such as independence) are often not justifiable, methods that model their dependence are being developed. A common challenge is that relevant historical data for specifying and quantifying a model are lacking. In this case, the dependence information should be elicited from experts. Guidance for eliciting dependence is sparse whereas particularly little research addresses the structuring of experts' knowledge about dependence relationships prior to a quantitative elicitation. However, such preparation is crucial for developing confidence in the resulting judgements, mitigating biases and ensuring transparency, especially when assessing tail dependence. Therefore, we introduce a qualitative risk analysis method based on our definition of conditional scenarios that structures experts' knowledge about (tail) dependence prior to its assessment. In an illustrative example, we show how to elicit conditional scenarios that support the assessment of a quantitative model for the complex risks of the UK higher education sector.

KEYWORDS

Uncertainty Modelling; Stochastic Systems; Dependence Modelling; Structured Expert Judgement; Cognitive Mapping

1. Introduction

Probabilistic modelling of uncertainties is of key importance for many decisions as well as in risk analysis problems. It supports decision-makers in understanding better the inherent uncertainties of a model and the decision(s) based on it. The common simplifying assumption of independence between uncertainties is, however, often not reasonable. Therefore, a robust analysis needs to account for the possible (often unexpected) effects that dependence relationships might cause. Despite advances in mathematical modelling of dependence, its elicitation is still much under development. Yet, whenever facing a lack of relevant historical data, having experts to assess the required dependence information is the only appropriate option for model quantification without introducing unjustified modelling assumptions.

While extensive guidance is available on formal processes for eliciting univariate uncertainties (Cooke, 1991; Dias et al., 2018; EFSA, 2014; Hanea et al., 2017; O'Hagan et al., 2006; Quigley et al., 2018), this is not the situation for the multivariate case.

Overviews on eliciting dependence for probabilistic modelling are given in Werner et al. (2017) and Werner et al. (2018a).

One of the most neglected parts of research in formal elicitation processes for multivariate uncertainty is that of structuring experts' knowledge about a joint distribution prior to its quantitative assessment. Nevertheless, this is essential for ensuring confidence in the elicitation, supporting transparency and reproducibility of the expert judgement study as well as mitigating experts' potential cognitive fallacies.

The few methods which specifically structure knowledge about joint distributions, are part of hybrid approaches to dependence modelling that specify qualitative relationships first (usually graphically), before assessing them probabilistically. These approaches comprise Knowledge Maps (Howard, 1989), Event and Fault Trees (Bedford and Cooke, 2001), Influence Diagrams (Howard and Matheson, 2005; Shachter, 1988) and Bayesian Belief Networks (Pearl, 1988). While these methods enjoy popularity in the decision and risk analysis literature due to their convenient graphical representation, they are not suitable in various modelling contexts which might be of interest. For example, they do not address potential tail dependencies as these models in themselves do not capture main characteristics of systems where extreme value distributions emerge, such as underlying vicious cycles (or reinforcing loops). Moreover, due to their hybrid nature, these models often require a large number of assessments for quantifying the underlying joint distribution.

In order to address knowledge structuring for joint distributions in a more flexible manner and also to separate the structuring process from the quantitative assessment, in particular for models which incorporate tail dependencies, we present a new structuring method in this paper which is based on conditional scenarios (as defined later). In doing so, we are in line with the arguments in favour of "sense-making" through qualitative scenarios when modelling uncertainty (French, 2015), whereas such a representation of a simplified part of reality has also been termed "small world" (Savage, 1954).

As far as we are aware this is the first method to specifically structure scenarios of (extreme) conditional distributions for use prior to experts' (tail) dependence assessments. There is a wide literature on scenarios and problem structuring, but for our purpose of supporting a dependence elicitation we build on findings and approaches from various research areas of decision and risk analysis. These include:

- *Probabilistic Risk Analysis* (PRA) (Bedford and Cooke, 2001): Fault and Event Tree Analysis (Ericson, 2015; Vesely et al., 1981), Root Cause Analysis (Wilson, 1993), Structured Expert Judgement (Bedford and Cooke, 2001), Multivariate Uncertainty Modelling (Bedford and Cooke, 2001), Bow-Tie Modelling (Ale et al., 2006)

together with research in *Risk Perception* (Renn, 1992)

- *Problem Structuring* (PSM or "soft OR") (Franco, 2006; Mingers and Rosenhead, 2004; Rosenhead, 2006): Cognitive Mapping (Ackermann and Eden, 2006) in particular as used for mapping systemic risks (Ackermann et al., 2014), Causal Loop Diagrams (Morecroft, 2010)
- *Scenario Planning* (SP) (Bradfield et al., 2005): Intuitive Logics (Van der Heijden, 2011)

A more detailed list is provided in Section 4.

The remainder of this paper is as follows. First (in Section 2), we define the variables of

interest for a quantitative dependence elicitation before we present a way to elicit them for a common model that can capture tail dependence. This shows the assessments that experts are required to make. In Section 3, we propose our definition of conditional scenarios and derive the main desiderata for a method that supports experts in structuring their knowledge on joint distributions through scenarios. Next (in Section 4), we discuss the applicability of existing problem structuring, scenario planning and qualitative risk analysis methods for our purpose and by that outline the features we require for our method. We introduce our method in Section 5. After that we then present an illustrative example in Section 6 which shows how the method was applied to structure and model tail dependencies in the UK higher education sector. Finally, we conclude the paper in Section 7 with a reflective discussion on the method’s current shortcomings and potential future enhancements.

2. Quantitative elicitation of dependence

Before proposing a definition for conditional scenarios and presenting our method, in this section we briefly discuss the quantitative assessment part of an elicitation process. This illustrates the quantitative judgements experts are asked to make and for which the method proposed in this paper aims to support them.

2.1. Tail dependence models and variables of interest for elicitation

Typically, the variables of interest in dependence elicitation (i.e. the elicited quantities) serve as inputs to some multivariate model. The form in which the information is elicited, e.g. through a conditional probability or a correlation coefficient, is chosen by an analyst who manages the elicitation process. Together with the underlying scenarios (the experts’ rationales) that determine a specific assessment, the value of an elicited quantity constitutes the knowledge (and type of information) we want to obtain from experts.

While Werner et al. (2017) discuss the prevalence of dependence in different modelling contexts in decision and risk analysis and its influence of how to model and assess it, a common mathematical representation of dependence is a collection of random variables describing measurable risk characteristics (McNeil et al., 2015), such as the number of lives lost, monetary losses and so forth. Formally (in the bivariate case), these random variables are denoted by X and Y , $F_{X,Y} : \mathbb{R}^2 \rightarrow [0, 1]$. For dependence, we are interested in the distribution of the random vector described by the joint probability distribution $F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$. Assuming probabilistic independence greatly simplifies the modelling process as we simply use the product of the marginal distributions to determine the multivariate distribution. However, whenever this assumption is not sensible, it follows that $F_{X,Y}(x, y) \neq F_X(x)F_Y(y)$. In this case, we need a dependence model that approximates the unknown distribution by capturing the most important features of the dependence relationship. For instance in the later case-study application, we are concerned that the dependence structure in the central part of a model might not be representative of the structure in the tails. Therefore, we include tail dependence, i.e. the strength of association in the tails of the joint distribution, if applicable. Neglecting this aspect of a random variables can lead to distorted model output and hence poor decision making. For example, MacKenzie and Spears (2014) discuss common (false) modelling assumptions in financial mathematics and their impact by reflecting on the famous *Wired* article about “the formula that

killed Wall Street” (Salmon, 2009).

Tail dependence can be modelled with a copula. For an introduction and discussion on the topic, see Joe (2014). Recall, we can decompose any multivariate distribution function into its univariate margins and a copula. This can be reversed in order to construct new multivariate distribution functions with a given copula, so that a convenient modelling feature is the separate treatment of the marginal distributions and the dependence relationship. Various common parametric copulas can be grouped into classes. For instance, *Elliptical* copulas are radially symmetric, i.e. their upper and lower tail dependence is the same, whereas *Archimedean* copulas do not show this symmetry. This is an important modelling property as for the former, large losses always occur together with large gains which in many situations might not be a realistic dependence characteristic. For example, Crook and Moreira (2011) show how copulas can be used to appropriately model asymmetric dependence of joint high default rates in a credit card portfolio.

Formally, lower tail dependence (which is of interest in our illustrative case-study) for the distribution functions F_X and F_Y of random variables X and Y is defined as:

$$\lambda_L(Y|X) = \lim_{u \rightarrow 0} P(Y \leq F_Y^{-1}(u) | X \leq F_X^{-1}(u))$$

when a limit $\lambda_L \in [0, 1]$ exists. Whenever $\lambda_L > 0$, X and Y are dependent in the lower tail whereas whenever $\lambda_L = 0$ they are tail independent. In other words, in the tail dependent case one is more likely to observe low values for Y given low values for X . From that, we can distinguish various copula types through their lower tail dependence coefficient $\lim_{u \rightarrow 0} \frac{C(u,u)}{u}$ (see Joe (2014) for its derivation) of which we make use in the elicitation as shown in the next section.

2.2. Tail dependence elicitation for parametric copulas

While an elicitation can be structured in various ways, we briefly present a method that allows for the explicit consideration of tail dependencies. Furthermore the formal definitions of tail dependence highlight the need for a structuring process that can help with elicitation.

The process below is a pragmatic solution to eliciting dependence information for choosing a copula that represents an experts belief. More specifically, it allows for distinguishing common parametric forms. Together with other elicitation methods it is discussed in Werner et al. (2017) in more detail while Werner et al. (2018b) propose a more flexible elicitation method which allows for considering non-parametric forms and which is also compatible with the scenario mapping method presented in this paper. Note, the below is only a brief description focussing on the actual elicitation, neglecting elements of pre- and post-elicitation, such as training experts and aggregating judgements:

- (1) The marginal distributions are specified either through historical data or an expert judgement method for univariate quantities, e.g. Cooke (1991); Quigley et al. (2018) or O’Hagan et al. (2006).
- (2) With (1) in place, we elicit various conditional exceedance probabilities¹ in

¹In the case-study of this paper we elicit conditional exceedance probabilities. Briefly, a conditional exceedance probability is the probability of a variable *exceeding* (or equalling) a certain percentile given that the variable we condition on also exceeds a certain percentile. The definition is the same for both variables being below a certain percentile threshold.

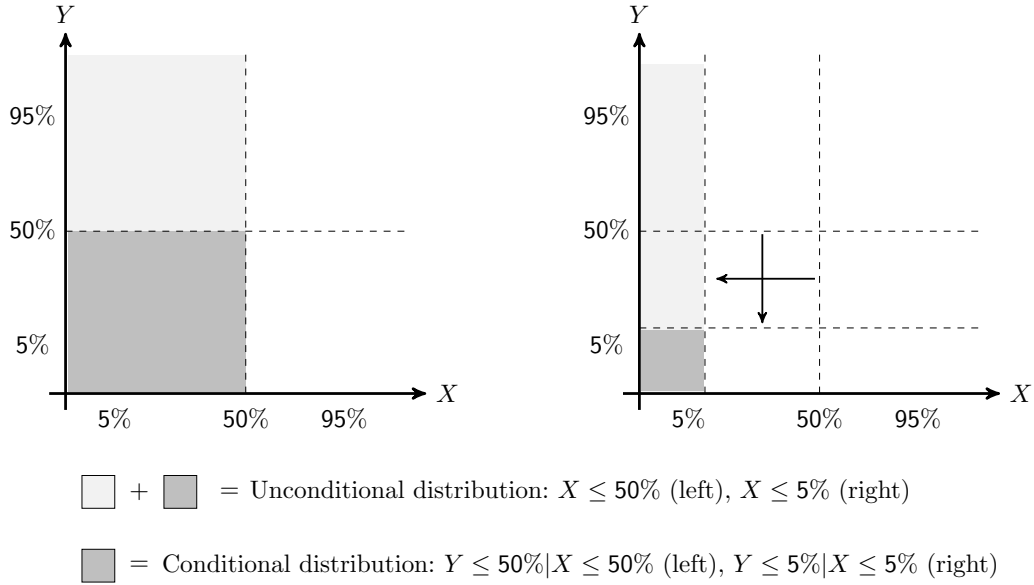


Figure 1. Schematic representation of eliciting the conditional 50% and 5% percentile.

the form of $P(Y \leq u\% \text{ percentile} | X \leq u\% \text{ percentile})$: For example, we start with the conditional median in the form of $P(Y \leq 50\% \text{ percentile} | X \leq 50\% \text{ percentile})$ for the variables of interest X and Y . This can be framed as: “Given that X is below your *median* for it, what is the probability that Y is also below your *median*?” (see Figure 1 on the left).

- (3) We elicit another percentile, one that corresponds to the (lower) distribution tail (e.g. the 5%), i.e. $P(Y \leq 5\% \text{ percentile} | X \leq 5\% \text{ percentile})$ which can be framed as: “Given that X is below your 5% percentile for it, what is the probability that Y is also below your 5% percentile?” (see Figure 1 on the right).

The $u\%$ percentiles correspond to the chosen severity levels of the risk characteristics that are of interest in an elicitation and hence risk assessment.

- (4) With the assessments of (2) and (3) in place, we can compare an expert’s judgements with different parametric copula forms. This is done by plotting the assessments against the converging conditional exceedance probabilities for selected parametric copulas simulated at the $u\%$ percentile. We will explain this in more detail and show this with actual assessment results in the illustrative case-study later.
- (5) With a first idea of which copula represents the expert’s information reasonably well given a specific rank correlation, we can test the robustness of that choice, e.g. by “feeding back” the probabilities for non-elicited percentiles and check an expert’s agreement for it.

Alternatively, in (3) we might elicit the conditional median for various percentiles. Thus, we elicit $P(Y \leq 50\% \text{ percentile} | X \leq u\% \text{ percentile})$, solely varying the $u\%$ percentile for X . Both ways of eliciting dependence information allow for choosing a copula that represents an expert’s input satisfactorily and both alternatives are in line with our scenario mapping method.

3. Definition of *conditional scenarios* and derivation of desiderata for a mapping method

In this section, we propose a definition for conditional scenarios which underpins our method and through which experts express the underlying rationale for their quantitative dependence assessment. This clarifies which information we obtain from experts through our scenario mapping method. Further, this definition allows for deriving the main desiderata for our method which have been used to guide its development by identifying the suitability of existing methods in the literature with respect to each of the desirable properties and combining these.

The term “scenario” is used differently within operational research, decision and risk analysis and related fields. Apart from our purpose of structuring experts’ knowledge as preparation for a quantitative uncertainty assessment and a better understanding of risk, other purposes, for which scenarios are formulated, include forecasting (Bunn and Salo, 1993), strategic planning (Van der Heijden, 2011), multi-criteria decision making (Stewart et al., 2013), various other applications in decision analysis, e.g. decision tree modelling (Chelst and Bodily, 2000; French and Rios Insua, 2000; Raiffa, 2002), and also defining risk itself (Kaplan, 1997; Kaplan and Garrick, 1981).

As one of the pioneers, Kahn and Wiener (1967) are credited in establishing scenario thinking as a structured approach to understanding uncertainties and risks better. They regard a scenario as “a hypothetical sequence of events constructed for the purpose of focussing attention on causal processes and decision points”. This is similar to Eden and Ackermann (1999) who describe scenarios as “hypothetical sequences of events constructed as causal chains of argumentation for the purpose of focussing attention on alternative futures”.

3.1. Definition of *conditional scenarios*

For our purpose of defining *conditional scenarios*, which are later used in a structuring method, first recall the commonly accepted definition of risk as the triplet $R = \langle s_i, p_i, x_i \rangle$, in which s_i denotes the i^{th} scenario, x_i its consequence and p_i the probability (Kaplan and Garrick, 1981). We base our definition of conditional scenarios on this by further refining the scenario part.

However, before doing this, note that the quantitative assessments presented in this paper and introduced in Section 2 are on multivariate consequences (x, y) and our core interest is in assessing the dependencies between these variables. Instead of considering complete scenarios, as in the definition of risk given above, we will need to consider the concept of a conditional scenario which is, formally speaking, the set of those scenarios that satisfy specified conditions. We are particularly interested in those conditional scenarios which are defined by having specific (given) values for one or both of our consequence measures. We wish to develop assessments of probabilities (and provide rationales) by using these conditional scenarios. For example, if we look at the conditional scenarios where there is no specification of variable y but the variable $x \leq x'$ for some particular value x' , then the overall probability assessment for these conditional scenarios would be equal to the probability that variable $x \leq x'$, that is, (when assessing over the full range of values for x') we are assessing the marginal distribution for the uncertainty of x . Of course, the assessments of probability have to be coherent (that is, they are consistent with the laws of probability), and we achieve this using the approach developed in (Werner et al., 2018b), which guides us

to make a sequence of assessments (relating in particular to tail events) which will be coherent.

The novelty here is to use conditional scenarios as an approach which enables experts to provide rationales for their judgements and to share those rationales with other experts. The conditional scenarios provide a way to structure experts' mental models, and in this way they can generate for themselves scenarios that lead, for example, to extreme events for one consequence measure and then explore through such conditional scenarios how likely extreme outcomes for the other consequence measure are.

It is important to note that we are using scenarios in a constructive, inductive, way to help the experts understand and provide rationales for how extreme events (as measured by the consequence variables) might occur. Hence our terminology is chosen to help with this process rather than to formally define a full mathematically complete structure. Indeed, the scenarios do not need to be developed further than is needed to gain sufficient understanding for the probability assessments of the events in question.

We introduce some notation as follows. The consequences are assumed to be future consequences within a system which is evolving through time. Hence we work with experts to identify current and future states of the system or critical events, which result in consequence values near those of interest. That is, these are plausible scenarios for the values under discussion. It is convenient to introduce notation for the *current state of the system* c_0 together with key *future states* c_1, \dots, c_k that follow in sequence at key points in the future 1 to k .

It is also convenient for experts to consider possible early *trigger conditions* $(c_{0,1}, \dots, c_{0,l})$ that could drive the system in different ways in the future. The 1 to k future states are described in terms of, possibly multiple, *enabling conditions*, vectors of the system's vulnerabilities and resilience for each future state, denoted as $(c_{1,1}, \dots, c_{1,m})$ to $(c_{k,1}, \dots, c_{k,o})$ for m enabling conditions in the 1st future state and o in the k^{th} .

Figure 2 shows the components of several scenarios schematically. Trigger conditions, enabling conditions and the consequence are represented by dark blue rectangles, light blue circles and a red rectangle accordingly. The bold arrows show the path of a single scenario whereas every path leading from a trigger condition to the consequence represents another scenario. The reader is also referred to section 6 for an example developed by an expert in a case.

Our scenario definition does not entail any past events, but in practical applications the experts may find it useful to consider them, either because they help to understand future developments or because they may not be fully known, and indeed the experts may specify trigger conditions which are not fully observable.

We address how facilitators can support experts to include trigger conditions even if they are only partially observable in Section 5.

As enabling conditions are the states of a system's variables that are affected by the presence of a trigger condition, they represent the system's vulnerabilities and resilience as stated above. In line with Haimes (2006), *vulnerabilities* are "the inherent states of the system (e.g., physical, technical, organizational and cultural) that can be subjected to a natural hazard or be exploited to adversely affect (cause harm or damage to) that system". *Resilience* refers to the "ability of the system states to withstand a major disruption within acceptable degradation parameters and to recover within an acceptable cost and time" (Haimes, 2006). Regarding the time

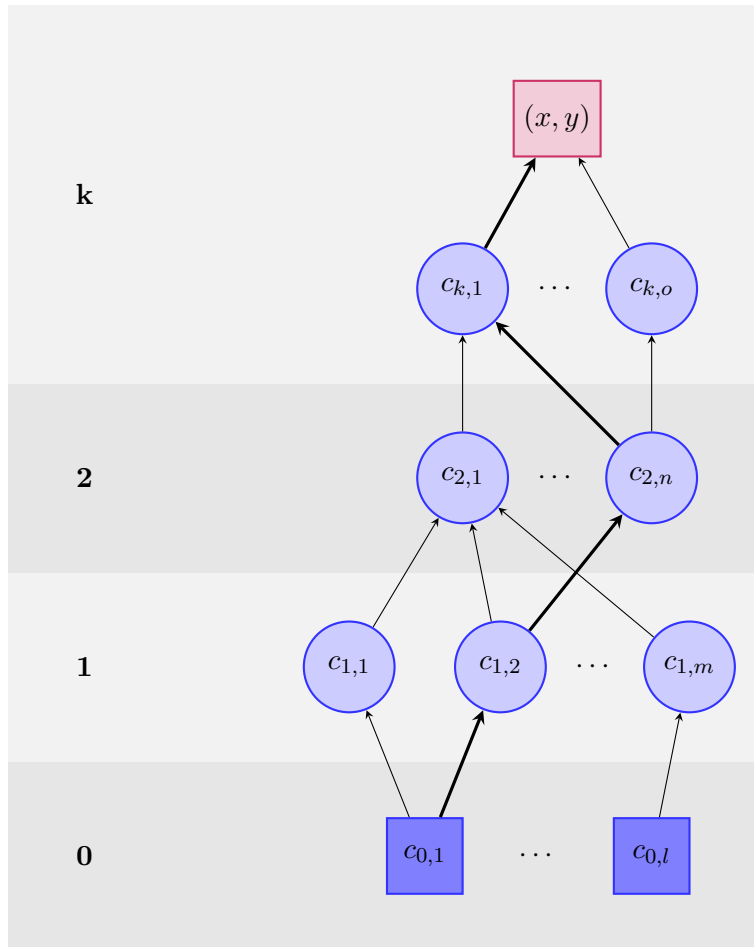


Figure 2. Basic components of our scenario definition.

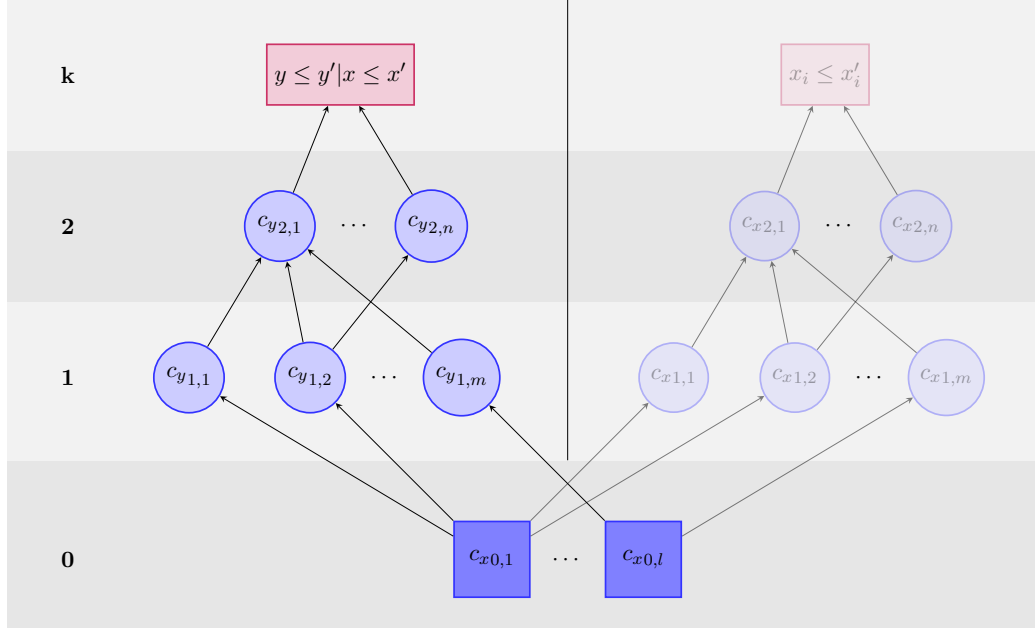


Figure 3. Representation of conditional scenario for consequence $(y|x)$.

frame, the last enabling condition at time k is closest to the consequence (x, y) . An important consideration for enabling conditions is the *causal field*. This is discussed in more detail in Werner et al. (2018a). It is the presumed background of necessary conditions that need to be in place together with enabling conditions to allow for a trigger condition to lead to the scenario's consequence while these are not regarded as causing the consequence themselves. When identifying enabling conditions it is important to distinguish the presumed background from these as shown later. Methodologically and philosophically, the terminology we use is related to a holistic, system-based understanding of risk discussed in Haimes (2009). Further, risk perception research (Renn, 1992) identifies similar event types for which initiating *hazard events* result in *secondary effects* that affect more people than the ones affected by the original hazard event. As example Renn (1992) mentions economic impacts together with social and political pressures. The adopted naming convention of our definition stems partly from linguistic expressions of causal relationships, such as *cause*, *enable*, *prevent* (Sloman et al., 2009).

Finally we return to the notion of a conditional scenario, which as we noted above is the set of all scenarios which lead to the specified consequences. We are particularly interested in assessing conditional tail probabilities of the form $y \leq y'$ for some fixed y' out of the set of scenarios for which $x \leq x'$ for some fixed x' .

Figure 3 illustrates the way we approach scenarios for conditional events.

Our structured approach enables experts to create a rationale for their conditional probability assessment, illustrated in the scenario mapping method later. It first considers the scenarios which relate to $x \leq x'$. Once their trigger conditions have been identified through a backwards logic structuring, these same trigger conditions are used to start the conditional scenarios (relating to y). The enabling conditions can

be changed to understand how well the trigger conditions lead to y being below (or above) some fixed y' . By exploring both sets of scenarios, experts are then well placed to assess a conditional probability.

Following the introduction of our scenario components that will be used to describe dependence within a system in our mapping method (section 5), it is worth noting that in scenario analysis processes more generally, identifying the factors that are relevant for explaining a system’s behaviour together with their interactions and their impact on its future evolution are the main steps in the scenario development phase when deriving a system’s model and in scenario generation (Tosoni et al., 2018).

A main challenge in scenario generation relates to the *completeness* and *comprehensiveness* of the selected scenarios sets (Tosoni et al., 2019, 2018). The first refers to including any factors that might impact a system whereas the latter is the more pragmatic idea of only including factors that significantly impact a system. Several types of uncertainties, aleatory and epistemic ones, influence one’s approach to addressing comprehensiveness. For us, in particular the uncertainties on a system’s structure, scope and evolution are of interest.

With regards to the parameterisation of these uncertainties, in some literature (Tosoni et al., 2019, 2018) a distinction is made between *pluralistic* and *probabilistic* scenario generation. The pluralistic approach asks for few scenarios, usually from experts, which represent their different assumptions on scenarios’ future behaviour. In contrast, the probabilistic approach scenario are treated as subsets from large random samples. Recall that we use scenarios qualitatively (prior to a quantitative assessment) and are therefore more in line with the first, the pluralistic approach and regard them as representative futures.

For addressing the comprehensiveness of the future behaviour of our scenarios, we apply the more general definition proposed by Tosoni et al. (2018), which in a nuclear safety context advises to focus on only those future scenarios in which corresponding performance targets are violated. Similarly, however in the context of a probability rather than a risk assessment, we aim to cover all these scenarios which affect the future consequence $x \leq x'$ being below a certain severity level as required in the assessment. This corresponds to our desideratum (section 3.2) of *relevance* and in particular in tail dependent systems avoids the inclusion of scenarios that are not regarded in a later assessment.

For addressing the comprehensiveness of the identified trigger and enabling conditions, we base our mapping method on debiasing techniques and the option to share scenarios without seeking consensus as shown in section 5 and the case-study discussion in 6.3. As such, experts derive partial scenarios that are detailed enough to give them comfort in assessing a joint tail probability. The sharing element of our process allows then for a judgement for whether one’s scenario is detailed enough.

In this paper, we focus on bivariate dependence as this is already cognitively complex for experts to assess. Therefore, our variables of interest are denoted as X and Y . Nevertheless, similarly to the bivariate case for which we structure conditional scenarios in accordance with the elicitation of $P(Y \leq u\% \text{ percentile} | X \leq u\% \text{ percentile})$, our method can be extended to structure conditional scenarios of larger conditioning sets, such as $P(Y \leq u\% \text{ percentile} | X_1 \leq u\% \text{ percentile}, X_2 \leq u\% \text{ percentile}, \dots, X_n \leq u\% \text{ percentile})$.

3.2. *Desiderata for a conditional scenario mapping method*

From our scenario definition above, we identify three main desiderata which a method structuring conditional scenarios for (tail) dependence assessment should possess.

First, a method should support experts to only map scenarios that are of *relevance*. That is, experts should only map scenarios which result in the consequences $x \leq x'$ and $y \leq y'$ that are specified, i.e. being above/below the $u\%$ percentile for x' and y' . A method, which elicits scenarios of which only few or even none are relevant or which explore other scenarios, is not suitable.

As second desideratum, a method should identify the *systemic impact* of threats. For instance, in risk analysis, consequences (or failures) can be attributed only very rarely to a single cause. Rather chains of events combine and interact to produce the outcome. This is reflected in our scenario definition in the possibility to have more than one trigger condition and enabling condition. A method should therefore support experts in identifying and capturing systemic interactions of various trigger conditions and enabling conditions in an intuitive and transparent way.

Third, a method should allow for capturing *inter-systemic dependence*, i.e. it should support experts in understanding the difference between unconditional and conditional scenarios. For that, our method needs to be able to incorporate the idea of a (sub-)system boundary together with a link between both (sub-)systems.

4. Applicability and features of existing methods in the related literature

With respect to the desiderata, we reviewed the literature on potentially applicable techniques and methods, mainly from the areas of Problem Structuring (“soft OR”), PRA and Scenario Planning. This helped to identify features from existing methods that we can use for our purposes. Thus, we examined how existing methods perform in terms of three different properties, (1) understanding impact and severities in an anticipatory way, (2) modelling the dynamic complexity of an underlying system and (3) capturing how common causes propagate through different (sub-)systems by defining their boundaries and links. The first property relates to the desideratum of identifying only relevant scenarios. The other two properties relate to the second and third desiderata of understanding and capturing the (complex) intra- and inter-systemic impact of trigger conditions. The second property focusses on the interaction of the scenario components of trigger conditions and enabling conditions within an individual scenario whereas the third property’s focus is on determining the impact of trigger conditions across dependent scenarios. Table 1 provides an overview on several methods’ applicability with respect to the three properties which is further detailed below.

Anticipatory severities. Applicable methods for the first property identify events that lead to a specific level of severity. This is done through *backwards logic*. It differs from forward logic approaches by not considering the possible development of scenarios from a fixed starting point (such as a threat in the current state of the world) but instead determining the threats that need to happen in a certain environment (or under certain conditions) for given outcomes to happen (at a specific point in future). In the SP literature, the terms *forecasting* and *backcasting* have been introduced for that, whereas Ducot and Lubben (1980) and Bishop et al. (2007) distinguish between *exploratory* and *anticipatory* scenarios. Wright and Goodwin (2009) propose backwards logic to enhance traditional SP methods, such as Intuitive Logics (Van der

Table 1. Features of existing structuring methods.

Technique	Reference	Technique models / captures...			
		anticipatory severities	dynamic complexity	inter-system propagation	
PRA ^a	Fault Tree Analysis	Vesely et al. (1981)	✓	×	×
	Event Tree Analysis	Ericson (2015)	×	×	×
	FMEA	Stamatis (2003)	✓	×	×
	HAZOP	Crawley and Tyler (2015)	×	×	×
	Root Cause Analysis	Wilson (1993)	✓	×	×
	Influence Diagrams	Howard and Matheson (2005)	×	✓	×
PSM ^b	Causal Loop Diagrams	Morecroft (2010)	×	✓	×
	Soft System Methodology	Checkland and Poulter (2010)	×	✓	×
	Causal/Cognitive Mapping	Ackermann and Eden (2006)	×	✓	×
	Knowledge Maps	Howard (1989)	×	✓	×
	Critical System Heuristics	(Ulrich, 2003)	×	×	✓
SP ^c	Intuitive Logics	Van der Heijden (2011)	×	✓	×
	Backward-Logic SP	Wright and Goodwin (2009)	✓	×	×
	Horizon Mission Methodology	Anderson (1993)	✓	×	×
	Impact of Future Technologies	Bishop et al. (2007)	✓	×	×

^a Probabilistic Risk Analysis; ^b Problem Structuring; ^c Scenario Planning

Heijden, 2011), as a way to focus particularly on rare events. Their method is motivated by crisis management approaches which aim at preparing organizations for high-impact/low-probability catastrophes.

Another SP method applying backwards logic is the *Horizon Mission* method (HM) (Anderson, 1993). HM originated within NASA to support engineers’ decision making about future research and development (R&D) pathways as before their scenarios often led to recommending incremental rather than breakthrough research. In the HM method, engineers first envision a horizon mission (infeasible given today’s technology) and then identify the new capabilities needed to achieve it. The *Impact of Future Technologies* method serves a similar R&D purpose at IBM (Bishop et al., 2007).

Similar to these SP methods, in PRA in particular *Fault Tree* and *Root Cause* analysis methods (Bedford and Cooke, 2001) investigate and assess specific future risks and severities. In PRA however, scenarios are captured by event sequences rather than narratives which are common for SP approaches.

Dynamic complexity. Regarding the second property, we observe from Table 1 that most PRA and SP methods do not allow for modelling dynamic complexities within a system. For instance, Paté-Cornell (2012) emphasises the need for traditional PRA methods, which often model engineered systems, to apply more holistic forms of analysis in order to address the challenge of more complex risks and their developments. Similarly, White (1995) examines the deficiencies of PRA methods with regards to systems thinking while Ackermann et al. (2014) and Ackermann et al. (2007) highlight the need for a holistic and systemic approach to risk analysis to account for “risk systemicity”. It refers to the idea that “the effect of two risks might be more than the sum of the two individual effects” (Williams et al., 1997).

While there is no agreement on the definition of a dynamic and complex system (see Ladyman et al. (2013) for a discussion), a commonly mentioned characteristic is *non-linearity* due to reinforcing (or vicious) feedback loops. Their identification is hence crucial when analysing a complex system. Various methods, summarised under the umbrella of systems thinking in PSMs/“soft OR” do this to try understanding a system

holistically rather than through its separate parts. Common methods that identify feedback loops by graphical representation of influences are *Causal Loop Diagrams* (Morecroft, 2010) and *Cognitive Maps* (Ackermann and Eden, 2006; Poplawska et al., 2017). Both methods allow for a participatory approach to modelling complex problems (Cunha and Morais, 2016) and have been used in mixed-method approaches for analysing structures of systems (Lowe et al., 2016).

Inter-system propagation. Related to understanding dynamic and complex systems is the assessment of common cause propagation through distinct (sub-)systems. This is the third property and it is not clear how most methods distinguish between (what is perceived/defined as) different (sub-)systems (see Table 1). It draws on a fundamental aspect of systems thinking in PSM, the idea of a *system boundary* (Churchman, 1970). For PSMs, identifying what lies inside a system and hence which factors are included in a model requires experience and judgement. This is why a modelling process is usually iterative and circular rather than linear (Morecroft, 2010). Once a first model version is constructed, experts might refine the model by re-assessing which factors to include (or exclude) based on a reflective understanding. Emphasising the importance of the circular modelling process and the judgemental nature of a model boundary is in agreement with Midgley (1992) who discusses the definition of a system’s boundary from the viewpoint of *Critical System Heuristics* (Ulrich, 2003). This is a framework for participatory and reflective practice on boundary judgements which requires to consider critically what a system includes and to examine it from multiple perspectives through a checklist/question-based approach. More generally, this is related to the issue of infinite regress when modelling (French, 1995), for which Phillips (1984) introduced the term of a *requisite model* that results from a circular and interactive modelling process.

In order to better understand the link between distinct (sub-)systems, i.e. capture how trigger conditions propagate through them, a *bow-tie* approach is applicable. In PRA, models with a bow-tie structure are used to assess how a hazard is caused by threats and at the same time is the cause of consequences (Ale et al., 2006). While this can be applied within a single system, it can be extended for hazards across different (sub-)systems.

5. Our method for mapping conditional scenarios

After having examined the applicability of existing methods for our desiderata and by that identified the main features of existing methods that comply with our desiderata, we now introduce our method which synthesizes specific elements of some of the methods discussed in Section 4 for our purpose of mapping conditional scenarios for (tail) dependence elicitation.

5.1. Overview of the conditional scenario mapping process

For mapping conditional scenarios, we propose an iterative process which is facilitated with each expert individually. The resulting scenarios are shared only in the final step of knowledge sharing. Figure 4 provides an overview of the overall mapping process and its individual parts.

In the *first step*, the facilitator ensures the expert’s familiarity with the task and set-

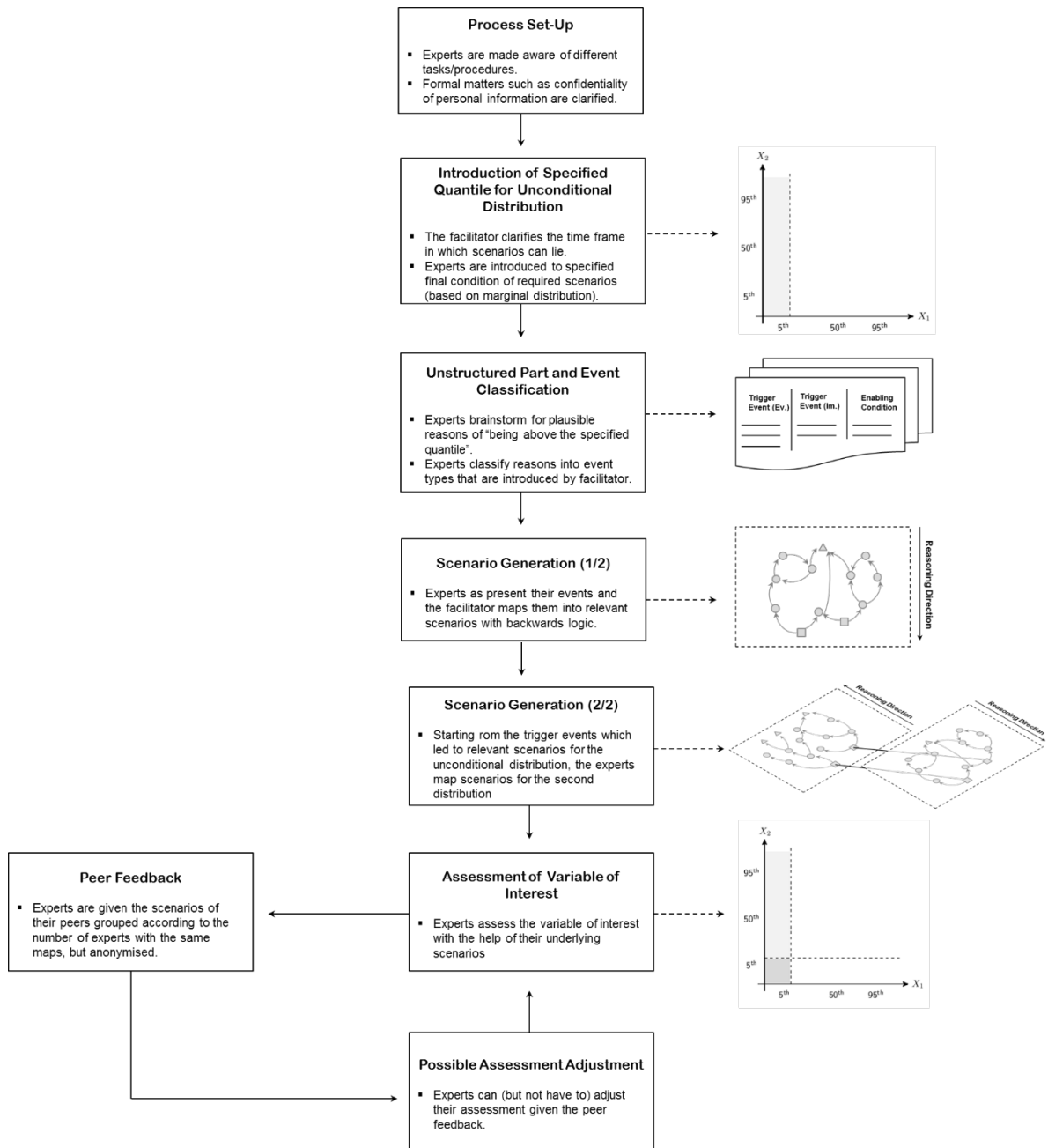


Figure 4. Overview of Scenario Mapping Process.

up. The facilitator introduces and explains the different steps of the mapping process together with the tasks needed at each of them. Further, common expert judgement formalities, such as confidentiality of personal information and the usage of the assessment and scenario mapping results, are clarified.

In the *second step*, the expert is introduced to the first variable of interest, X , that of the unconditional distribution. This usually involves clarifying the marginal distribution either specified by data or a previous elicitation. In both cases it is important the expert is in agreement with X 's distribution. Then, we (as facilitator) introduce the expert to the consequence of the unconditional scenario, x_i , which states that the unconditional distribution is above or below a certain percentile, and ensure that the expert is clear with its interpretation. Further, we clarify and emphasise the specified time frame in which scenarios can lie. Note, we regard our method more suitable for shorter rather than longer time frames, especially whenever the focus is on tail dependencies which might be not recognisable for events too far into the future.

In *step three*, the expert is given time to brainstorm and note down the different reasons of why the variable of interest can lie above or below the specified percentile. We use the word "reason" here as this part of the process is unstructured and an expert might express these reasons in any form. For example, they might mention specific key words or give a narrative. In both cases, the facilitator should take notes if the expert does not do so herself. Note, this step employs *backward logic* by reasoning from a specific consequence to potential causes.

Fourth, the facilitator classifies (in agreement with the expert) the reasons, identified in the previous step, into being either enabling conditions or trigger conditions. Verbal conventions used for the classification and distinguishing the two event types are:

Trigger conditions. Remember, a trigger condition is a plausible initiator of a scenario contained in the current state of the world and it may or may not be (fully) observable. For clarification we might add words like "start", "outbreak", "attack", "eruption", "shock" and so on. For example, "disease outbreak", "terrorist attack", "volcanic eruption" or "oil price shock".

For observable trigger conditions, it is possible to neglect any preceding events as we condition on them knowingly. However, for trigger conditions that are only partly observable, we need to include immediate preceding events (which led to the trigger condition) for ensuring a richer set of scenarios. For instance, suppose that an expert identified "oil price shock" as a trigger condition. In this case, the facilitator might have to clarify whether this is due to geo-political risks involving OPEC countries or due to a change in usage of alternative energy sources, as for both versions, very different scenarios can unfold in the future.

Another remark on correctly identifying a trigger condition is that experts cannot position it into the future. If this occurs, it is important that the facilitator supports the expert in re-considering why such an event will happen in future (as a response to another event) in order to identify its corresponding trigger condition in the current state of the world.

Lastly, note that some trigger conditions (such as the previous examples) are happening in a short time. Other trigger conditions are characterised by a longer development of an event. For these, it might be helpful to insert words like "development", e.g. "development of (long lasting) rain showers" as a trigger condition for a certain flood severity rather than "rain start".

Enabling conditions. Complementary to trigger conditions, an expert should also identify enabling conditions. Remember, these follow from the trigger conditions and capture evolving trends/changes in a system based on the conditions and responses that need to be in place for a trigger condition to reach the scenario’s consequence. For that they might be labelled as “higher”/“lower” for classification and clarification, e.g. “lower economic growth”, “higher risk of infection” or “higher migration”. The categorisation into trigger conditions and enabling conditions is essential for mapping conditional scenarios from the current unconditional ones as shown in the following steps.

In the *fifth step* of the process, experts determine the links between trigger conditions and enabling conditions in order to map out the path to the final enabling consequence entailing the specified consequence. This part makes use of cognitive mapping and causal loop diagram methods as we can identify feedback loops and the overall interconnectivity of the events. Links are set according to an expert’s belief while we omit the assignment of polarities as these are already captured in the labelling of enabling conditions. Further, this step serves as a robustness test for the enabling conditions as experts might change their labelling based on the links. Testing robustness in graphical models is commonly embedded in the modelling process (Lane, 2000; Montibeller and Belton, 2006). This part can be supported by mapping software to allow for a direct visualisation of the unfolding scenarios. Once an expert is satisfied with the resulting set of unconditional scenarios, the final picture can be captured.

Next (in *step six*), the expert maps how (and if at all) the trigger conditions that are relevant for the unconditional distribution propagate in the conditional one. For that, we simply import the previous trigger conditions into the conditional scenario space of $Y|X$. Then, an expert identifies the necessary enabling conditions and links (in *forward logic*) from the imported trigger conditions with respect to the specified aspect and time frame for the conditional scenario.

Now, an expert has a thorough understanding of how both variables of interest are affected by the same trigger conditions and can proceed with the quantitative assessment in the form introduced earlier.

Finally, the *last step* of the process allows for (anonymously) sharing experts’ scenarios and letting them change the quantitative assessments based on the other experts’ scenarios if desired.

By first facilitating the scenario mapping sessions with experts individually, before providing each expert with the scenarios of other experts, this process shows similarities with elicitation processes, such as Delphi (Rowe and Wright, 2001). A difference is however that we do not seek consensus. Rather, our process builds on findings that the accuracy of individual assessments improves upon receiving the reasoning and rationales other expert’s judgements (Wintle et al., 2013; Yaniv and Choshen-Hillel, 2012). An expert judgement process for univariate quantities which encourages a second round of assessments after individual assessments have been shared and discussed with other experts is the IDEA protocol (Hanea et al., 2017, 2018). We similarly aim for assessment improvements by sharing each expert’s scenario maps among all experts and then offering to adjust assessments.

The overall process of mapping scenarios is repeated for all percentiles of interest.

6. Illustrative case-study: Assessing dependence for risk management in the UK Higher Education sector

The higher education (HE) sector in the United Kingdom (UK) has been frequently of interest for applying operational research techniques, mainly for problems of performance measurement and resource allocation (DeWitte and López-Torres, 2017; Johnes, 2015; Mayston, 2003). In the context of the latter, it is crucial to analyse risk in this sector. However, little experience is available for that. This is the case, even though the general management of HE in the UK has been extensively studied and is well-understood (HEFCE, 2001, 2005).

As factors that might pose a risk to an HE institution, HEFCE (2005) outlines variable tuition fees, increasing competition and changing students' expectations, an increased exposure to and reliance on overseas markets, large investments in infrastructures to facilitate institutional expansion as well as potential loss of market share due to new technologies. Further, Altbach and Knight (2007) view tuition fee income as a main driver for internationalisation whereas some uncertainties affecting its development are political realities together with national security concerns, such as changing visa requirements in the face of international terrorism, government policies influencing the cost of studies, the potential expansion of domestic capacity for sending countries to meet education demands, the increasing importance of English as lingua franca, the alignment and accreditation of degrees and the future impact of e-learning offerings. In order to better understand such complex uncertainties, our scenario mapping method has been applied in a dependence assessment as part of a risk analysis that supports decision-makers in charge of managing the postgraduate taught course portfolio at the authors' home institution. The dependencies assessed and mapped are between the tuition fee income of the established *MSc Business Analysis and Consulting* (BAC) course and the newly introduced *MSc Data Science* course within the next four years.

6.1. Background: elicitation of marginal distributions

Before presenting the dependence elicitation together with the application of our scenario mapping method, we briefly outline how the marginal distributions for that have been elicited. This was necessary as these are predictions, hence no other data exist. The full elicitation protocol with both elicitations and each expert's scenarios can be found in Werner (2017).

In the elicitation, in total five experts, who are in charge of managing postgraduate taught courses and implementing the new course, MSc Data Science, participated. The variables of interest are defined as the generated income from each course through tuition fees in 2020/21. For the marginal distributions, we used a common method for eliciting univariate distributions, the Classical method (Cooke, 1991; Quigley et al., 2018). For that, experts quantify their uncertainty for the variables of interest and calibration variables (also known as seed questions). For the latter, the true values are known to the facilitator but unknown to the experts. They allow for combining the experts' assessments after being validated against empirical data. Experts are treated as statistical hypotheses and combined as to maximise the statistical accuracy and informativeness of the combined assessment (for details see the above references). In total, we elicited ten calibration questions on the financial performance of the (existing) MSc BAC, the student population in this courses as well as the university student popu-

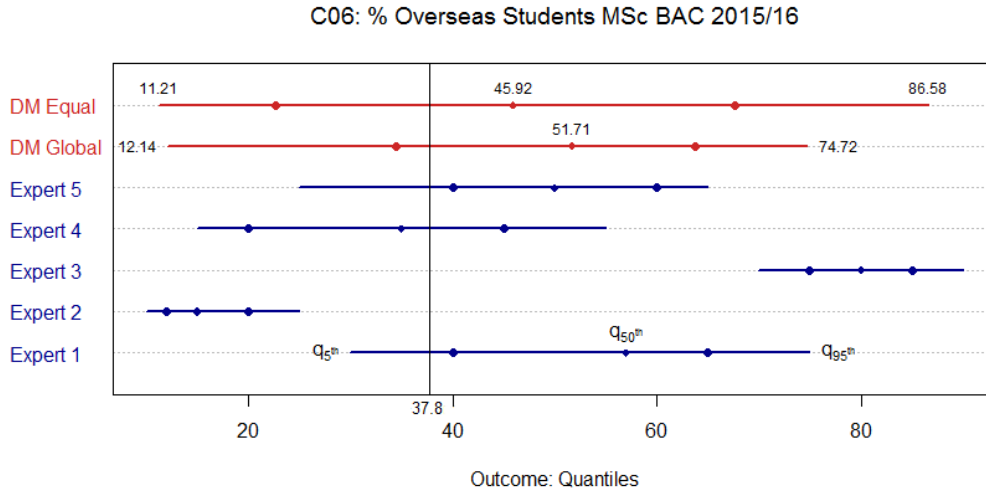


Figure 5. Exemplary calibration question.

lation and the expected future employment data for data scientists. Figure 5 shows exemplary the assessment for one of the calibration questions:

“What was the percentage of oversea students for the MSc BAC in 2015/16?”

The experts’ assessments (blue) and weighted combinations (red) are shown together with the true value which is given by the vertical line. The “DM Global” is the combined assessment according to the Classical method and the “DM Equal” is the equal-weighting combination, in which every expert receives the same weighting.

From the calibration questions, a main outcome is that the combinations with the Classical method and the equal-weighting combinations had similar mean values while the former has been consistently more informative, i.e. the 5% and 95% percentiles were closer to the mean and the overall distribution less spread out. This is also shown in the combination of the target questions, variables of interest (target questions):

- (1) *“For the academic year 2020/2021, i.e. in 4 years from now, what will be the generated income from the MSc Business Analysis and Consulting?”*
- (2) *“For the academic year 2020/2021, i.e. in 4 years from now, what is the generated income form the MSc Data Science (reminder: this course start from next year onwards)?”*

The corresponding resulting distributions are shown in Figure 6.

6.2. Dependence elicitation and conditional scenario mapping

After having assessed the marginal distributions assessed, we then mapped the conditional scenarios for the variables of interest with each expert before eliciting their dependence through the following target questions²:

²The experts were given as well the corresponding monetary values for the specific percentiles by the facilitator.

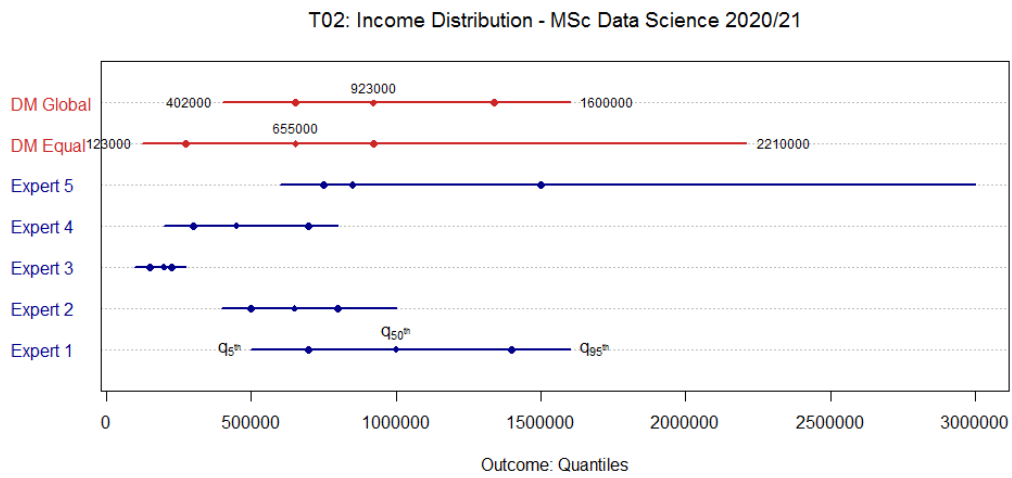
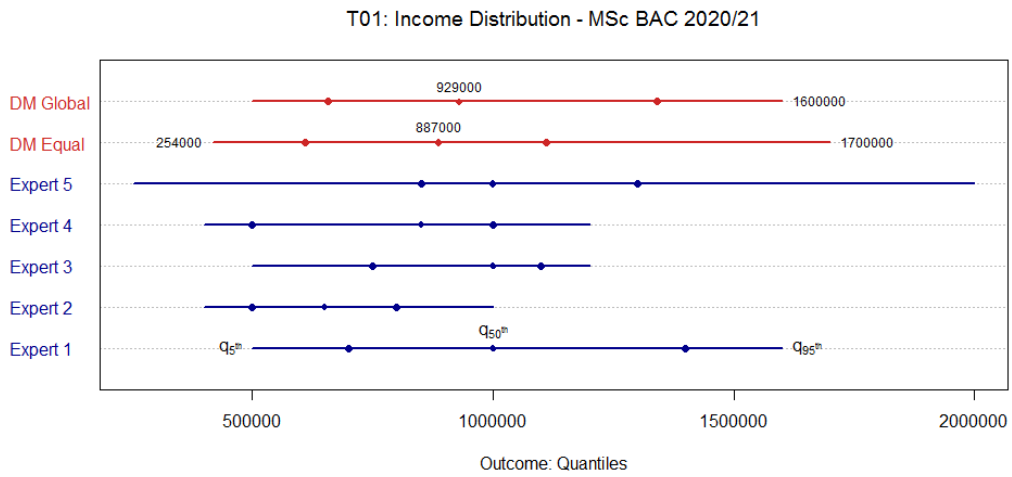


Figure 6. Assessment results for variables of interest of marginal distributions.

- (1) “Given that the generated income of the MSc BAC is below its median in the academic year 2020/21, what is the probability that the MSc Data Science is also below its median?”
- (2) “Given that the generated income of the MSc BAC is below its 5% percentile in the academic year 2020/21, what is the probability that the MSc Data Science is also below its 5% percentile?”

We facilitated the mapping of the unconditional and conditional scenarios for being below the median first (with the process of Section 5), then repeated the same process for being lower than the 5% percentile. Figures 7 and 8 show the unconditional and conditional scenarios for the former (50% percentile) of one of the experts that were elaborated within approximately 20 minutes.

As we can see, for the 50% percentile the expert believes that most trigger conditions of the unconditional scenarios will affect the conditional distribution similarly. This is due to most trigger conditions being on changes affecting HE in UK (and even globally) as a whole. A slight difference is the future demand of MSc Data Science graduates (in particular) due to a more important data science market. As such, the expert considers MSc BAC students to be less attractive in an (then) established data science job market in which certain technical abilities (added from elicitation protocol: such as programming and more profound statistical, mathematical modelling and OR knowledge) are desired by employers. This is on the other hand not a trigger condition for the conditional scenarios as the MSc Data Science offers these skill. For the 5% percentile, this expert considers the trigger conditions, which are part of the unconditional scenarios, to have the same impact on the conditional distribution. In other words, once the income generated from tuition fees by the MSc BAC is below a certain threshold, the scenarios must be relevant on a more global level, so that the MSc Data Science will be affected similarly. Most experts’ scenarios and assessments consider such a potential tail dependence to be present in the course portfolio as they view the risk of being below the 5% percentile as a result of events that affect the UK HE sector more broadly rather than the different courses individually. Due to such similar beliefs, no expert changed her/his assessment after reviewing the other experts’ scenarios in the last round of the process (Figure 4).

As a result, when aggregating the experts’ assessments, both, the equal weighting combination as well as the performance-based one (based on the statistical accuracy of experts for the marginal distributions), indicate the fit of a tail dependent copula. Figure 9 shows how both combinations fit well with a Clayton and Survival Joe copula of rank correlation of 0.55). In contrast, tail independent copulas, such as the Gaussian one, do not fit well, even when altering the rank correlation values. As mentioned in Section 2 this is has been a pragmatic way of eliciting dependence and distinguish between some main parametric copulas.

6.3. Discussion of our scenario mapping method in the case-study

Before concluding the overall paper, we briefly reflect on the scenario mapping method’s application in the illustrative case-study. While this discussion is based on only one case-study and hence the below points are not necessarily generalisable, they at least indicate how the method has performed when used with experts in a real world risk analysis context.

In the introduction, we outlined that structuring experts’ judgement prior to a quan-

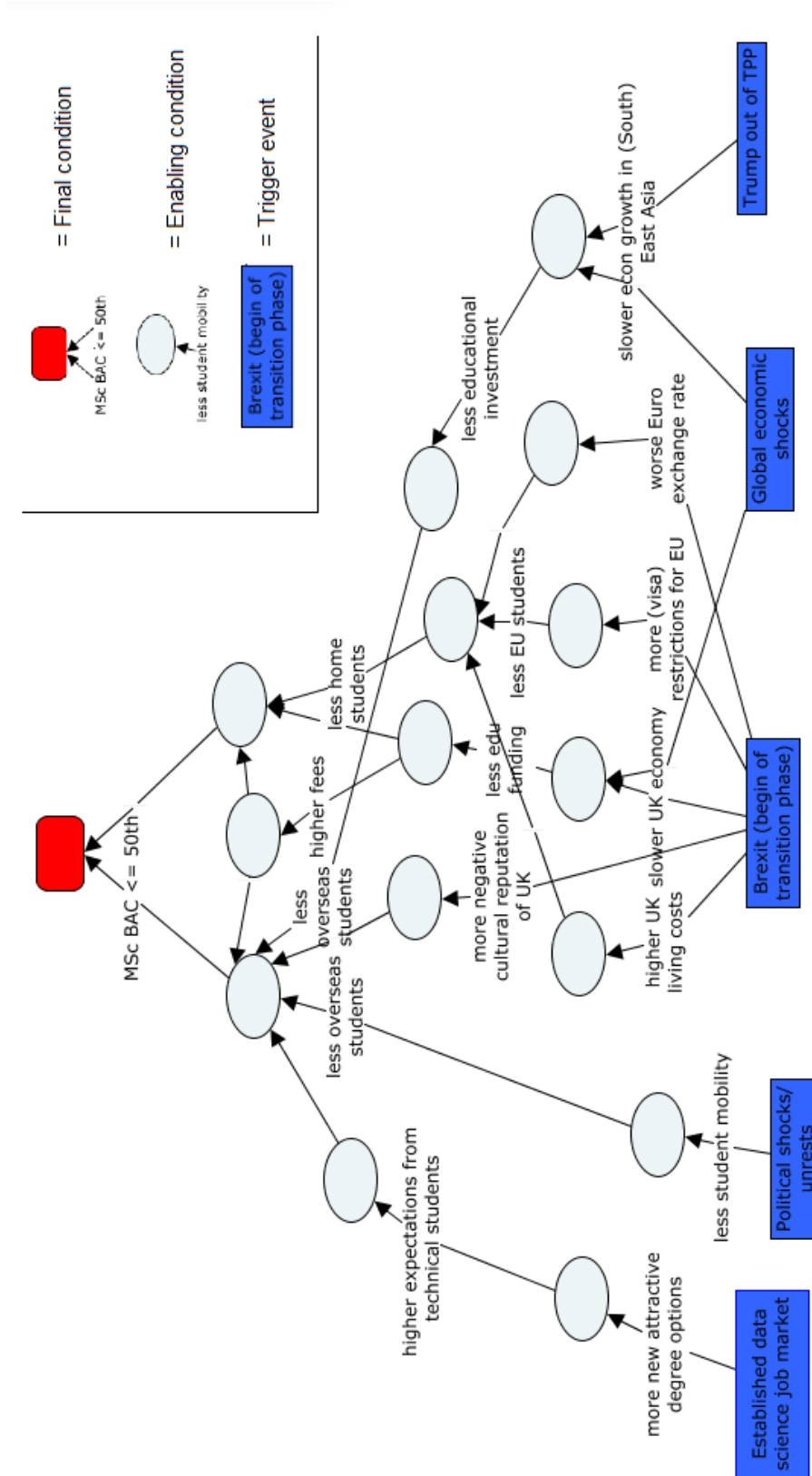


Figure 7. Unconditional scenarios for 50% percentile.

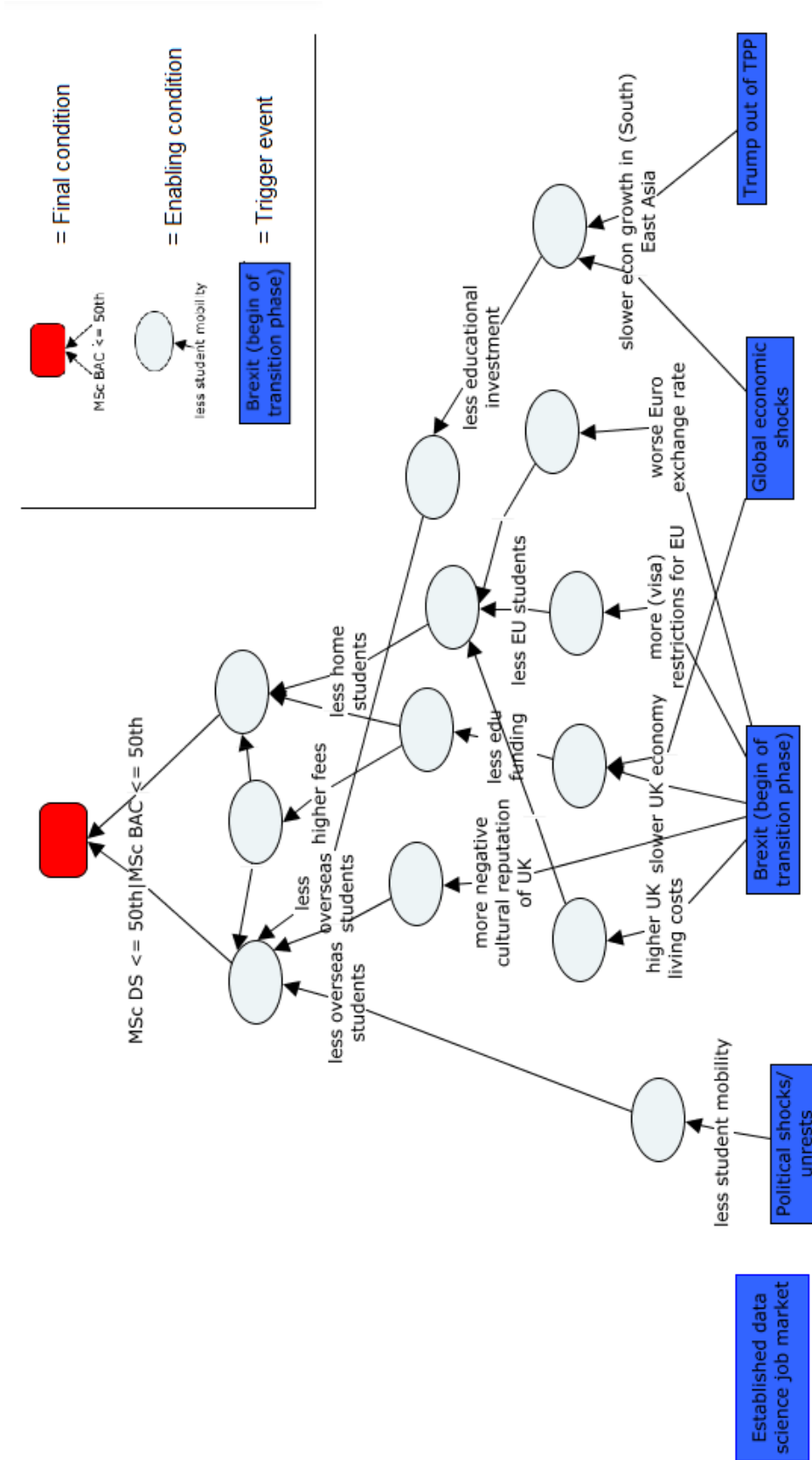


Figure 8. Conditional scenarios for 50% percentile.

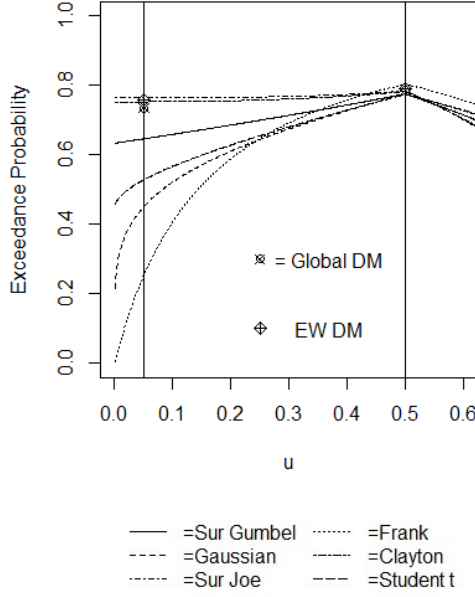


Figure 9. Fitting parametric copula forms to combined assessments (rank correlation=0.55).

titative assessment allows for (1) increasing confidence in the quantitative assessment, (2) mitigating potential cognitive fallacies and (3) supporting transparency together with reproducibility of the expert judgement study.

With respect to (1), increasing confidence for the later quantitative assessment, the experts' feedback shows that they regard our method as helpful for obtaining a better understanding of the dependence relationships and hence have more confidence to assess dependence quantitatively after using it. This is attributed to the possibility of expressing their thinking in natural language whereas the classification into the different event types is regarded as intuitive. Further, the decomposition of the dependence relationships, which allows for seeing how the trigger conditions that are elaborated in backwards logic (for the unconditional distribution) are relevant when thinking about the conditional distribution, is appreciated by the experts as a structured way to think about complex scenarios by regarding the influencing factors of dependence relationships separately.

With regards to (2), the mitigation of biases, the case-study results indicate that two main cognitive fallacies for dependence assessment can be mitigated through our method. For an overview on the prevalent heuristics and biases in dependence assessment, see Werner et al. (2018a). Further, it is indicated that our method accounts for findings from the area of descriptive decision making research on assessing probabilities of extreme (low-probability) events.

When eliciting dependence in conditional form, a common fallacy is the *confusion of the inverse* (Dawes, 1988; Hastie and Dawes, 2001). A conditional probability, such as $P(Y|X)$, is confused with $P(X|Y)$. One explanation states that this fallacy is related to the better known *representativeness heuristic* (Kahneman and Frederick, 2002) as people assess how similar or representative characteristics of X are for Y (rather than the conditional probability). Another explanation refers to an incorrect perception of a causal relationship between X and Y , for instance due to a reversed temporal order of X and Y (Bechlivanidis and Lagnado, 2013; Einhorn and Hogarth, 1986). Our

method decomposes both variables of interest and by doing so, it challenges the representativeness heuristic through alternative scenarios and clarifies the perceived causal and temporal orders. Several experts fed back that our method helped to consider alternative scenarios and that the order of the conditional variables is clarified.

In other cases, people *confuse conditional and joint probabilities* (Einhorn and Hogarth, 1986). Most explanations attribute this confusion to linguistic ambiguities introduced by describing conditional dependence relationships through “given that” or “if” (Einhorn and Hogarth, 1986). Based on anecdotal evidence, our method helped one of the experts to clarify the meaning of “given that” (i.e. that we elicit conditional probabilities in this case-study) when thinking out loud about the probability of a “perfect storm” and it being a joint probability relationship. As such, the expert first thought about the probability of a “perfect storm” event happening as being extremely small before realising that she is assessing a conditional probability which was then assessed higher after reflecting on her scenario map.

When eliciting tail dependencies (i.e. joint extreme events), there is indication that the mapping of unconditional scenarios with the backwards logic allows for mitigating some main potential cognitive fallacies. While the advantages of applying backwards logic are also discussed in Wright and Goodwin (2009), our experts confirmed that this way of reasoning helped them to consider the necessary scenarios for the extreme final, enabling conditions entailing the consequence.

In this context, a first fallacy is *frame blindness*, i.e. forecasting the wrong event. With backwards logic, experts elaborate scenarios by starting from the consequence, so that they will not explore scenarios that are not relevant.

Another challenge is that low probability events (in the tail of a distribution) by definition lack a reference class for similar events. Therefore, we cannot assess whether experts have well-calibrated assessments or are (for example) frequently overconfident. Nevertheless, by seeing which experts’ scenarios are coherent and plausible we can understand better with which assessment experts might struggle and further experts can explore how final enabling condition entailing the specified consequence can (or cannot) be reached from the current state of the world and if not discard them. In that way, the idea of coherent and plausible scenarios provided our experts with a guidance for structuring the importance of future events which are necessary for an extreme scenario’s consequence to occur.

For (3), ensuring transparency of the quantitative assessment results, we found the graphical representation of the experts’ rationales as crucial. Experts’ scenario maps can be associated with later assessments which makes the outcome of the elicitation more transparent for anyone not involved in it. As such, the quantitative assessments of this elicitation study can be easily compared with the corresponding scenario maps.

7. Conclusions

In this paper, we have proposed a scenario mapping method for structuring experts’ knowledge about dependence relationships prior to a quantitative assessment. This is important due to the high cognitive complexity that experts face when assessing joint distributions. The aim is to offer a flexible method that can be used together with various dependence models, including ones for tail dependencies. Structured approaches for supporting experts in expressing their rationale have already been recommended for eliciting univariate uncertainty (Cooke and Goossens, 2004, 1999). For multivariate uncertainty however, no methods were available before the one proposed in this paper

according to the authors' knowledge. As shown in Section 4, various methods proposed for similar purposes, such as scenario planning and qualitative risk assessment, cannot be used for structuring conditional scenarios (prior to dependence assessments) as they do not consider dependent scenarios/events or extreme events. Nevertheless, some of their elements have been integrated in our method. While we already discussed achievements of the method in Section 6.3 when reflecting on the case-study, in this conclusion we focus on the method's current limitations and future research.

In this regard, a first remark is that more applications in case-studies are desirable in order to gather more insight on the suitability of our scenario mapping method's in other real world contexts. This will show in particular whether other experts (and more specifically experts with different backgrounds/practitioners in different industries) regard it as supportive and whether the previous findings (from Section 6) can be confirmed and even generalized after many more applications. Further, more applications in case-studies allow for a better understanding of the trade-off between intuitiveness and detail of the resulting scenario maps. For example Haines (2009), on whom our scenario definition is partly based, also discusses the importance of considering resilience in a system-based definition of risk. While resilience can be included in the current method, in future it might be of interest to explicitly incorporate events that make a system resilient, e.g. as a particular type of enabling conditions. This can then either be considered by experts as another, additional event type within their own conditional scenarios or different experts could assess resilience in that way for scenarios which are not their own - in that way taking the position of a *devils-advocate*. Further developing the method with the objective of becoming more intuitive and offering more detailed scenarios can be achieved by adding more scenario components or at least do more research on alternative components.

In addition to other event types/components and modified roles of the experts, another way of enhancing our method worth exploring in future applications is by more research on a suitable time frame. We mentioned previously that we consider this method as being more suitable for shorter time periods. However, for which time frame the method becomes unreliable is not well-explored. This research question might be combined with the one on alternative scenario component types by introducing components which are explicit for specific time steps into the future.

Lastly, additional research is of interest on the method's performance for extreme conditional scenarios, i.e. tail dependencies. While supporting experts in better understanding tail dependencies has been one of the main motivations for developing our method, some might argue that such extreme (joint) events cannot be mapped nor assessed sensibly. Similar to the difficulty of calibrating experts' performance on low-probability quantitative judgements, the corresponding scenario maps are also difficult to validate. In particular, one can argue about the plausibility of scenarios when they are on extreme events. Therefore, Wright and Goodwin (2009), the only scenario planning method focussing on extreme events, argues that an extreme scenario is plausible if it can be reached from the current state of the world. Similarly, it is essential for the future development of our method to include steps that challenge scenarios' plausibility in the feedback round when sharing the maps with the other experts. Currently experts have only used this part of the process for adjusting one's own assessments if desired. In that context, it is also important to further develop the method in line with new findings of descriptive research on assessing probabilities of extreme events as discussed in the previous section. This includes future research on complimentary ways to ensuring plausible scenarios.

Considering the above points in more case-study research supports the future devel-

opment of the method as an important step for dependence modelling.

References

- Ackermann, F. and Eden, C. (2006). Using causal mapping: individual and group; traditional and new. In: M. Pidd (ed.), *Systems Modelling: Theory and Practice*, pages 127–145. Chichester: John Wiley & Sons
- Ackermann, F., Eden, C., Williams, T. and Howick, S. (2007). Systemic risk assessment: a case study. *Journal of the Operational Research Society*, 58(1):39–51
- Ackermann, F., Howick, S., Quigley, J., Walls, L. and Houghton, T. (2014). Systemic risk elicitation: Using causal maps to engage stakeholders and build a comprehensive view of risks. *European Journal of Operational Research*, 238(1):290–299
- Ale, B. J. M., Bellamy, L. J., Cooke, R. M., Goossens, L. H. J., Hale, A. R., Roelen, A. L. C. and Smith, E. (2006). Towards a causal model for air transport safety—an ongoing research project. *Safety Science*, 44(8):657–673
- Altbach, P. G. and Knight, J. (2007). The internationalization of higher education: Motivations and realities. *Journal of Studies in International Education*, 11(3-4):290–305
- Anderson, J. L. (1993). Horizon mission methodology - a tool for the study of technology innovation and new paradigms. *NASA Working Paper AIAA Paper 93-1134*, Washington: NASA
- Bechlivanidis, C. and Lagnado, D. A. (2013). Does the why tell us the when? *Psychological Science*, 24(8):1563–1572
- Bedford, T. and Cooke, R. M. (2001). *Probabilistic risk analysis: Foundations and methods*. Cambridge: Cambridge University Press
- Bishop P., Hines, A. and Collins, T. (2007). The current state of scenario development: an overview of techniques. *Foresight*, 9(1):5–25
- Bradfield, R., Wright, G., Burt, G., Cairns, G. and Van Der Heijden, K. (2005). The origins and evolution of scenario techniques in long range business planning. *Futures*, 37(8):795–812
- Bunn, D. W. and Salo, A. A. (1993). Forecasting with scenarios. *European Journal of Operational Research*, 68(3):291–303
- Checkland, P. and Poulter, J. (2010). Soft systems methodology. In: M. Reynolds and S. Holwell (eds.), *Systems approaches to managing change: A practical guide*, pages 191–242. London: Springer
- Chelst, K. and Bodily, S. E. (2000). Structured risk management: Filling a gap in decision analysis education. *Journal of the Operational Research Society*, 51(12), 1420–1432
- Churchman, C. W. (1970). Operations research as a profession. *Management Science*, 17(2):3–37
- Cooke, R. M. (1991). *Experts in uncertainty: opinion and subjective probability in science*. New York: Oxford University Press
- Cooke, R. M. and Goossens, L. H. J. (2004). Expert judgement elicitation for risk assessments of critical infrastructures. *Journal of Risk Research*, 7(6):643–656
- Cooke, R. M. and Goossens, L. J. H. (1999). *Procedures guide for structured expert judgment. Project Report to the European Commission 18820*
- Crawley, F. and Tyler, B. (2015). *HAZOP: Guide to best practice*. Amsterdam: Elsevier
- Crook, J. and Moreira, F. (2011). Checking for asymmetric default dependence in a credit card portfolio: A copula approach. *Journal of Empirical Finance*, 18(4):728–742
- Cunha, A. A. R. and Morais, D. C. (2016). Analysing the use of cognitive maps in an experiment on a group decision process. *Journal of the Operational Research Society*, 67(12):1459–1468
- Dawes, R. M. (1988). *Rational choice in an uncertain world: the psychology of judgement and decision making*. Thousand Oaks: Sage Publications
- De Witte, K. and López-Torres, L. (2017). Efficiency in education: a review of literature and a way forward. *Journal of the Operational Research Society*, 68(4):339–363

- Dias, L., Morton, A. and Quigley, J. (2018). Elicitation: The Science and Art of Structuring Judgement, volume 261. Springer International Series in Operations Research and Management Science, Cham: Springer
- Ducot, G. and Lubben, G. J. (1980). A typology for scenarios. *Futures*, 12(1):51–57
- Eden, C. and Ackermann, F. (1999). The role of GDSS in scenario development and strategy making. In: String Processing and Information Retrieval Symposium and International Workshop on Groupware, pages 234–242. IEEE
- EFSA (2014). Guidance on expert knowledge elicitation in food and feed safety risk assessment. Technical report: European Food and Safety Authority, (EFSA) Journal 12(6)
- Einhorn, H. J. and Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, 99(1):3–19
- Ericson, C. A. (2015). Hazard analysis techniques for system safety. Hoboken: John Wiley & Sons
- Franco, L. A. (2006). Forms of conversation and problem structuring methods: a conceptual development. *Journal of the Operational Research Society*, 57(7):813–821
- French, S. (1995). Uncertainty and imprecision: Modelling and analysis. *Journal of the Operational Research Society*, 46(1):70–79
- French, S. (2015). Cynefin: uncertainty, small worlds and scenarios. *Journal of the Operational Research Society*, 66(10):1635–1645
- French, S. and Rios Insua, D. (2000). Statistical Decision Theory. New York: Wiley
- Haines, Y. Y. (2009). On the complex definition of risk: A systems-based approach. *Risk Analysis*, 29(12):1647–1654
- Haines, Y. Y. (2006). On the definition of vulnerabilities in measuring risks to infrastructures. *Risk Analysis*, 26(2):293–296
- Hanea, A. M., McBride, M. F., Burgman, M. A., Wintle, B. C., Fidler, F., Flander, L., Twardy, C. R., Manning, B. and Mascaro, S. (2017). Investigate Discuss Estimate Aggregate for structured expert judgement. *International Journal of Forecasting*, 33(1):267–279
- Hanea, A. M., Burgman, M. and Hemming, V. (2018). Idea for uncertainty quantification. In: L. C. Dias, A. Morton, and J. Quigley (eds.), Elicitation: The science and art of structuring judgement, volume 261, chapter 5, pages 95–117. Springer International Series in Operations Research and Management Science, Cham: Springer
- Hastie, R. and Dawes, R. M. (2001). Rational choice in an uncertain world: the psychology of judgement and decision making. Thousand Oaks: Sage Publications
- HEFCE (2001). Risk management: A guide to good practice for higher education institutions. Technical report: Higher Education Funding Council for England, UK
- HEFCE (2005). Risk management in higher education: A guide to good practice. Technical report: Higher Education Funding Council for England, UK
- Howard, R. A. (1989). Knowledge maps. *Management Science*, 35(8):903–922
- Howard, R. A. and Matheson, J. E. (2005). Influence diagrams. *Decision Analysis*, 2(3):127–143
- Joe, H. (2014). Dependence modeling with copulas. New York: CRC
- Johnes, J. (2015). Operational research in education. *European Journal of Operational Research*, 243(3):683–696
- Kahn, H. and Wiener, A. J. (1967). The Year 2000: A Framework for Speculation on the Next Thirty Three Years. New York: MacMillan
- Kahneman, D. and Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The psychology of intuitive judgment*, 49:49–81
- Kaplan, S. (1997). The words of risk analysis. *Risk Analysis*, 17(4):407–417
- Kaplan, S. and Garrick, B. J. (1981). On the quantitative definition of risk. *Risk Analysis*, 1(1):11–27
- Kolmogorov, A. N. (1933). Grundbegriffe der Wahrscheinlichkeitsrechnung, Ergebnisse der Mathematik. Zurich: Springer
- Ladyman, J., Lambert, J. and Wiesner, K. (2013). What is a complex system? *European Journal for Philosophy of Science*, 3(1):33–67
- Lane, D. C. (2000). Diagramming conventions in system dynamics. *Journal of the Operational*

- Research Society*, 51(2):241-245
- Lowe, D., Martingale, L. and Yearworth, M. (2016). Guiding interventions in a multi-organisational context: combining the viable system model and hierarchical process modelling for use as a problem structuring method. *Journal of the Operational Research Society*, 67(12):1481-1495
- MacKenzie, D. and Spears, T. (2014). The formula that killed wall street: The Gaussian copula and modelling practices in investment banking. *Social Studies of Science*, 44(3):393-417
- Mayston, D. J. (2003). Measuring and managing educational performance. *Journal of the Operational Research Society*, 54(7):679-691
- McNeil, A. J., Frey, R. and Embrechts, P. (2015). Quantitative risk management: Concepts, techniques and tools. New Jersey: Princeton University Press
- Midgley, G. (1992). The sacred and profane in critical systems thinking. *Systems Practice*, 5(1):5-16
- Mingers, J. and Rosenhead, J. (2004). Problem structuring methods in action. *European Journal of Operational Research*, 152(3):530-554
- Montibeller, G. and Belton, V. (2006). Causal maps and the evaluation of decision options - a review. *Journal of the Operational Research Society*, 57(7):779-791
- Morecroft, J. (2010). System dynamics. In: M. Reynolds and S. Holwell (eds.), *Systems approaches to managing change: A practical guide*, pages 25-85. London: Springer
- O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E. and Rakow, T. (2006). *Uncertain judgements: eliciting experts' probabilities*. Chichester: John Wiley & Sons
- Paté-Cornell, E. (2012). On 'black swans' and 'perfect storms': risk analysis and management when statistics are not enough. *Risk Analysis*, 32(11):1823-1833
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. San Francisco: Morgan Kaufmann
- Phillips, L. D. (1984). A theory of requisite decision models. *Acta Psychologica*, 56(1):29-48
- Poplawska, J., Labib, A. and Reed, D. M. (2017). From vicious to virtuous circles: problem structuring for quantified decision making in operationalization of corporate social responsibility. *Journal of the Operational Research Society*, 68(3):291-307
- Quigley, J., Colson, A., Aspinall, W. and Cooke, R. M. (2018). Elicitation in the classical model. In: L. C. Dias, A. Morton, and J. Quigley (eds.), *Elicitation: The science and art of structuring judgement*, volume 261, chapter 2, pages 15-36. Springer International Series in Operations Research and Management Science, Cham: Springer
- Raiffa, H. (2002). Decision analysis: a personal account of how it got started and evolved. *Operations Research*, 50(1):179-185
- Renn, O., Burns, W. J., Kasperson, J. X., Kasperson, R. E. and Slovic, P. (1992). The social amplification of risk: Theoretical foundations and empirical applications. *Journal of Social Issues*, 48(4): 137-160
- Rosenhead, J. (2006). Past, present and future of problem structuring methods. *Journal of the Operational Research Society*, 57(7):759-765
- Rowe, G. and Wright, G. (2001). Expert opinions in forecasting: the role of the delphi technique. In: *Principles of forecasting*, pages 125-144. London: Springer
- Salmon, F. (2009). The formula that killed wall street. *Wired*, 3:16-20
- Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley
- Shachter, R. D. (1988). Probabilistic inference and influence diagrams. *Operations Research*, 36(4):589-604
- Sloman, S., Barbey, A. K. and Hotaling, J. M. (2009). A causal model theory of the meaning of cause, enable, and prevent. *Cognitive Science*, 33(1):21-50
- Stamatis, D. H. (2003). *Failure mode and effect analysis: FMEA from theory to execution*. Milwaukee: ASQ Quality Press
- Stewart, T. J., French, S. and Rios, J. (2013). Integrating multicriteria decision analysis and scenario planning - review and extension. *Omega*, 41(4):679-688
- Suppes, P. (1970). *A probabilistic theory of causality*. Amsterdam: North-Holland Publishing

Company

- Tosoni, E., Salo, A., Govaerts, J. and Zio, E. (2019). Comprehensiveness of scenarios in the safety assessment of nuclear waste repositories. *Reliability Engineering & System Safety*, 188: 561-573
- Tosoni, E., Salo, A. and Zio, E. (2018). Scenario analysis for the safety assessment of nuclear waste repositories: A critical review. *Risk Analysis*, 38(4): 755-776
- Ulrich, W. (2003). Beyond methodology choice: critical systems thinking as critically systemic discourse. *Journal of the Operational Research Society*, 54(4):325-342
- Van der Heijden, K. (2011). *Scenarios: the art of strategic conversation*. Chichester: John Wiley & Sons
- Vesely, W. E., Goldberg, F. F., Roberts, N. H. and Haasl, D. F. (1981). *Fault Tree Handbook NUREG-0492*. Washington DC: Nuclear Regulatory Commission
- Werner, C. (2017). Report and R-tool manual: Conditional scenario mapping and dependence elicitation in higher education risk assessment. Technical report: University of Strathclyde, UK
- Werner, C., Bedford, T., Cooke, R. M., Hanea, A. M. and Morales-Nápoles, O. (2017). Expert judgement for dependence in probabilistic modelling: a systematic literature review and future research directions. *European Journal of Operational Research*, 258(3):801-819
- Werner, C., Bedford, T. and Quigley, J. (2018b). Sequential refined partitioning for probabilistic dependence assessment. *Risk Analysis*, 38(12):2683-2702
- Werner, C., Hanea, A. M. and Morales-Nápoles, O. (2018a). Eliciting multivariate uncertainty from experts: Considerations and approaches along the expert judgement process. In: L. C. Dias, A. Morton, and J. Quigley (eds.), *Elicitation: The science and art of structuring judgement*, volume 261, chapter 8, pages 171-210. Springer International Series in Operations Research and Management Science, Cham: Springer
- White, D. (1995). Application of systems thinking to risk management: a review of the literature. *Management Decision*, 33(10):35-45
- Williams, T., Ackermann, F., Eden, C., and Howick, S. (1997). Project risk: systemicity, cause mapping and a scenario approach. *Managing Risks in Projects*, 343-352
- Wilson, P. F. (1993). *Root cause analysis: A tool for total quality management*. Milwaukee: ASQ Quality Press
- Wintle, B. C., Fidler, F., Vesk, P. A. and Moore, J. (2013). Improving visual estimation through active feedback. *Methods in Ecology and Evolution*, 4(1):53-62
- Wright, G. and Goodwin, P. (2009). Decision making and planning under low levels of predictability: Enhancing the scenario method. *International Journal of Forecasting*, 25(4):813-825
- Yaniv, I. and Choshen-Hillel, S. (2012). Exploiting the wisdom of others to make better decisions: Suspending judgment reduces egocentrism and increases accuracy. *Journal of Behavioral Decision Making*, 25(5):427-434