DTU Library

Assessing the capabilities of advanced risk quantification methods for engineering systems management

Tegeltija, Miroslava

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Tegeltija, M. (2018). *Assessing the capabilities of advanced risk quantification methods for engineering systems management*.

Assessing the capabilities of advanced risk quantification methods for engineering systems management

Miroslava Tegeltija

PhD Thesis

May 2018.

Technical University of Denmark

Management Engineering Department

Engineering Systems Division

Main supervisor: **Josef Oehmen**
Associate Professor, PhD
Technical University of Denmark
Department of Management Engineering
Kongens Lyngby, Denmark

Supervisor: **Igor Kozin**
Senior Researcher, PhD
Technical University of Denmark
Department of Management Engineering
Kongens Lyngby, Denmark

Preface

The PhD thesis is the product of the PhD project carried out by Miroslava Tegeltija at Technical University of Denmark, Management Engineering Department, Engineering Systems Division. The three-year project started on February 1st, 2015 and ended on January 31st, 2018. This thesis is primarily manuscript-based. However, two of the publications (Tegeltija *et al.*, 2018a and Tegeltija *et al.*, 2018b) have been integrated into this thesis as chapters 7 and 8 respectively with only minor edits to harmonize the language. The following publications are part of the work presented in this thesis.

Peer-reviewed publications (published):

1. Post-Probabilistic Uncertainty Quantification: Discussion of Potential Use in Product Development Risk Management

Tegeltija, Miroslava; Oehmen, Josef; Kozin, Igor; Geraldi, Joana
Part of: 14th International Design Conference - Design 2016, pages: 533-542, 2016, Design Society,
Presented at: 14th International Design Conference, 2016, Dubrovnik
Type: Article in proceedings (Peer reviewed)
Status: Published | Year: 2016

2. Project risk management: potential in the field and the NUSAP scheme

Tegeltija, Miroslava; Kozin, Igor
Presented at: SRA Europe 2nd Nordic Chapter Meeting, 2016, Gothenburg
Type: Conference abstract for conference (Peer reviewed)
Status: Published | Year: 2016

3. Risk Management Challenges in Large-scale Energy PSS

Tegeltija, Miroslava; Oehmen, Josef; Kozin, Igor
Part of: Procedia CIRP, pages: 169-174, 2017,
Presented at: 9th CIRP IPSS Conference: 2017, Copenhagen
Type: Article in proceedings (Peer reviewed)
Status: Published | Year: 2017

4. Exploring Deep Uncertainty Approaches for Application in Life Cycle Engineering (Integrated as Chapter 7)

Tegeltija, Miroslava; Oehmen, Josef; Kozin, Igor; Kwakkel, Jan
Part of: The 25th CIRP Conference on Life Cycle Engineering
Presented at: 9th CIRP IPSS Conference: 2018, Copenhagen
Type: Article in proceedings (Peer reviewed)
Status: Published | Year: 2018

5. Risk management practice in construction: Case study Landssimareitur, Reykjavik

Tegeltija, Miroslava; Kozin, Igor; Gunnarsson, Bjarni
Presented at: SRA Europe 3rd Nordic Chapter Meeting, 2017, Eespo, Helsinki
Type: Conference abstract for conference (Peer reviewed)
Status: Published | Year: 2017

6. Tailoring Risk Management in Design (Integrated as Chapter 8)

Tegeltija, Miroslava; Oehmen, Josef; McMahon, Chris; Maier, Anja; Kozin, Igor; Skec, Stanko
Part of: 15th International Design Conference - Design 2018, pages: 533-542, 2018, Design Society
Presented at: 15th International Design Conference, 2018, Dubrovnik
Type: Article in proceedings (Peer reviewed)
Status: Published | Year: 2018

Acknowledgements

It would not have been possible for me to complete this thesis without the help of many people. I would like to extend my appreciation to all those who assisted me in this endeavor.

First and foremost, I would like to thank my main supervisor, Josef Oehmen, for his endless efforts, guidance, proactive supervision, ideas and patience (especially in the last stage of the project). I would also like to thank my other supervisor, Igor Kozin, for his undoubtable support, encouragement and consistent suggestions for improvement. I feel truly privileged and I am fortunate to have had them as my supervisors.

Not only did the project significantly benefit from the inspiring collaboration, but I would also like to thank Graeme Maxwell Keith for his input, inspiration and sharing his expertise, which motivated me to explore various scientific domains. For the challenging discussions, raising fundamental questions and support in pursuing this direction, I am grateful to Martin Bo Clausen, who contributed to both my professional and personal development. Thanks also to Søren Randrup-Thomsen, who provided me with many insights from a practitioner's point of view and in this way enabled me to strengthen the findings, discussions and contributions.

Very special thanks to Scott Ferson, who opened the doors of Applied Biomathematics, hosted my external research stay and contributed to this research at key stages. Working with Chris McMahon was an honor and it was a true privilege to receive his knowledgeable insights and advice. Thanks also to Kristin Wood for his input and suggestions. Robert Taylor's enthusiasm within the field stimulated a number of discussions in which I had the chance to hear about, and learn more from, his extensive experience.

I would also like to thank all my colleagues in the Engineering Systems Division at DTU Management Engineering, as well as many others across the Department and the whole University. In particular, I would like to acknowledge the valuable feedback and support I received from Sebastiano Piccolo, Pedro Parraguez, Melanie Kreye, Alexis Laurent and the head of our Division, Anja Maier.

To those who made this journey possible – Miljan Mikić, Giorgio Locatelli and Naomi Brookes – thank you.

Thanks to Žarko Zobenica, Mladen Miljuš, Slobodan Kacanski, Søren Friis and the absolutely adorable Katharina Oehmen, who made this experience unforgettable.

And last, but not least, I would like to thank my brother, my mother and my father for more than they will ever be aware of.

Thank you all.

“Prediction is very difficult, especially about the future”

- Niels Bohr -

Abstract

When risk management considerations are integrated into the engineering systems design, both overall system performance and quality of the developed solutions improve. A central part of integrating risk management in engineering systems design is to ensure that the design process benefits from employing risk and uncertainty methods with different levels of sophistication. That is, namely through the application of risk analyses that model risk and uncertainty in different ways. Traditionally, especially in engineering fields, risk analyses have largely been expressed in a quantitative, probabilistic form. However, such quantitative information, either as customized input to decision making or as general-purpose statistics, is itself becoming increasingly problematic and afflicted by severe uncertainty. Both the precision in estimates and the quality of background knowledge, on which probabilities are based, have been challenged in practice and academia.

This PhD thesis investigates advanced risk and uncertainty quantification methods in the context of engineering systems to better address, reflect, and utilize available information and background knowledge in design. The investigation was guided by the four research questions focusing on: 1) challenges in current design risk management quantification, 2) advanced risk and uncertainty methods, introduced under the non-probabilistic framework: the first group of methods is based on imprecise probabilities, the second represents a group of semi-quantitative approaches and the third group of methods is based on exploratory modeling, 3) prototypical applications of the non-probabilistic methods in different engineering systems design contexts, and 4) the transfer and integration of these methods and their results into overall risk management and associated processes. The results are presented in corresponding chapters from which four core findings are extracted: 1) currently widely used risk and uncertainty methods do not appropriately describe all uncertainty - especially uncertainty due to lack of knowledge, so called epistemic uncertainty – that remains a challenge, 2) advanced methods have been developed in other fields (i.e. outside of engineering design) to deal with similar issues and have provided valuable results in those fields, but have not yet been applied or tested in engineering design contexts, 3) for the engineering design situations and scenarios tested in this thesis, the non-probabilistic methods provided more credible representation of uncertainty, and 4) finding and employing a satisfactory quantification method from the available options is context dependent, and a broader process view needs to be considered when tailoring risk management to specific design situations.

This study contributes in four ways to the extension of our knowledge base on risk management in engineering systems design. First, the study synthesizes the challenges in current risk management from the literature and through empirical work regarding: modeling, quality of background knowledge and use and integration of results (research question 1/contribution 1). Second, this study systematically collects and categorizes advanced methods from the literature in other domains, conceptually develops them for the design context and provides a unique platform for their application through the non-probabilistic framework (research question 2/contribution 2). Third, it transfers these methods into usable tools through examples in case study applications in the oil and gas industry, followed by their comparison with several traditional probability approaches in representative situations (research question 3/contribution 3). Fourth, this study facilitates and enables a more adequate choice of a quantification method depending on the design context in question by developing a risk management tailoring approach (research question 4/contribution 4). The overall conclusion is that non-probabilistic methods have a high potential in engineering systems design, but their integration to the overall risk management and associated processes must be carefully and knowingly planned and carried out, to harness this potential and to achieve an actual design impact in practice.

Dansk Sammenfatning

Denne afhandling udforsker forskellige måder, som skal forbedre det videnskabelige felt risikostyring. Mere specifikt er målet med denne afhandling, at give en praktisk rettesnor og et solidt videnskabeligt fundament til at adressere, reflektere over og udnytte den tilgængelige information/viden indenfor kvantificering af risiko og usikkerhed. Når risikostyring integreres i engineering systems design styrker det kvaliteten af både systemets performance og kvaliteten af de løsninger der udvikles. Et central aspekt i at integrere risikostyring i engineering systems design er at sikre at design processen drager fordel af at benytte metoder til håndtering af risiko og usikkerhed som har forskellig grad af detaljering og sofistikation. Fordelen handler ofte om at benytte metoder der arbejder med risiko og usikkerhed på forskellige måder. Risiko styring har oprindeligt været overvejende kvantitativ og benyttet sandsynligheder, specielt inden for ingeniørkunst. Kvantitativ information som benyttes som input til beslutninger eller statistik, er imidlertid blevet mere problematisk at benytte og påvirkes af voldsom usikkerhed. Brugen af sandsynligheder er blevet udfordret af akademika og i praksis med hensyn til præcision i estimater og kvaliteten af den baggrundsviden som brugen bygger på.

Denne Ph.d. afhandling undersøger avancerede metoder til kvantificering af risiko og usikkerhed i relation til engineering systems, for bedre at kunne adressere, reflektere og benytte den information der er tilgængelig samt baggrundsviden inden for design. Afhandlingen blev guidet af fire forskningsspørgsmål som fokuserede på: 1) udfordringer ved nuværende kvantificering i risikostyring i design, 2) avancerede metoder til risiko og usikkerheds håndtering – som introduceres i forbindelse med ikke-kvantitative metoderamme i tre dele: Den første gruppe af metoder er baseret på upræcise sandsynligheder, den anden gruppe repræsentere en gruppe af semi-kvantitative tilgange og den tredje gruppe er baseret på eksplorative modeller, 3) prototypiske anvendelser af de ikke-probabilistiske metoder i forskellige engineering systems design kontekster, og 4) overførsel og integration af disse metoder og deres resultater til risikostyring og de tilknyttede processer. Resultatet præsenteres i tilsvarende kapitler hvorfra fire centrale resultater uddrages: 1) nuværende metoder til risikostyring og usikkerhed som bruges i vid udstrækning, beskriver ikke al usikkerhed på en passende måde – især vedrørende usikkerhed som skyldes manglende viden, såkaldt epistemisk usikkerhed, som forbliver en udfordring, 2) avancerede metoder er blevet udviklet inden for andre felter (uden for engineering design) for at adressere lignende udfordringer og har vist sig at være værdifulde inden for disse felter, men er endnu ikke blevet anvendt eller testet i engineering design kontekster, 3) det var de ikke-probabilistiske metoder som gav mest pålidelige resultater i de engineering design kontekster som blev testet, 4) at identificere en tilfredsstillende metode afhænger af konteksten og et mere vidtrækkende syn på processen er nødvendigt for at kunne tilpasse eller ’skræddersy’ risikostyringsprocessen til specifikke design sammenhænge og situationer.

Afhandlingen bidrager til forskningen og viden inden for risikostyring i engineering systems design på fire måder. For det første syntetiseres udfordringerne ved nuværende risikostyring fra litteraturen og gennem case studier vedrørende: modellering, kvalitet af baggrundsviden og brug af samt integration af resultaterne (forskningsspørgsmål 1/bidrag 1). For det andet samler afhandlingen systematisk avancerede metoder fra litteraturen inden for andre domæner og kategorisere dem, udvikler dem til brug i en design kontekst og bidrager med en unik platform til at benytte metoderne ved hjælp af det ikke-probabilistiske rammeværktøj (forskningsspørgsmål 2/bidrag 2). For det tredje overføres disse metoder til anvendelige værktøjer ved hjælp af eksempler fra case studierne indenfor olie og gas industrien. Metoderne sammenlignes med traditionelle tilgange som gør brug af sandsynlighed i repræsentative scenarier (forskningsspørgsmål 3/bidrag 3). For det fjerde facilitere afhandlingen valg af metode til kvantificering i relation til den konkrete design kontekst. Der udvikles en fremgangsmåde og rammeværktøj til tilpasning / skræddersyning af risikostyring (forskningsspørgsmål 4/bidrag 4). Den mest gennemgribende konklusion er at ikke-probabilistiske metoder har stort potentiale inden engineering systems design, men deres integration i risikostyring og de tilknyttede processer bør planlægges nøje og udføres refleksivt for at kunne tøjle og udnytte potentialet, samt opnå en reel indflydelse i praksis.

Table of Contents

Table of Figures

List of Tables

1. Introduction: Re-thinking risk quantification in engineering systems design

"Great Design sprouts when good research grows"

\- M. Cobanli -

This first chapter introduces the PhD thesis, starting with the overall motivation for the research and problem framing (Section 1.1), as well as outlining the industrial context surrounding the study. Section 1.2 highlights the challenges of designing engineering systems, the current industrial needs and knowledge gaps regarding advanced risk quantification methods, and their integration into the overall risk management and design process. Thereafter, Section 1.3 lays out the research objectives and the main research questions, based on the need and knowledge gaps introduced in the previous section, and links them to the research methodology that is further described in Chapter 2. Finally, Section 1.4 provides a brief outline of the structure of the thesis.

1.1. Motivation and problem framing: Risk and uncertainty quantification in engineering systems design need improvement

This section highlights the following key messages: 1) Many engineering projects fail to deliver in terms of time, cost and/or performance; 2) This means that there is great incentive to find methods – risk management – that help us assess the risk of failing to deliver and of mitigating those risks; risk management is part of project management and most engineers' training; 3) Such methods often need a quantitative estimate of the probabilities of adverse events and yet these are difficult to identify, especially in the early stages of engineering systems design; 4) This deficiency (lack of reliable quantitative probability estimates) is made especially difficult today by the trends toward "systems of systems" in engineering, as well as by megaprojects. Ironically, it is in megaprojects that the need for risk management is the greatest. For these reasons, this thesis investigates methods to improve risk and uncertainty quantification in current practices.

There is an ongoing and lively discussion in research communities on why engineering systems projects often face challenges to be on time, on budget, and on specifications. In particular, much focus is placed on those whose goal is a design of an engineering system (de

Weck, Roos, & Magee, 2011). According to de Weck, Roos, and Magee (2011) engineering systems are defined as: "*A class of systems characterized by a high degree of technical complexity, social intricacy, and elaborate processes, aimed at fulfilling important functions in society.*" The design, delivery and operation of engineering systems are normally executed through projects, including products in the systems – aircraft, computers, communications equipment etc.

The issue raised in this thesis is of key relevance for industry. The Project Management Institute estimated that more than 40% of engineering projects fail to meet their goals. "*We see US$122 million wasted for every US$1 billion invested due to poor project performance, a 12 percent increase over last year*." (PMI, 2016). The main message of their reports is that the cost of low performance is high. Apart from this substantial economic factor, engineering systems that fail to meet design specifications could potentially have a tremendous effect on the quality of thousands or millions of people's lives.

Engineering systems play an important role in society, but are also extremely risky (Locatelli & Mancini, 2010). Causality with a weak (i.e. often sub-optimal) phase of project planning has been researched before, since it leads to underestimations of the costs, overestimation of short-term benefits, and strategic misrepresentations (Flyvbjerg, 2006b). Merrow (2013) shows that the vast majority of large-scale projects could be considered a failure when considering adherence to schedule and budget as well as benefits in operation (Locatelli, 2018).

On the other hand, globalization brought integration of multiple engineering systems accompanied by designed and integrated services. However, it also brought higher competitiveness in the market. This consequently leads to time pressures, tighter budgets and a greater need for higher accuracy in estimates at an early design stage. This is even more the case when system-ilities are taken into account – properties concerning wider system impacts with respect to time and stakeholders (e.g. resilience, flexibility, adaptivity) (Chalupnik, Wynn, & Clarkson, 2013). In the case of large-scale engineering systems, the complexities, numbers of people involved, long life-cycles and enormous societal effects are even more evident. Therefore, such engineering systems are inherently more difficult to describe, understand, design, manage, and operate (de Weck, Roos, & Magee, 2011).

For instance, one example of poor performance is certainly the Berlin Brandenburg Airport. A number of challenges and changes in the design, and a rather long list of identified shortcomings led to tripled costs, now estimated to be €5.4 billion (Hammer, 2015). The learnings from this example pose the question: What can we do to better support decision making in similar cases? There is no doubt that making decisions in engineering systems design is a challenging task, but can we do better than this, and how can we make sure that the overall quality improves?

It has been suggested by both researchers and practitioners that the way we manage design solutions should keep pace with the complex and changing nature of engineering systems (Chang, Lee, & Chen, 2014). Risk management is an important tool to assess the environmental, financial, legal, technical and societal impacts of product, system and service designs to support achieving predefined goals. These changes lead to the increased importance of addressing uncertainty throughout the whole life cycle of a product, system or service. Uncertainty considerations are particularly relevant for the accuracy of planning models, and thus research, in that direction is of great significance for the field.

During the last decades, the management of risk in engineering systems design and associated projects and services has drawn attention from researchers and practitioners in areas such as engineering design (Lough, Stone, & Tumer, 2009), project management (Raz & Michael, 2001), and safety-related risk management (Paté-Cornell, 1996; Glendon, Clarke, & McKenna, 2016). The Project Management Institute represents the largest professional organization dedicated to the project management field, and identifies risk management as one of the ten main areas of project management (PMI, 2008). Furthermore, risk management courses are usually a part of most training programs for project managers. In accordance with the current view of project management as a life cycle process, project risk management is often perceived as a process that accompanies a project, from the initiation through the planning, execution, monitoring and control phases all the way to the completion and closure (Raz & Michael, 2001). Arguably, risk management has become an integral part of many formalized design processes for complex technical or socio-technical systems.

Despite this formalization of risk management in organizations, Flyvbjerg (2007) observed that the main challenges of large projects, including the design of engineering systems, are incomplete, inadequate, unreliable or misleading information. Decisions made during the design process have a significant impact on the strategic value of the asset delivered, and these decisions depend on the quality of the information on which they are based (Eweje, Turner, & Müller, 2012).

Furthermore, it has been shown by empirical studies (Levi, 1990; Sahlin, 2012) that the amount and quality of information used to develop probability and utility functions is an important factor when making decisions. For example, people tend to make different decisions if they are aware of the amount and quality of the data on which probability and utility assessments are based. Given that uncertainty plays an important role in decision making, it is notable that its quality improves if uncertainty is carefully addressed (e.g. Prelec & Loewenstein, 1991; Riabacke, 2006).

For the last few decades, the probability theory has gained popularity in many applications such as modeling and quantifying uncertainty in engineering systems. The development of probability as a measurement of uncertainty is based on an axiom that precise measurements of uncertainties can be made (Bernardo & Smith, 2009). However, the complexity of today's engineering systems has been increased by various requirements, such as high performance, efficiency, and cost reduction. Since a probabilistic risk and uncertainty quantification analysis requires extensive information, both scientific and engineering communities have recently realized that there are limitations to using probabilistic frameworks in their systems, and the precision of estimates has been challenged. Therefore, there is a need for exploring advanced methods that could overcome these challenges.

1.2. Identified need and knowledge gaps: The critical role of knowledge in uncertainty quantification

This section highlights that risk management tries to identify what we know about the ways in which things could go wrong, and the likelihood of such occurring; this is crucially dependent on our knowledge of the project, the engineering, the people involved etc. In large-scale engineering systems design this knowledge is very distributed among the participants. Hence, this section introduces the reader to epistemic and aleatory uncertainty, and ambiguity. Probabilistic methods function well in relation to aleatory uncertainty, but to a lesser extent with the other two. Engineering systems design in particular needs methods to deal with a lack of knowledge, and I will thus examine non-probabilistic methods that are nevertheless compatible with Bayesian/classic probability in this regard. The non-probabilistic methods have demonstrated reliable results in other

fields facing similar challenges, and are therefore worth exploring in the engineering systems design context. Methods that focus on ambiguity are outside the scope of the thesis. I will examine three different groups of non-probabilistic methods and test their application through case studies. This will show how they may be incorporated into usable design tools.

Almost a century ago, Knight (1921) made a distinction between risk and uncertainty. Both concepts are described in detail in Chapter 3, where I introduce the definition underlying this research. Risk can be defined as "*the effect of uncertainty on objectives"* (ISO, 2009).

Furthermore, two types of uncertainty can be distinguished: epistemic uncertainty and aleatory uncertainty (e.g. Helton & Burmaster, 1996). Epistemic uncertainty arises due to lack of knowledge and can be reduced by collecting and acquiring new knowledge. This is in contrast to aleatory uncertainty that is of stochastic nature, and therefore cannot be reduced. In addition to the types of uncertainty, there is the concept of ambiguity: it describes how factual statements may be interpreted differently by different individuals (Klinke & Renn, 2002).

Arguably, one of the key challenges in design risk management today is that uncertainty quantification relies heavily on probabilistic models (Flage *et al.*, 2014). While these are fully capable of describing aleatory uncertainty, they have been challenged when used to model epistemic uncertainty (Dubois, 2010) or ambiguity. If used this way, probabilistic approaches lead to violations of their initial assumptions and provide arguable precision in their results. This thesis examines the current state of the art in practice in six leading, large-scale companies in engineering systems design, and documents the existing challenges. That represents the basis for the first claim: current risk management practices need improvement, since we only use a subset of the quantification methods.

Understanding current design risk management challenges is a key element to providing usable tools to best support industry needs. Thus, this thesis relies heavily on established collaborations with practitioners from various engineering systems design domains. The collaborations were essential for understanding their risk management process requirements.

Based on this, better support for decision making in situations dominated by weak available information is documented and this is found to be a common issue for multiple engineering sectors. It is therefore essential to explore methods to better assess uncertainty caused by a lack of knowledge. This sets the basis for the second claim: literature could provide arguments to build theoretical reasoning, and bring formality to the choice of risk management methods and their application.

For the reasons mentioned above, it is necessary to investigate other risk and uncertainty quantification methods to advance the support of decision making in situations dominated by lack of knowledge. After introducing the research methodology in Chapter 2, this thesis first investigates the current state-of-the-art in engineering systems design risk management (Chapter 3). Second, the thesis systematically collects and introduces three groups of "non-probabilistic" risk and uncertainty quantification methods that promise to better address epistemic uncertainty, and discusses their possible application in the context of engineering systems design risk management (Chapter 4). These theories are not in conflict with Bayesian or classical probability but rather provide tools that complement probabilistic methods for risk assessment of systems when data are scarce. However, advanced methods to better deal with ambiguity in uncertainty quantification are beyond the scope of the thesis.

Third, the non-probabilistic methods represent three different angles of adding to the existing engineering systems design risk management thinking. These angles are presented below (each is introduced and analyzed through one representative approach):

1. Imprecise probabilities through Coherent upper and lower probabilities (Walley, 1991): expand the possibilities of established probabilistic risk quantification to reason more reliably with limited information on actual probability distributions. The approach allows decision makers to review and discuss coherent and plausible ranges of probabilities.
2. Semi-quantitative approaches through the NUSAP scheme (Funtowicz & Ravetz, 1990): this can be seen as an extension of established probabilistic modeling of uncertainty. NUSAP adds qualitative information to the uncertainty and risk analysis in a structured manner, informing the modeling, analysis and decision making process by making issues such as data origin, quality and key assumptions transparent.
3. A family of related approaches for dealing with uncertainty with their roots in exploratory modeling, here introduced through robust decision making (Lempert, Popper, & Bankes, 2003). The main principles of these methods are to explore a wide variety of relevant uncertainties, connect short-term targets to long-term goals, commit to short-term actions while keeping options open, continuously monitor the environment, and act if necessary.

Fourth, improving risk and uncertainty quantification is only one part toward achieving higher accuracy in estimates. The choice of risk quantification method and its integration into the overall risk management process play a crucial role and also need to be considered (Chapter 8).

1.3. Outline of research objective and research questions

Chapter 3 presents the theoretical background that will substantiate the choice of research questions in detail, complemented by empirically documented needs. The development of the research questions was an iterative process (see Chapter 2 for details), where the detailed research questions, particularly research questions 3.1, 3.2, and 3.3, were developed and refined during the exploration phase of this PhD project.

The overall research objective is derived from three interconnected research areas and defined as better design and delivery of engineering systems utilizing advanced risk and uncertainty quantification (introduced in this thesis under the non-probabilistic framework). Chapter 2 and Table 1 describe in detail all steps and rationale behind each of them.

1.4. Thesis Structure

The thesis consists of ten chapters complemented by references and appendices. In the following, the thesis structure is described in relation to the current state-of-the-art and its limitations, methodology, data, and research questions.

The remainder of this thesis is structured as follows:

- **Chapter 2** describes the employed research methodology in reference to the Design Research Methodology and its stages, details the research questions and their rationale, and provides information about the empirical studies.

- **Chapter 3** addresses research question 1. It provides an overview of the main schools of thought in uncertainty quantification. Moreover, it provides an understanding of the limitations of the currently most widely employed methods, followed by a number of challenges in practice that are documented through case studies. The chapter explains the motivation and sets the basis for investigating advanced methods for risk and uncertainty quantification.

- **Chapter 4** addresses research question 2. It introduces a number of advanced methods for risk and uncertainty quantification that promise to better cope with the challenges identified. These methods are introduced under the 'non-probabilistic' framework and are structured into three groups. The first group of methods is based on imprecise probabilities, the second represents a group of semi-quantitative approaches, and the third group of methods is based on exploratory modeling.

- **Chapter 5** addresses research question 3.1. It illustrates the problem of imprecision and how we can employ the first group of non-probabilistic methods (imprecise probabilities) to better support decision making. The chapter compares an imprecise probability method, i.e. a probability bound analysis, to several traditional subjective probability approaches for a case study in the oil and gas industry.

- **Chapter 6** addresses research question 3.2. It focuses on the representation of the background knowledge in risk and uncertainty assessment. Based on a case study, a set of methods from the second group of non-probabilistic approaches (semi-quantitative approaches) are applied to visualize uncertainty surrounding data and results. In addition, calculations are developed to quantify and correct biases in expert judgment in risk assessments, as well as qualitative approaches to inform decision makers' levels of trust in risk quantifications.

- **Chapter 7** addresses research question 3.3. It further investigates the third group of methods (exploratory modeling) through one representative approach, robust decision making, for the challenges related to the life cycle aspects. The challenges of using the method in the engineering systems design context are documented, and conceptual suggestions to overcome them are proposed.

- **Chapter 8** addresses research question 4. It provides practical guidance for tailoring risk management. Different needs for risk and uncertainty quantification are discussed, and concrete suggestions are provided for designing a risk management process and choosing a risk or uncertainty quantification method through a number of representative examples.

- **Chapter 9** deals with the broader discussion of the integration of this work with the current state-of-the-art. The chapter compares the presented methods with several other, widely used methods. This is followed by recommendations and overall research limitations.

- **Chapter 10** concludes and summarizes this thesis and includes a reflection on the theoretical and industrial contributions, research and managerial implications, and recommendations for future research.

2. Research methodology

"Not everything that is faces can be changed, but nothing can be changed until is faced"

- James Baldwin -

This chapter describes the research methodology applied to this doctoral study. In addition to the system of methods employed to acquire, analyze, and interpret empirical data, this chapter introduces the logic behind the selected methods in connection to the theoretical research approach and its limitations (Blessing & Chakrabarti, 2009).

The following discusses the methodological approach and research design and provides a detailed description of each research stage. The chapter is structured as follows: Section 2.1 provides a short review of theoretical and empirical considerations related to the methodology and frameworks. Section 2.2 further describes the research objectives and questions introduced in Chapter 1. Section 2.3 describes the overall stages of the applied Design Research Methodology (DRM), and Section 2.4 introduces the case studies and the strategies utilized for data gathering, analysis and interpretation. Section 2.5 provides a summary of the presented research methodology.

2.1. Theoretical and empirical approach

Given that the main research paradigm is within design research, the thesis's methodology is built on the Design Research Methodology (DRM). DRM is also highly suitable, as it supports both the literature-based (i.e. Chapters 3, 4, 8) and the empirically-based elements (i.e. Chapters 5, 6, 7) of this thesis. Also, the projects used in the empirical components of this thesis are either directly situated in a design context, or design new products and/or services in their respective contexts (for the overview of the empirical data sources and design challenges in question see Section 2.4). The methodology allows a systematic approach for conducting design research, with the overall aim "*to make design more effective and efficient in order to enable design practice to develop more successful products…*" (Blessing & Chakrabarti, 2009). Such design research/design science is aimed at improving, which is expressed as "*the purpose of design science is to raise quality of designing and designs …*" (Argyris & Schon, 1989). This is accomplished by a focus on both creating an *understanding* of the phenomenon in design, and the development and validation of *support* to improve design practice. Therefore, the DRM framework allows researchers to generate insights into design practice, and by developing different support tools strive for a changed, further developed, and improved design (practice) (Figure 1).

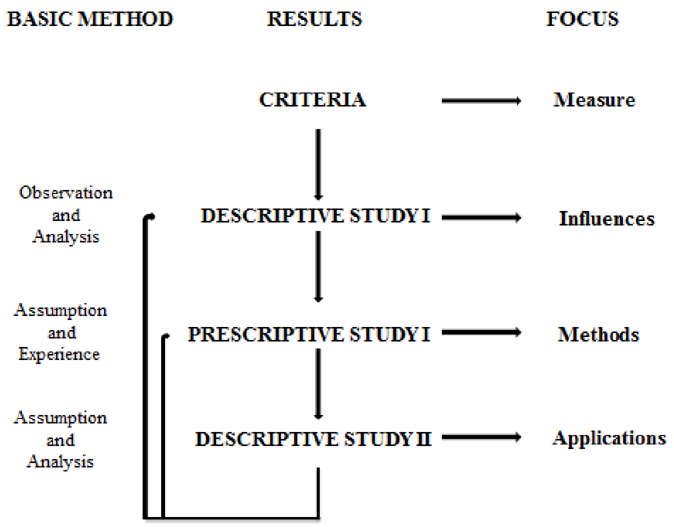


Figure 1 Design Research Methodology (Blessing & Chakrabarti, 2009).

Research design

As Chapter 1 indicates, this study uses both qualitative and quantitative methods. This research paradigm is known as a mixed-methods approach. The mixed-methods approach has its strength in getting the best from the two worlds (Johnson & Onwuegbuzie, 2004). Even though the research process seems like a linear sequence of steps, in practice it followed a series of iterative steps and customizations toward industry needs.

The initial literature review on the various limitations in terms of current risk management practices sets the basis for collaboration with industry. In fact, during the project, interaction with six participating companies took place over the course of three years in multiple ways, at different levels, and with varying goals. The research project's first explorative studies were accompanied by the use of semi-structured interviews. This generated insights from the industry and real-world practice, provided details on specific risk quantification analyses, and enabled the articulation of knowledge gaps and current risk methods limitations. These examples lead the interviews in directions of interest to both the interviewees and the interviewers. Hereafter, follow-up meetings, student projects, conferences, forums, single-person interviews, and two synthetic case studies for the actual test of the developed support from this project increased the involvement with and verification from industry.

A number of analysis methods were used throughout the research project, with iterative exploration as the backbone. The process included exploratory meetings with industry stakeholders, and an initial literature review in order to identify needs and knowledge gaps. Identification and organization of the main industry needs and knowledge gaps helped to align academic and industry needs and goals. Initial formulation of research objectives and research questions clarified the scope of the project. Exploratory case studies and a second more focused literature review were carried out to additionally enrich the non-probabilistic framework. Furthermore, the case studies were introduced in which the methods were applied. At the end, analysis of the results and evaluation of the framework's ability to address industrial needs and knowledge gaps were documented and discussed with practitioners.

2.2. Research objectives and research questions

Before implementing the research design and verifying whether the research questions could be answered within the scope of this project, the constructs used in the questions had to be developed, detailed and described. This was done mainly to investigate if the constructs can be measured, and even more so to establish which methods were suitable for the use in each of the DRM research stages (see Table 1).

The identified needs and knowledge gaps acted as drivers for the thesis and were used to determine the goal of this study as well as concrete and feasible research objectives. The main unresolved issue for each of the four research objectives was phrased as a research question, which also defined the shape of the expected outcomes. Finally, the results/outcomes were evaluated based on academic and industrial success criteria.

The overall research aim, derived from three interconnected research areas, is defined as **better design and delivery of engineering systems utilizing advanced risk and uncertainty quantification**. These research areas are first introduced in Sections 1.2 and 1.3 and further detailed in Chapter 3. As explained in Chapter 1, in the research questions I refer to the methods as risk quantification in order to shorten/simplify the wording. In order to achieve the aim, four more specific aims were articulated: A) to understand and document the current state-of-the-art in current engineering systems design risk management; B) to collect and conceptually develop the non-probabilistic framework; C) to provide prototypical applications of the introduced (non-probabilistic) methods, and D) to discuss and develop the integration of these methods into the overall risk management processes.

Table 1 Research questions and respective research aims and objectives

Thesis	**RQ 1**	**RQ 2**	**RQ 3**			**RQ 4**
Research question	What are the challenges with current engineering systems design risk management methods?	What advanced risk quantification methods exist that have not been widely used in the engineering systems design context?	How can advanced risk quantification methods be transferred into usable tools? During the thesis the question was refined into the following sub-questions (see Chapter 4):			How can we effectively integrate advanced risk quantification methods into the overall risk management process?
			How to use (deploy) imprecise probabilities in expert judgment elicitation?	How to use the NUSAP tool to treat and manage uncertain assumptions?	What methods exist to support long-term decision-making in early design when facing severe uncertainties and scarce information?	
Aim and objective	An overview of the current state-of-the-art in the field	Collection of the methods and framing of the non-probabilistic framework	Prototypical application of an advance quantification method	Representing background knowledge and information	Exploring approaches for coping with deep uncertainty	Tailoring risk management (based on the maturity of risk management)

A) An overview of the current state-of-the-art in engineering systems design risk management

The first research objective is to provide an overview of the current engineering systems design risk management, articulate the challenges in practice, and conceptually guide the remainder of the research. In terms of knowledge gaps, this objective stems from difficulties in quantifying and managing uncertainty (i.e. epistemic uncertainty) in design due to its nature, and a lack of information at the early stages of design projects. From an industrial point of view, the objective emerged from insufficient understanding of uncertainty types and respective quantification requirements of associated risks, as well as inability of the available methods to address all the challenges practitioners face.

The main problem was translated into **research question 1: What are the challenges with current engineering systems design risk management methods?** The answer to the first research question should articulate problems with current risk management processes and most employed risk quantification methods, document the challenges in practice, and clarify the main risk quantification theoretical concepts.

B) Conceptual development of the non-probabilistic framework and collection of the methods

The second research objective was to collect advanced risk quantification methods developed in other fields that have a potential to address documented challenges in design, and to conceptually develop the non-probabilistic framework that provides a clear structure for gathering advanced methods. In terms of knowledge gaps, this objective was triggered by the latest advancements in other domains (such as mathematics, artificial intelligence, safety engineering, water management, etc.). From an industrial point of view, this objective emerged from the need to more thoroughly analyze the way we cope with epistemic uncertainty and to achieve that in a more systematic way.

The main problem was translated into **research question 2: What advanced risk quantification methods exist that have not been widely used in the engineering systems design context?** The answer to this question should provide the unique non-probabilistic framework that offers a clear structure and collection of the advanced methods.

C) Prototypical applications of the non-probabilistic methods

The third research objective was to provide means for the non-probabilistic approaches to be applicable in practice by transferring them into usable tools. These can be seen as a toolkit from which, depending on the design challenge or the risk quantification aspect we want to improve, the use of a specific method is recommended. In terms of knowledge gaps, case studies were conducted to demonstrate the potential of non-probabilistic approaches, as they need to be adjusted to the particular design needs. This objective emerged from an industrial point of view, as the more mathematically advanced methods need to be manageable for the practitioners; they can utilize the potential of the powerful computers now available in companies.

The main problem was translated into **research question 3: How can advanced risk quantification methods be transferred into usable tools?** During the thesis the question was refined into the following sub-questions, each corresponding to one group of methods from the non-probabilistic framework:

Research question 3.1: How to use (deploy) imprecise probabilities in expert judgment elicitation? The answer to this question should introduce the imprecise method reasoning and provide an example of how a method from the first group of non-probabilistic methods can be applied in practice.

Research question 3.2: How to use the NUSAP tool to treat and manage uncertain assumptions? The answer to this question should provide an example of where and how a method from the second group of non-probabilistic approaches can be used.

Research question 3.3: What methods exist to support long-term decision-making in early design when facing severe uncertainties and scarce information? The answer to this question should raise awareness of the advancements in IT that now allow advanced simulations. A set of methods is introduced and one particular approach from the third group of non-probabilistic methods is discussed – robust decision making.

D) Tailoring risk management (based on the maturity of a risk management process) to specific quantification needs

The fourth research objective is to provide a wider process view of the integration of risk quantification methods into the overall risk management process. In terms of knowledge gaps, this objective emerges from a broad range of risk management processes as well as different practical design challenges. From an industrial point of view, companies can have different risk management maturity levels, which is why support is needed in planning process improvements.

The main problem was translated into **research question 4: How can we effectively integrate advanced risk quantification methods into the overall risk management process?** The answer to this question should provide a tailoring approach that ties risk quantification methods to the overall risk management process (proposed in ISO 31000) as a basis for systematic improvement of risk management. What needs to be improved is not only the quantification itself, but also its communication and its integration into the overall process.

2.3. Design Research Methodology stages

The DRM consists of four stages that were followed to structure the thesis and guide the research process: research clarification (Chapters 1 and Section 3.1), descriptive study I (Section 3.2 and Chapter 4), prescriptive study, and relative initial stage of descriptive study II (Chapters 5, 6, 7, 8). While a descriptive study focuses on investigating and describing problems, a prescriptive study develops support that addresses those problems. The first three DRM stages were associated with at least one research question (Table 2). The last stage, descriptive study II, focuses on "*the impact of the support and its ability to realize the desired situation*" (Blessing & Chakrabarti, 2009), and therefore this stage evaluates whether the success criteria were met. For this reason, this stage is discussed as the final part of each of the Chapters 5, 6, 7, and 8.

Table 2 Research questions and research stages according to DRM

Thesis Chapter	Research Question	RC	DS I	PS	DS II
Chapter 3	**RQ 1:** What are the challenges with current engineering systems design risk management methods?	●	●		
Chapter 4	**RQ 2:** What advanced risk quantification methods exist that have not been widely used in the engineering systems design context?		●		
Chapter 5	**RQ 3.1:** How to use (deploy) imprecise probabilities in expert judgment elicitation?			●	●
Chapter 6	**RQ 3.2:** How to use the NUSAP tool to treat and manage uncertain assumptions?			●	●
Chapter 7	**RQ 3.3:** What methods exist to support long-term decision-making in early design when facing severe uncertainties and scarce information?			●	●
Chapter 8	**RQ4:** How can we effectively integrate advanced risk quantification methods into the overall risk management process?	●		●	●

The goals and the work carried out at each stage are summarized as follows:

Research clarification (RC)

The goal of this stage was to define the key research problems, research objectives, theoretical focus and research questions, as well as to identify potential models and methods to answer the research questions. The research clarification process was conducted iteratively, defining a set of research goals and questions and adjusting these after Descriptive study I. A preliminary literature study, based on state-of-the-art design risk management research combined with discussions with industry practitioners and researchers, supported the research clarification process and the literature study through the entire research project (reported in Chapter 3).

This stage included an inductive process of increasing the level of abstraction, where literature gaps and the discovered needs in practice are framed into approachable research questions and linked to the research methodology. The intention was to provide a holistic exploration of the problem space that could subsequently be transferred to a suitable model. The main results of this stage can be found primarily in Chapter 1, Chapter 2 and Chapter 3.

Descriptive study I (DS I)

The goal of this stage was to develop, refine and propose the non-probabilistic framework. In order to do so, two exploratory in-depth case studies and interviews (Yin, 2013) were carried out in six companies, as well as a literature review (Webster & Watson, 2002). This particular stage was also vital in the sense that it was used to align research and industry goals, and to establish a sound platform to articulate the problems in a form that corresponds to the scope of the project.

In order to gain an in-depth understanding of the industrial state-of-the-art in terms of design risk management, two in-depth case studies (Yin, 2013) were developed. As described in Section 3.2, the exploratory case studies look into two different, but currently well used risk management tools (Primavera and RamRisk). The studies document some of the limitations in their application as well as the challenges regarding the input for decision making.

These studies were accompanied by a literature review (Webster & Watson, 2002) on different risk quantification theories and their limitations. Reviewing the theoretical foundations set the basis for articulating the need to more thoroughly explore various ways to represent epistemic uncertainty.

In addition, interviews were carried out with the two case companies but also with four other companies (Yin, 2013) in order to broaden the understanding of various design challenges and related uncertainties and the applicability of the collected and developed methods. If was further confirmed that the tools employed in exploratory studies are part of current best practices. Details regarding each set of interviews are available in Section 3.2. The interviews were coded and analyzed in ATLAS.ti using grounded theory approach (Corbin & Strauss, 1990).

The results of this stage, conceptual development of non-probabilistic framework in design, are found in Section 3.2 and Chapter 4. This stage primarily addresses research question 2 by building upon research question 1, because the objective was to develop the framework and collect advanced risk and uncertainty quantification methods through iterative work with the case companies.

Prescriptive study (PS)

In this thesis PS is the basis for answering research question 3 (i.e. 3.1, 3.2 and 3.3). The objective of this stage was to develop concrete means to support design risk management and decision making by utilizing advanced risk and uncertainty quantification. To achieve that, the research question was divided into three sub-questions, each corresponding to one group of methods from the non-probabilistic framework developed in Chapter 4.

This stage includes the development of two case studies and one conceptual development of the approaches toward the needs in the field. The first step of the development of the case studies was to generate the data needed for the analyses in accordance with the company's processes. Second, advanced risk and uncertainty analyses were performed and compared with some of the already existing approaches. Finally, the results were presented to the practitioners; the feedback is documented and elaborated in the corresponding Chapters 5 and 6. In the case of the only conceptually developed support for the design needs (Chapter 7), the actual synthetic case development was not included due to the limitations of this study and data availability. Chapters 5, 6, and 7 include the main results of this stage.

To answer research question 4, this stage aimed to develop decision making support in the form of a tailoring approach (Chapter 8). To do so, first the literature review on risk management maturity models was introduced as a basis to further expand on an existing maturity model. Additionally, the developed tailoring approach was tied to the overall ISO 31000 risk management process, allowing immediate implementation of the practices basing their risk management process on this standard. The main insights from this stage are elaborated in Chapter 8.

Descriptive study II (DS II)

The objective of this stage was an initial evaluation of the support developed during the prescriptive studies. A qualitative assessment sought to discover whether the support did or could improve the companies' processes during the prescriptive case studies. This was organized through the follow-up interviews and presentations at two companies, followed by practitioners' feedback. However, the confidentiality, data accessibility, and time considerations of this doctoral study did not allow for a comprehensive evaluation, including a wide implementation of the proposed approaches in the companies' processes.

This stage aimed to demonstrate the overall and preliminary findings, and prepare the results for others to pursue additional studies in this direction, highlighting some potential limitations or challenges. This corresponds to the definition of an initial study by Blessing and Chakrabarti (2009): "*An initial study closes a project and involves the first few steps of a particular stage to show the consequences of the results and prepare the results for use by others.*" The results can be found in the final subsections of Chapters 5 to 8 (Sections 5.5, 6.5, 7.5 and 8.4).

The following elaborates more on the empirical work conducted in this thesis.

2.4. Empirical studies

This section introduces the companies involved in this study, the developed case studies and the conducted interviews (Table 3).

Table 3 Overview of industry engagement and corresponding research method

Method of empirical data collection	RQ1 + RQ2		RQ 3.1	RQ 3.2	RQ 3.3	RQ 4
	Exploratory in-depth case studies	Interviews	Case study 1	Case study 2	Interviews	Interviews
Company 1		●			●	●
Company 2		●	●	●		●
Company 3	●	●				●
Company 4						●
Company 5		●				●
Company 6	●	●				●

Interviews

Interviews were used as one of the main research methods throughout the research project (RAND Corporation, 2009). Semi-structured interviews (RAND Corporation, 2009) were organized in such a way that information regarding the companies' risk management process, design process and current challenges were documented. The questions were developed around open, hypothetical, or comparative lines of questioning (Kvale, 2008). Moreover, follow-up interviews (Kvale, 2008) were organized when needed.

In total, 36 interviews were conducted. Details regarding each of the interviews, coding principles and interviewees can be found in Section 3.2 and Chapter 8. The following sections will briefly introduce the six companies that were involved in the interviews and/or case studies, in particular their area of work and design challenges.

Company 1: Design of large-scale engineering systems

The first case relates to a large Danish company with extensive experience in designing and managing large-scale engineering projects, such as long-life cycle infrastructure systems. It works on projects of different sizes (from megaprojects to small design solutions). For instance, it designs for first-of-a-kind engineering projects in which it faces severe uncertainties, but also helps small practices achieve their goals. Its risk management approach needs to provide support for the whole spectrum of different design activities and to ensure proper and timely response and monitoring.

Company 2: Oil and gas exploration, designing new systems

The exploration and commercial production of oil and gas is the main business of the second case company. A significant risk in the design and early execution of a new production project is the placement of exploration drill wells. The objective is to find a new oil or gas field, based on a sound analysis of the prospect's risks and potential hydrocarbon volumes: what is the chance that a well will find (contain) hydrocarbons, how much could be there? The design challenges are to understand the best process and infrastructure design to explore and exploit these fields. The company explores different locations and prospects, and its performance depends directly on the success rate of drilling, determined in the early design phase of the project. Test drillings are very expensive and represent a significant investment. To increase the success rate with regard to identifying prospective oil deposits, the opinions of multiple experts are solicited as part of the early project design risk management. Given that the subsequent detailed design of the whole production system is based on these analyses, attaining higher accuracy in the estimates is of great engineering and financial importance.

Company 3: SME, design in construction

The third organization is an engineering and consulting SME that provides design services for construction projects. It experienced several risks in the design phase and sees severe delays in its currently most challenging project.

Company 4: Consultancy for the Design phase

This international, multidisciplinary engineering consultancy company is an example of an organization that provides design services for construction projects. It provides consultancy services for projects such as design of airports, design of transportation systems, hospitals and similar. It also constructs some projects of its own.

Company 5: Public Organization

This international organization provides design services for a number of different projects and systems. It provides services for other NGOs, governments, the private sector and private foundations. It mainly focuses on procurement services, project management, and infrastructure. In addition, it offers some financial management services (such as managing grants) and human resources (some organizations sometimes outsource their recruitment process). It is currently designing its risk and quality framework. The biggest challenge is to design a framework for the whole spectrum of its practice (applicable and manageable for those working in the field in war zones as well as for desk workers).

Company 6: Large-scale high-tech infrastructure design in energy sector

The sixth case company is involved in designing and deploying large-scale high-tech infrastructure in the energy sector. Designing and operationalizing both onshore and offshore systems is part of its expertise.

Brief description of exploratory case studies (with companies three and six)

The exploratory case studies (Yin, 2013) form a coherent body of work with a company for each study. This includes a number of continuous empirical engagements that are detailed below.

Exploratory case study 1:

To document the current challenges in design risk management practice, an exploratory in-depth case study (Yin, 2013) was conducted with a case company involved in designing and deploying large-scale high-tech infrastructure in the energy sector. Its risk management is recognized as one of the best practices due to its advanced way of dealing with risk and uncertainties throughout the process, as well as the adopted and developed tools and decisionmaking processes. The collaboration also included interviews with the company's senior project risk manager, as well as the analysis of the implementation of a complex, quantitative engineering design and deployment project risk model in Primavera. The key insights of the interviews and the analysis are described in Section 3.4.

Exploratory case study 2:

The second exploratory case study was developed to document the potential, but also the limitations of currently one of the most employed risk tools in Nordic risk management practice. RamRisk was used to conduct the analysis for a design phase of a construction project. Details regarding the project, RamRisk, the analysis and the findings are available in Section 3.4.

Case study 1 and 2 (with Company two)

Overall, the research project includes two different synthetic case studies. Due to confidentiality reasons, there was no opportunity to analyze real project data. However, based on the work with the industry partners, we developed similar and representative cases without revealing any confidential information.

Yin (2013) describes a case study as a research strategy within social science research, with different case study types to be selected. Case studies are chosen to: "*Investigate a contemporary phenomenon within a real-life context. Especially when the boundaries between phenomenon and context are not clearly evident.*" Further details and design of the case studies, as well as risk management context of the studies, are available in Chapter 5 and Chapter 6. The work was carried out with the large Danish oil and gas company.

2.5. Chapter summary

The chapter covers theoretical and practical considerations related to the key methodological choices, and in particular the rationale behind the focus on risk and uncertainty quantification and its representation. Furthermore, the chapter details the research aims and objectives, and research questions that narrowed the scope and organized research. Also, the design research methodology stages are described and linked to chapters in the thesis, outlining the research questions. Finally, this chapter provides information about the case companies and the specific methods used during development and application of the non-probabilistic framework.

3. Theory of risk quantification and current state-of-the-art in risk management practice

"In theory, theory and practice are the same. In practice, they are not."

– Albert Einstein –

This chapter introduces the key theoretical foundations that create a basis for the work presented in the following chapters. More specifically, it presents current perspectives and trends on risk and uncertainty quantification. The aim is not an exhaustive presentation, but rather a presentation of what recent authoritative sources describe as state-of-the art thinking on risk.

Risk assessment research has traditionally focused on the development of probabilistic methods, tools and procedures for risk management and risk analysis (see e.g. Kaplan & Garrick, 1981; Dubois & Prade, 2009; Goerlandt & Reniers, 2015). This consequently led to tendencies in both research and practice to make risk assessment into a well-defined operation for evaluating different hazards, technologies and safety issues (Renn, 1998). The problem with such routinization of risk assessment is that formal analysis may obscure a number of the conceptual foundations and limitations of the methods used (Aven & Anthony, 2015). Additionally, it can also lead to a false degree of certainty when dealing with human actions and interventions (Ferson & Ginzburg, 1996). This chapter highlights main strengths and weaknesses of the current view on risk and asks whether that view is still feasible. Three main aspects are discussed: modeling, data, and human behavior; other general challenges are summarized. The chapter concludes that given the increasing scope of large-scale systems (or systems of systems), the field needs to more thoroughly consider concepts and theories that promise to overcome current limitations in the way we deal with uncertainties. Chapter 4 further explores possibilities for overcoming these challenges through advanced risk and uncertainty quantification.

This thesis investigates advanced risk and uncertainty quantification methods. In that context, this chapter represents the first step: it describes the reasons, needs and motivation for advanced risk and uncertainty quantification by documenting the challenges in current practice (answering research question 1), based both on the current state in literature, as well as on empirical work.

The chapter is structured as follows: Section 3.1 provides a short introduction to risk. Section 3.2 further introduces and describes uncertainty. Section 3.3 outlines the current state-of-the-art in engineering systems risk management. In Section 3.4 the exploratory in-depth case studies are described and challenges in current practice are documented. In Section 3.5 I open the discussion (that continues in Chapter 4 where the methods are introduced) for the need to explore alternative approaches.

3.1. Definitions of Risk

Attempts to manage risks should first start from attempts to answer the question: "*What is risk?*" Risk is ubiquitous in almost every human activity (Bernstein, 1996). We talk about the risk of a terrorist attack, risk of losing an investment, risk of falling from a ladder, risk of being involved in a traffic accident, risk of contracting a disease, risk of bankruptcy, risk of extinction of certain plant species, and so on. These are very different situations, but they share some common elements. First, people talking about them care about the outcomes. They are concerned about a terrorist attack that can happen and jeopardize their own or other people's lives and property; they can lose their savings or investments, which can even result in bankruptcy; they can fall and injure themselves; they can become involved in a traffic accident and either be injured or lose their lives; they can have a disease that may influence the quality of their lives, etc. That is to say, talking about risk is pertinent when a person, a group of people, an organization, or a whole society can be exposed to something they do not want to be exposed to (Fischhoff, 1995). They want to avoid being exposed to negative consequences of their or others' activities, that is, they do not want to lose or jeopardize something that they value: their lives, property, health, environment, valuable items, including money, etc.

Ironically, being exposed to the possibility of unwanted events can be a voluntary and desirable thing (Rowe, 1975). In the past, the risk of a ship sinking or being robbed by pirates on the way from Europe to India was offset by the rewards from selling cargo brought back to Europe. People risk their lives in return for the benefits they can get. High risks can simply be taken in return for emotional pleasure, honor, and fame. An example is extreme sports. In this regard, there is a strong consensus among risk theorists that a risk definition should accommodate both undesirable and desirable outcomes (Aven & Renn, 2010).

Second, people do not know whether the future unwanted events they may be exposed to will happen or not. This means that there is uncertainty about these events happening. Hence, ***exposure*** and ***uncertainty*** are two essential components that constitute risk (Holton, 2004).

Suppose a man leaps from an airplane without a parachute. If he is certain to die, he faces no risk. Risk requires both exposure and uncertainty. There is no uncertainty here. Or we can assume he does not value his life. Hence, he faces no risk in losing something that does not have any value to him.

We can look at the concept of risk in different ways. As soon as we have some objectives, aims, or targets we run a risk of not fulfilling, achieving, or hitting them. A project's objectives defined upfront may not become fulfilled at the project end. Hence, we can say there is risk of not fulfilling them. Aims, objectives and targets and uncertainty about their achievements are also the components of risk. From this angle and concisely, risk can be defined as *the effect of uncertainty on objectives* (ISO, 2009).

There are a number of other definitions of risk that are in line with the above commonly accepted foundational components of risk. What is common for them all is that they address exposure and uncertainty and accommodate both desirable and undesirable outcomes. A range of different definitions used today is summarized by Aven (2011), Kreye (2011) and Aven *et al.* (2015). Renn (1998) summarizes three underlying questions that a proper definition should cover (a similar set of questions is formulated by Kaplan and Garrick (1981)):

1. What are undesirable outcomes and who determines what undesirable means?
2. How can we specify, qualify or quantify the possibilities of undesirable outcomes?
3. How do we aggregate different classes of undesirable outcomes into a common concept that allows comparisons and the setting of priorities?

In the following, I provide an overview of various definitions of risks (following Aven, 2011):

1) Risk = Expected value (loss)

a) The risk of losing any sum is the reverse of expectation, and the true measure of it is the product of the sum adventured multiplied by the probability of the loss.

b) Risk equals the expected loss.

c) Risk equals the product of the probability and utility of some future event.

d) Risk equals the expected disutility.

2) Risk = Probability of an (undesirable) event

a) Risk is the chance of damage or loss.

b) Risk equals the probability of an undesirable event.

c) Risk means the likelihood of a specific effect originating from a certain hazard occurring within a specified period or in specified circumstances.

3) Risk = Objective uncertainty

a) Risk is the objective correlative of the subjective uncertainty; uncertainty considered as embodied in the course of events in the external world.

b) Risk is measurable uncertainty, i.e., uncertainty where the distribution of the outcome in a group of instances is known (either through calculation a priori or from statistics of past experience).

4) Risk = Uncertainty

a) in regard to cost, loss or damage.

b) about a loss.

c) of the happening of an unfavorable contingency.

d) of outcome, of actions and events.

5) Risk = Potential / possibility of a loss

a) Risk is the possibility of an unfortunate occurrence.

b) Risk is the possibility of an unfavorable deviation from expectations.

c) Risk is the potential for realization of unwanted, negative consequences of an event.

6) Risk = Probability and scenarios / consequences / severity of consequences

a) Risk is a combination of hazards measured by probability; a state of the world rather than a state of mind.

b) Risk is a measure of the probability and severity of adverse effects.

c) Risk is equal to the triplet (s_i, p_i, c_i), where s_i is the *i*th scenario, p_i is the probability of that scenario, and c_i is the consequence of the *i*th scenario, i=1,2, …N; i.e. risk captures: What can happen? How likely is that to happen? If it does happen, what are the consequences?

d) Risk is the combination of probability and extent of consequences.

7) Risk = Event or consequence

a) Risk is a situation or event where something of human value (including humans themselves) is at stake and where the outcome is uncertain.

b) Risk is an uncertain consequence of an event or an activity with respect to something that humans value.

8) Risk = Consequences/damage/severity of these + uncertainty

a) Risk = Uncertainty + Damage.

b) Risk is equal to the two-dimensional combination of events/ consequences (of an activity) and associated uncertainties.

c) Risk is uncertainty about and severity of the consequences (or outcomes) of an activity with respect to something that humans value.

d) Risk is the deviations from a reference level (ideal states, planned values, expected values, objectives) and associated uncertainties.

9) Risk is the effect of uncertainty on objectives (ISO).

While there is no single agreed definition of risk, the understanding of the risk concept has evolved over the last decades. The latest renewed interest of researchers in a definition of risk comes from the opinion that if the definition is shaky, the application is shaky (Aven *et al.*, 2014). The risk concept (Aven, 2011) should be distinguished from how we measure or describe that concept. Several initiatives were carried out in order to bring formality and unity to the terminology. The latest one, conducted by The Society for Risk Analysis (SRA), which brings together representatives from both academia and industry, suggested a new SRA glossary (Aven *et al*., 2015). The new glossary allows for different perspectives, distinguishing various concepts, for which overall qualitative definitions are provided, and the measurements of those concepts, for which examples of metrics are provided. There are different ways to measure, but they are all based on the same concept of risk having two features – uncertainty and consequence. The novel description of risk, first introduced by Aven, Baraldi, Flage and Zio (2014) is based on three features: uncertainty, consequence and knowledge. I see this research trend as yet further proof that there is a need to carefully address the quality of information and background knowledge when analyzing risks in engineering systems and their design.

It is normal in daily life to compare risks when choosing between alternatives. For example, one may have reason to assume that the risk of arriving late to an appointment is greater if you drive by car through a city center compared to taking the metro. In daily life, we can simply do comparison by gut feeling or rather simple contemplation. However, gut feelings would be a very shaky basis for responsible decisions that may result in large losses (Hubbard, 2009). Engineers make decisions by articulating with numbers. To be able to compare risks we need some measures. As there are basically two components of risk: human values that are at stake and uncertainties, they must be defined in a way that is susceptible to measuring (Hubbard, 2009).

For that reason, I will continue with understanding and defining the concept of uncertainty.

3.2. Definitions of Uncertainty

Based on Holton's (2004) review of common usage, ***uncertainty*** is a state of not knowing whether a proposition is true or false. Alternately, uncertainty is defined as the complement to certainty (Smithson, 1989) (Figure 2). That is, uncertainty is the lack of certainty. If given in this way, it is sensible to ask: what is certainty? Or what statements can we be certain about? Rene Descartes'[1] best known philosophical statement in this regard was: "*Cogito ergo sum*", which translates from Latin into English as "*I think, therefore I am.*" According to him, this is the only certain statement one can make. Any others can be doubted; thus they are uncertain. Other contemporary thinkers were also rather skeptical about being completely certain. For example, Jacob Bronowski's[2] often quoted thought is similar. He says that achieving "*knowledge is an unending adventure at the edge of uncertainty.*" On this very general note, uncertainty can be seen as the state of knowledge between complete ignorance and certainty (Smithson, 1989).

However, the division of knowledge into the three categories is too coarse. It can be nuanced. Some more recent developments are presented in Chapter 7 (Table 9). As Bertrand Russel[3] says: "*When one admits that nothing is certain one must, I think, also admit that some things are more nearly certain than others.*" Assume you are going to watch two tennis matches: one in which the rank of the contenders is very different: player A is ranked as number 5, while player B is ranked as number 85; the second match is played by players C and D about whom you do not know anything. You will perhaps be much more certain about the statement

[1] Rene Descartes is a French philosopher, 1596-1650
[2] Jacob Bronowski is a Polish-born British mathematician, historian, theatre author, poet and inventor, 1908-1974
[3] Bertrand Russel is a British philosopher, 1872-1970

“Player A will win over player B” than about the statement “Player C will win over player D” (Hubbard, 2009).

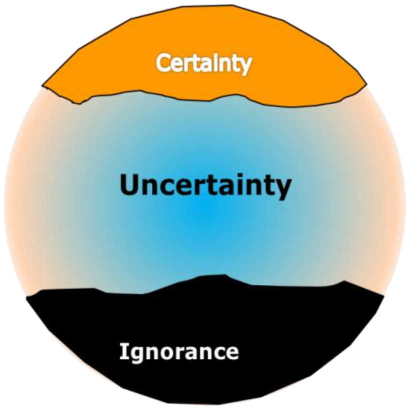


Figure 2 Uncertainty as a lack of certainty (following Smithson 1989).

In order to intelligently deal with uncertainty, we need to be able to present it and reason about it. There are primarily three ways of reasoning about it: ***formal logics***, ***fuzzy reasoning*** and ***statistical reasoning*** (Nisbett, 1993). By applying a formal logic to a possibly complex statement, represented as a set of formulas, one can deduct whether the statement is true or false. However, deducting to what extent one statement is more certain than another is not possible with this bi-valued logic (Ramsey, 2009). This type of formal logic-based reasoning is also referred to as the logic of Aristotle (Carruccio & Quigly, 2006). An alternative is to introduce the third value for a statement that is neither true nor false (Lejewski & Łukasiewicz, 1967). It is also possible to derive four-, five-, and even infinite-valued logic (Dunn & Epstein, 1977). However, the derivations, which are based on rather complex axiomatic systems, are not easily comprehensible and adaptable to observations, and they are exercised in the framework of formal and traditional epistemology (Smithson, 1989).

Lofty Zadeh suggested in his seminal work (Zadeh, 1978) an infinite-valued logic known as fuzzy set theory, and its extension as fuzzy logic or fuzzy reasoning. As the key concept, the theory proposed the membership function as a degree to which a statement is true and false (Lakoff, 1975). A completely new calculus was proposed, which is an alternative way of thinking and modeling complex systems using a higher level of abstraction originating from our knowledge and experience. Fuzzy logic resembles human reasoning in its use of imprecise information to make decisions. It allows expressing imprecise subjective knowledge such as *very old* and *a long time* and mapping it into exact numbers within the range [0, 1]. The theory captures both the uncertainties associated with human cognitive processes and uncertainties resulting from a lack of knowledge on the subject matter of interest (Taroun, 2014). The type of uncertainty that can be captured by fuzzy logic is also called ***ambiguity*** (Klinke & Renn, 2002).

The third way of reasoning about uncertainty is statistical reasoning. It involves various methods for representing uncertainty, assessing the measures of uncertainty, modifying the assessments to take account of new information, and combining them to calculate other quantitative measures and to draw conclusions (Dani & Joan, 2004). The most common representation of uncertainty uses probability, but it is by no means the only one (Swart *et al*., 2009; Dubois & Prade, 2012). The reason for having many other representations is the rather complex nature of uncertainty. In the risk community there is general agreement that there are at least two types of uncertainty that should be addressed in risk analyses: ***aleatory*** and ***epistemic*** *uncertainty* (Helton & Burmaster, 1996).

Aleatory uncertainty has a stochastic nature and cannot be reduced by acquiring relevant knowledge (Bernardo & Smith, 2009). For example, the wind speed in a given geographic point at a given point of time in the future is an uncertain value that has a pure stochastic nature. Regardless of the number of measurements, one cannot become more certain of the speed's value in a relatively remote future. This type of uncertainty is best represented by probability.

However, the other type of uncertainty – epistemic – arises due to a lack of knowledge and can be reduced by collecting data and information, and by acquiring new knowledge (Paté-Cornell, 1996). For example, the parameters of the probability distribution of time to failure of an electronic device produced in large numbers and exploited for a long time in very similar environments can be known very precisely. But exploiting the same device in different and more aggressive environments will introduce uncertainty over the parameters of the distribution that can only be removed by collecting and observing the time to failure in the new conditions. The uncertainty about the parameters of a probability distribution is epistemic.

In contrast to aleatory uncertainty, there are different views on how epistemic uncertainty should be represented, which has given rise to the development of different mathematical structures that can capture both aleatory and epistemic uncertainty (Nilsen & Aven, 2002). At present, there is no clarity about which of the existing theories should be employed when available evidence is provided in very different forms.

Many of the theories that can capture the two types of uncertainty are covered by the generic term 'Imprecise Probability'. They are mathematical models that measure chance or uncertainty without sharp numerical probabilities. These models include Dempster-Shafer belief functions, comparative probability orderings, convex sets of probability measures, interval-valued probabilities, possibility measures, plausibility measures, and upper and lower expectations or previsions. To capture aleatory uncertainty these models use different types of probability, while the probabilities are not presented as point-valued quantities in order to capture epistemic uncertainty.

At large, probabilities have two broad categories of interpretations: ***frequentist*** and ***subjective*** (Bernardo & Smith, 2009). The frequentist probabilities are associated with random physical phenomena like weather conditions, and systems such as roulette wheels and rolling dice. Frequentists posit the probability of an event as its relative frequency of occurrence after repeating the attempts to observe the event many times under similar conditions. This is how aleatory uncertainty can be characterized numerically. If a fair coin is repeatedly tossed many times, the empirical frequency of the two outcomes (head and tail) converges to the limit ½ as the number of trials tends to infinity.

If we denote n_A the number of occurrences of an event A in N trials, then the probability of this event is $P(A) = \lim_{n\to\infty} \frac{n_A}{N}$.

The frequentist view has its problems when we are concerned with events in the future that have never been observed, but are considered possible, or that have been observed but only rarely (possibly only once) or multiple times but under different conditions (Jaynes, 2003). These are in fact the situations a risk analyst faces. To be able to resolve the conflict of having only few (or no) observations of an event, the risk of which we want to measure, the subjective interpretation of probability is invoked (Jaynes, 2003). In this case probability is regarded as a measure of the degree of belief of the individual assessing the uncertainty. This probability captures both aleatory and epistemic uncertainty, and is often referred to as Bayesian probability (Jaynes, 2003). In principle it can be assigned to any statement, even when no random process is involved, for example to the likelihood that a suspect has committed a crime based on the existing evidence.

In process risk analysis both interpretations play a role, as the number of collected observations of a failure event can be representative to make an assessment of its objective probability as the expected frequency (Poirier, 2014). However, conditions for the failure may be unknown, so the "experiment" is no longer random and well-defined. Hence a resultant attributed probability can be 'corrected' by an expert, which makes it eventually subjective. Or the subject can adopt probabilities assessed as frequencies, which is a way of reconciling the frequentist and subjectivist view. Generally, risk analysis experts state that all risk calculations for slightly complex systems are subjective, because the results cannot be tested against experiments.

The use of Bayesian probability has caused both philosophical and practical debates on whether beliefs must follow the laws of probability, whether they should be expressed as a single number even though the knowledge support is very poor, or whether it is justifiable to use them in safety risk assessments (Beard, 2004). In the latter case, individual assessments by those who are not exposed to safety risks influence the safety of the people exposed to them.

Finally, subjective probabilities can be broadly classified into two different categories (interpretations): ***behavioral*** and ***evidential*** (Jaynes, 2003). The behavioral interpretation is given to probability if it is elicited by observing the choices of an individual, or it is provided by an individual who commits to acting accordingly when making choices. If we assume that the individual is rational, conclusions can be made on his subjective probabilities by offering him different options.

Evidential subjective probabilities reside on a different interpretation "*in which the probability* (…) *measures a logical* (…) *relation between the hypothesis and available evidence*" (Kyburg, 1987). Individuals provide evidence in some form that can be transformed into probabilities by employing sets of axioms or conventions. There is more than one way of doing this. Examples are the Dempster-Shafer theory of evidence and Walley's theory of coherent imprecise previsions (Walley, 1991).

Finer distinctions within the class of subjective probabilities are described by Walley (1991). Furthermore, exposure to negative consequences is the second essential component of risk. Negative consequences in engineering are as varied as individual industries and construction projects. Figure 3 brings together all the concepts and views introduced above on what risk is and how we can reason about it.

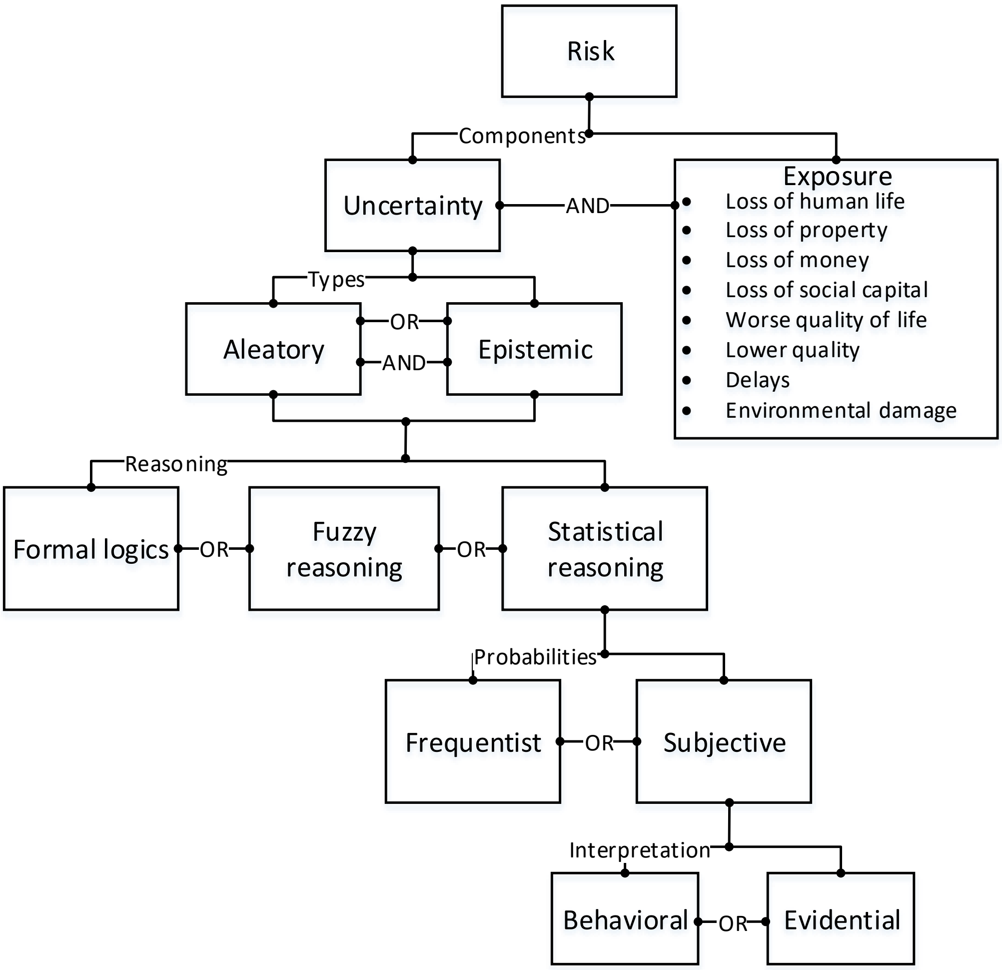


Figure 3 Overview of the concepts introduced so far: Risk components and their taxonomy (following Kozin, 2017).

3.3. Current state-of-the-art in risk management

During the last decades, management of risk in engineering design and associated projects and systems has attracted attention from researchers and practitioners in areas such as engineering design (Lough, Stone, &. Tumer, 2009), project management (Raz & Michael, 2001), or safety-related risk management (Paté-Cornell, 1996). The Project Management Institute (PMI) represents the largest professional organization dedicated to the project management field, and lists risk management as one of the ten main areas of project management (PMI, 2008). Furthermore, risk management courses are usually a part of most training programs for project managers. In accordance with the current view of project management as a life cycle process, project risk management is often perceived as a process that accompanies a project from the initiating through the planning, execution, monitoring and control phases all the way to the completion and closure (Raz & Michael, 2001).

Arguably, risk management has become an integral part of many formalized design processes for complex technical or socio-technical systems (Unger & Eppinger, 2011). The comparisons of risk management process steps under various design frameworks (Raz & Hillson, 2005; Oehmen *et al.*, 2014), including the generic ISO 31000, illustrate several points (Table 4): 1. All risk management process frameworks quantify risks, including qualitative ways of representing risks and uncertainty, as in some cases and for some practices that is only what is needed or feasible to achieve (e.g. high-medium-low evaluation). 2. Quantification of risks is directly linked to improved decision making, program stability and problem solving. 3. Risk and uncertainty quantification is only a part of risk analysis (which in turn is only a part of the overall risk management process). This implies that we not only need to improve the "quality" of numbers we generate during risk quantification, but also the way they are integrated into the overall risk management and associated decision making processes. In addition, the quality of input data, knowledge and information on which we base our assessments also has profound implications on the overall outcome.

Many of the issues that occur during the design of engineering systems are due to a lack of knowledge. It can be argued that epistemic uncertainty might sometimes be reduced by additional research and information gathering. However, that might lead to additional and hidden costs and delays, thus making it not feasible. This leads to "real-life" situations in design where actions that have significant impact on the subsequent processes and outcomes have to be taken on the basis of incomplete information. A major weakness of risk management today is that the methods used do not fully capture epistemic uncertainty (Aven & Zio, 2011). The previously mentioned various ISO standards and different professional and regulatory guidelines represent a significant progress in risk management practice. However, it is still open to debate how applicable, appropriate, and effective those guidelines are (Pender, 2001; Zwikael & Ahn, 2011).

Table 4 Comparison of risk management process steps under various design frameworks (adapted from Oehmen *et al.*, 2014)

ISO 31000 (ISO, 2009)	**PMI** (PMI, 2008)	**NASA** (Dezfuli et al., 2010)	**DoD** (DoD, 2006)	**INCOSE** (INCOSE, 2007)	**SEI** (Gallagher, 1999)
Communi-cation and consultation	*Implicit*	Commun-icate document	*Implicit*	Planning	Communicate
Establishing the context	Plan RM	*Implicit*			*Implicit*
Risk identification	Identify risks	Identify	Risk identification	Risk identification	Identify
Risk Analysis	Risk analysis	Analyze	Risk Analysis	Risk Assessment	Analyze
Risk Evaluation					
Risk treatment	Plan risk response	Plan	Risk mitigation planning	Risk Analysis	Plan
			Risk mitigation plan implementation	Risk Handling	
Monitoring and review	Monitor and control risks	Track	Risk tracking		Track Control

The question: "What is an acceptable way to quantify epistemic uncertainty?" is the underlying challenge motivating this study. Unlike for aleatory uncertainty, there is no general agreement on how to address epistemic uncertainty (e.g. Oberkampf *et al.*, 2004; Beer, Ferson, & Kreinovich, 2013). However, the scientific communities agree that the two types of uncertainty should be modeled differently, and yet, there is still a tendency in practice to employ one approach (the approach developed for aleatory uncertainty) for coping with all types of uncertainty. In other words, it is desirable to have a single method capable of quantifying all uncertainty (Goerlandt & Reniers, 2015), or at least a structured method to select appropriate methods of uncertainty quantification based on the specific uncertainty profile of the situation being analyzed.

Over the last decades probabilistic methods have been the predominant choice for risk assessments. A number of different approaches have been developed and applied in different contexts (for more details, please refer to section Table A.3 in Appendix 3 for a comparison of different methods). Practitioners' standards and guidelines have been developed according to such applications and their advancements.

However, a number of challenges have been identified in the established methods and in the attempt to use a single method to quantify all uncertainty (the corresponding findings from the literature accompanied by the empirical findings from this thesis are presented in Section 3.4), and new questions have emerged: Is probabilistic treatment appropriate for addressing epistemic uncertainty? Do we need more than probabilistic approaches to quantify uncertainty?

As a consequence of the identification of these issues, scholars from a probabilistic background have begun to acknowledge the limitations of their approaches and suggest developing extensions to the approaches that allow epistemic uncertainty to be better addressed (Hubbard, 2009).

There are also others that challenge that view and instead suggest that other mechanisms are needed. As they see probabilistic techniques as inadequate, they developed alternatives – non-probabilistic methods – that have certain merits and will be discussed later in the thesis. The response from the probabilistic community is that those approaches themselves are limited, as uncertainty can only be properly handled in a probabilistic view (Colyvan, 2008). They subsequently extended probabilistic techniques in an attempt to deal with the limiting issues that were identified (e.g. research on second-order uncertainty (Barrett & Lampard, 1955), etc.).

We are currently in the situation where sets of competing methods are proposed by the academic community – probabilistic and non-probabilistic methods – and without a clear understanding of how and when to use what method in order to cope with the challenging risk management role in design.

The thesis helps to address this gap through the following steps:

1. **The thesis provides a summary of the limitations, as identified in the literature, of the currently widely used (probabilistic) risk approaches. Furthermore, it introduces the reader to the ways probabilistic methods have been extended – and to the experiences with those approaches** (research question 1 and contribution 1)

The increased need to adequately cope with epistemic uncertainty also comes from the fact that large-scale engineering design solutions today often cover several systems and their interconnections, operating over a longer period. Such an evolving, iterative, social, and complex nature of engineering systems design corresponds to a multiplicity of plausible futures, several variants for system models, a range of outcomes, and associated weights or preferences regarding the various outcomes. In order to apply a certain method, we need to simplify a real situation to a “practical” model. More assumptions need to be made in order to have calculable data. This is especially the case when using probabilistic methods, as it has been proven that those methods face challenges when dealing with a higher level of uncertainty (Helton & Oberkampf, 2004; Walker, Lempert, & Kwakkel 2013). It is not justifiable to make significant assumptions when the overall level of ignorance is high. To provide more concrete insights and to document some of these modeling challenges, and inspired by the aphorism “*all models are wrong, but some are useful*”, the exploratory case studies described in Section 3.4 consider more details on when the current best-practice modeling becomes arbitrary for use and what the industry needs are.

In their study, Aven *et al.* (2014) showed that probabilities can always be assigned under the subjective probability approach, but that the origin and amount of information supporting the numbers is not reflected by the numbers produced. Their example clarifies that one may subjectively assess that two different events have probabilities equal to, say, 0.7, but in one case the assignment is supported by a substantial amount of relevant data, whereas in the other by effectively no data at all. This is the main argument in the critique of the probability-based approach to dealing with epistemic uncertainty. There is a particularly interesting case in situations when there is no information at all, in which case probabilistic approaches assign 0.5 probability by default (Bernardo & Smith, 2009).

Most other challenges are described in the corresponding chapters where the proposed alternative method is introduced. For instance, some of the challenges include the choice of a prior function in probabilistic modeling (Ferson, Ginzburg, & Akcakaya, 1996), subjectivity in expert opinions (Cooke, 1991), interpretation of results (Fortin & Gagnon, 2006), etc.

Within the probabilistic view of uncertainty, the research mostly considered further improvement of the developed approaches, further verifying and enabling higher precision in estimates (e.g. Kwiatkowska, Norman, & Parker, 2011). However, such focus on higher accuracy in modeling (providing more decimals or complete/detailed distributions) often leads to a false degree of precision (Ferson, 1996). Also, research suggests that predictions that are provided by action-outcome probabilities entail a certain degree of (first-order) uncertainty and that these probabilities themselves embody second-order uncertainty. Some advancements involve complicated sensitivity studies that are often cumbersome and can be difficult to interpret (Renn, Klinke, & Van Asselt, 2011). Conditional probabilities also attract special attention from researchers (Hogg, McKean, & Craig, 2005). However, even such extensions can be challenged (Sims, 2001).

Others significantly contributed to the field by introducing new paradigms/terms such as “known unknowns” and “unknown unknowns” (e.g. Pawson, Wong, & Owen, 2011). Tamed and wicked problems aimed to provide the context under which a certain modeling is more appropriate than others (Atie, 2008).

A range of different uncertainty and risk management methods has been applied to the mentioned problems. Group processes, such as the Delphi technique (Rowe & Wright, 1999), have helped large groups of experts combine their expertise into narratives of the future. This can be understood as an advanced method, where plausible future scenarios are developed without necessarily quantifying the associated uncertainties. In their work (Ferson & Ginzburg, 1996) illustrate examples in risk analysis for which classical Monte Carlo methods yield incorrect answers when used to quantify higher levels of uncertainty. IT development has brought statistical and computer simulation modeling that allows capturing quantitative information about the extrapolation of current trends and the implications of new driving forces. On the other hand, formal decision analysis can systematically assess the consequences of such information. Some more recently developed approaches, such as scenario planning, help individuals and groups accept the fundamental uncertainty surrounding the long-term future and consider a range of potential paths, including those that may be inconvenient or disturbing for organizational, ideological, or political reasons (Schoemaker, 1995).

2. **The thesis explains the opposing views and why they emerged** (research question 2 and contribution 2)

However, despite this rich legacy of approaches, one key aspect remains a problem. The commonly used methods that are briefly outlined above face challenges when dealing with long-term multiplicity of plausible futures, unknown causal structures, assigning probabilities and difficulty in identifying preferred solutions. In the following, the thesis briefly introduces why and which alternative approaches emerged for coping with such situations. More detailed descriptions and a broader overview are provided in Appendix 4.

The shift toward the application of analytical tools started in 1980s, mainly through the following two approaches. In his seminal work, Lotfi Zadeh (1978) described the concept of the Fuzzy Sets Theory introduced in the previous section. On the other hand, the Analytic Hierarchy Process (e.g. Dey, 2001; Singh & Singh, 2011) was recognized for its merits in relation to qualitative problems and factors that are often complicated and/or conflicting. Such a systematic approach allows decision makers to avoid addressing their problems intuitively, which suffer from inaccuracy and inconsistency.

An initial work on imprecise probabilities in the engineering design context was published by Chris Paredis and his colleagues (Aughenbaugh & Paredis, 2005). Even though their focus is rather on the technical aspects of design, their approach was not broadly accepted. However, after a decade we should revise this direction, as circumstances (computational support and maturity of the field) have changed. In contrast, some researchers focused on identifying situations for which more information is needed to reliably continue with the simulations (Goh, McMahon, & Booker, 2007).

Sandia National Laboratories (SNL, 2016) recognized the need to investigate research on epistemic uncertainty, and therefore initiated workshops in which approaches (some of which are described in this thesis) were presented. They produced reports (Sentz & Ferson, 2002), as well as guidelines (Ferson *et al.*, 2003) with a slightly more technical and modeling focus.

Other initiatives for example include workshops and publications on 'Decoding rings' – the attempt to clarify and unify the terminology on risk and uncertainty (e.g. Aven *et al.*, 2015); SRA focused on creating a special issue on foundational issues (Aven & Anthony, 2015) and questioning the definitions of risk (Aven, 2011; Kaplan & Garrick, 1981).

In addition to the previously described theoretical and methodological challenges of the most widely used methods, there are challenges that go beyond the quantification challenge itself. Different studies report the misuse and/or misrepresentation of probabilistic results (Love, Edwards, & Irani, 2012; Cantarelli, Flyvbjerg, Molin, & van Wee, 2010; Flyvbjerg, 2007; Flyvbjerg, Morris, Pinto, & Söderlund, 2013). Misuse can be due to two reasons: 1) inadequate or insufficient understanding/knowledge/skill of the practitioners when using a method, or 2) deliberate manipulation of results to support a desired outcome.

Misrepresentation can occur on two levels: 1) externally, for bidding purposes and winning different contracts, or 2) internally, sending a desired message to the top management.

Such findings represent one of the main reasons for clarifying concepts, discussing their adequate usage and informing on alternatives. Furthermore, these aspects have been confirmed through the empirical work presented in Section 3.4. One of the interviewees explained: "*I had two types of experiences: when risk owners would ask me to highlight their risk so they can use that as an argument to get more funding for their project, and when risk owners would ask me to remove a risk from reporting as their boss said anyone with a top-level risk will be fired.*"

This created the need to research and scientifically confirm methods that could help modelers overcome a broad spectrum of issues, both from the practical and academic point of view. To involve higher transparency in processes, methods application, background knowledge, results and their limitations, but also to clarify assumptions is central to achieving this goal. After being introduced to the non-probabilistic framework, one of the interviewees described it as: *"a more honest approach to the reality of challenges and complexities faced when quantifying risk and uncertainty*."

3. **The thesis highlights the merits from applying the non-probabilistic methods and presents empirical work on documenting challenges** (research question 3 and contribution 3)

Other than providing an understanding for the need to further research this topic, the thesis highlights the importance of risk and uncertainty quantification in the field. It further details the existing challenges and identifies the critical situations for which non -probabilistic methods offer better results (Section 9.2). Each group of methods is separately analyzed and findings are presented in corresponding chapters (the overall summary is presented in Chapter 10).

4. **The thesis provides advice for the communities on implementing non-probabilistic risk management techniques** (research question 4 and contribution 4)

After presenting the work carried out during this PhD project (in Chapters 4, 5, 6, 7 and 8), a discussion on its integration with the current state-of-the-art is presented in Chapter 9.

Recommendations for specific situations are provided, followed by a broader discussion on the characteristics of a good quantification practice (Section 9.2).

The next section reports on the empirical work regarding the challenges in current risk management practice.

3.4. Understanding and documenting challenges of existing methods from practitioners' point of view

This section presents the two exploratory case studies. The objective is to address research question 1. The research leads to contribution 1, i.e. current risk management faces challenges, and the most widely used risk quantification methods need improvement.

Exploratory case study 1: research design, data collection, coding and analyses

In order to investigate the industry needs, I conducted the first exploratory in-depth case study (Yin, 2013) with a case company involved in designing and deploying large-scale high-tech infrastructure in the energy sector. The purpose is to describe the current state and to address the initial phase of theory building. For this reason, theoretical sampling is appropriate (Eisenhardt & Graebner, 2007) as the case company is selected and the case is developed for illuminating and extending the relationships and logic among the constructs. As explained in Chapter 2, confidentiality concerns limit to what extent I can elaborate on the empirical work presented in this chapter. Following Eisenhardt (1989), I build on the argument that replication logic is key to theory building from case studies. Furthermore, the exploratory approach is applicable as this is the early stage of the theory-building cycle (Cash, 2018). The case study data were collected using a multimethod approach consisting of quantitative (probabilistic) modeling and interviews over the course of nine months.

We jointly developed and analyzed the implementation of a complex, quantitative engineering design and deployment project risk model in Primavera. A junior risk analyst working daily with the tool also helped in this process. The case was designed to illustrate their modeling process and availability and the quality of data needed for such types of analyses. The practitioners monitored the data used to ensure that the synthetic input to our model corresponded to real-world scenarios. Possibilities for different analyses were investigated and discussed, such as tornado diagrams, risk correlations, cluster risks, and cascading effects.

Interviews were organized to allow for potential triangulation, enrich the overall setup of the performed modeling, support the modeling data, and document additional challenges with current approaches (both related to this type of modeling but also more general in risk management). I conducted nine interviews with their senior project risk manager, each lasting between one to two hours. Three sessions were excluded from the analysis as they served more for 'building relationship' purposes than establishing content. Additional meetings were organized when needed to clarify modeling. It was agreed there will be no recording, and I took notes during the meetings. Two senior researchers also attended the first four meetings, ensuring that the design and progress of the study were carried out carefully, stepping in with additional 'What if' questions as appropriate. The follow-up calls were arranged when more clarification was needed. The semi-structured interviews (Kvale, 1996) were designed to investigate the current risk management practice by asking the participants in three ways. First, I asked about what he perceived to be instances of good and bad risk management, encouraging him to go into as much detail as possible with the experiences. Second, we discussed modeling challenges in current best practices. Third, we discussed the potential of the non-probabilistic approaches in current practices.

The interviews were iteratively coded and analyzed in ATLAS.ti based on a grounded theory approach (Corbin & Strauss, 1990). The process of coding was to use open codes and line by line coding as well as second-stage-codes that were then grouped. For an example see the figure in Appendix 1. The preliminary conclusions and emerging coding constructs were discussed and validated with the company in informal and formal meetings throughout the process. In addition, I requested an independent senior researcher from the field to read through the data and codes in the process (to reduce the chance of bias and hear about possibilities to extend the list of questions/topics).

It is worth noting that the interviewed manager has more than 20 years of experience. He was able to provide many insights from different industries he had the opportunity to work in (telecommunication, construction, energy) and highlight areas that would not necessarily emerge from the initial literature review. For instance, the importance and impact of assumptions made when choosing distributions, gaming and political aspects of decision making affecting the analysis, correlations among risks, computational capacity and a "lucky manager" problem.

Key insights from Exploratory Case Study 1

Quality of probabilistic models: The key challenges occurred around the issue of model size and complexity. The engineering activities at hand generated a large number of activities and resources that needed to be modeled, including their dependencies, each of which was analyzed in terms of schedule risk. A large number of probabilities and probability distributions is required to run risk assessment simulations, e.g. regarding the duration of each task. Quality issues arose as to how representative the model actually is of the underlying project. It was difficult to justify simplifications that were made during the modeling process, particularly regarding the impact on the outcome of the risk assessment.

Quality of data and results: The data used to generate probabilities and probability distributions are perceived to play a critical role in the outcome of the risk assessments. While some probability distributions were developed based on similar past projects, others relied on expert opinion and group consensus, based on various elicitation techniques. However, their representations in the system are identical, and do not reflect the quality or reliability of the input data. They are also required to be put in as fixed probabilities and / or probability distributions. In addition, various mathematical and computational tools are used during the simulations, without always fully appreciating their prerequisites or limitations.

Use and integration of results: Most analyses rely on advanced mathematical concepts employed during the simulation and computation of the risk assessment. Their meaning and implications cannot be fully appreciated without a deep understanding of the tools and methods used. The same applies to the origin and quality of the data, which can often no longer be judged from the presentation of the results. Finally, existing tools do not explicitly address the "gaming" aspects of tailoring risk analysis approaches to produce the desired results, or interpret results one-dimensionally to suit a particular preconceived notion of a desirable outcome.

An interesting part of the interviews relates to the 'lucky manager' problem. The problem relates to the possibility of perceiving a manager (or management) lucky in terms of a project/portfolio/system performance rather than competent in his/her (their) capacity to handle a broad range of activities, including risk management. This was to an extent studied by Geraldi, Lee-Kelley and Kutsch (2010). The interviewee raised the question on how to distinguish between managers who conducted thorough risk management, but faced major or unlikely risks materializing, and managers that undertook no risk management at all, but did not suffer from any larger risks occurring during their projects. After two iterations, it was decided that the topic is outside of the thesis' scope, and it was not coded separately, nor was further research carried out. However, I conducted an initial literature review on the topic and asked the independent senior researcher if he had any experience with it.

The senior researcher commented: "There is no such a thing as a lucky manager. One can fail in managing large projects in so many ways that an incompetent, but lucky person could potentially be lucky in one situation, but for sure not in others. Therefore, if they executed a large-scale project efficiently, they knew what they were doing."

Exploratory case study 2: research design, data collection, coding and analyses

To investigate the industry needs in a completely different area of design work and with a completely different risk management process, I conducted the second in-depth exploratory case study. This study analyzed a design phase of a large construction project in the city center of Reykjavik worth almost a billion Icelandic kronor and with project delivery scheduled for late 2019. The project includes a hotel, apartments, bars, restaurants and a music hall and is managed by an engineering and consulting SME that provides design services for construction projects. In general, the study confirms the findings described above.

The study was designed on the same pillars as the previous one – an exploratory case study for which the data were collected as a multimethod study consisting of quantitative (probabilistic) modeling and interviews. While the research design was the same, the difference is that this case study took fewer iterations than the first one, as it came later in the theory-building process. The interviews were recorded and transcribed, each lasting from one to one-and-a-half hours. The study was conducted over the period of six months. More on the specifics of the interviewees is available in Section 8.4. As the company does not have an established risk management, individual interviews were conducted with an engineer with risk management training, a fire and safety engineer, the project manager, a structural engineer, the HVAC design manager, an electrical engineer-designer, an architect and design manager, and the project owner. Semi-structured interviews (Yin, 2013) were designed to cover three topics: the interviewees' background and experience and relation to managing risk; their role in the project and potential issues (risks) they considered; the main challenges they experienced in this project and in their practice; their view on the proposed modeling (RamRisk); quality of data in their work. ‘What if’ questions were used as an additional way to draw out the risk management challenges for the participants. The wrap-up question encouraged reflections on clashes experienced in relation to risk management in projects in which they were involved, but also on what would be a way to improve their practice.

The second exploratory case study investigated RamRisk in order to document the potential, but also the limitations, of one of the currently most widely employed risk tools in Nordic risk management practice. Given that the company had no established risk management procedures, by using the tool RamRisk it was feasible to demonstrate the benefit of an early involvement of such risk analysis, as well as its practicability for the project outcome. An initial risk register was taken from available online sources in terms of project cost and schedule. The register was then updated and discussed with practitioners who helped assign probabilities for the specific risks. Possibilities for different analyses were investigated and discussed, such as tornado diagrams, FN curves, and risk matrix.

The tool allows assigning responsibilities of identified risks to different users, which was one of the main needs of the project contractor. However, the insights regarding the limitations of the tool are aligned with the previous study: **1) there are challenges in modeling and assigning probabilities, 2) the quality and availability of data is a main constraint, as there is no developed culture toward documenting and articulating risks, 3) use and integration of results is seen as rather challenging,** as they first need to establish a culture that values and understands the need for risk management, but also that the employees have the adequate educational level (possibly achieved through courses and seminars on risk). Moreover, a number of behavioral aspects were mentioned, such as lack of interest from the managerial side to implement formal procedures, lack of response during data gathering, and cross-sectoral learning/knowledge sharing.

Additional Interviews: data collection, coding and analysis

The findings from the two exploratory case studies were supplemented by interviews in three more companies. These interviews, on the other hand, provided a number of additional challenges in current risk management practice that are generally aligned with the literature findings described in Section 3.3. I conduced semi-structured interviews according to the topics discussed (current tools and methods, quality of data used for analyses, limitations of the current methods, communication of the results to decision makers, risk-informed decision making). The types of questions were developed around open, hypothetical, or comparative lines of questioning (Kvale, 2008). The findings from the interviews are summarized as follows (please see Appendices 1 and 2):

Current tools and methods: The choice of the case companies was conducted as to cover a broad range of design activities and solutions. Accordingly, there was also a large variety of methods employed in their practices. Yet, none of the methods or processes used specifically focus on epistemic uncertainty.

Quality of data used for analyses: Only the interviewees in two highly specialized companies in terms of risk quantification expressed the importance of the quality and availability of data in the whole process. The others did not consider if and how the quality of data impacts their risk management process and related decision making. Two companies analyzed if they could store data from various projects and implement knowledge sharing across departments.

Limitations of the current methods: The two companies where risk quantification practices are well established reported the inability of their methods (such as the ones used in the exploratory in-depth case studies) to represent the quality of data on which they performed the analyses. They find it crucial for this to be communicated to decision makers together with the results. One company that was only establishing a risk management process explained that the issue with implementing formal, quantitative methods lies in the educational level of its employees.

Communication of the results to decision makers: In terms of more advanced quantification, communication was seen as challenging as decision makers cannot necessarily comprehend all available results. In terms of less quantitatively oriented practices, communication was also seen as challenging, as there was not enough awareness and appreciation for discussions regarding risk. Moreover, time pressure was a commonly reported issue.

Risk-informed decision making: What was also commonly reported is that regardless of the analyses or the results of a risk assessment, a number of things influence the final decision. Managers' personal ambitions and gaming aspects are recognized as important elements influencing the final decision.

Reflections

Some of the limitations resulting from the choice of research methods include:

Case studies typically combine data collection methods such as archives, interviews, questionnaires, and observations to enrich the findings. However, as their validity can be challenged, ways how rigor can be improved are elaborated in Section 9.3.

The coding scheme in interviews presents several potential problems including non-reproducibility and subject selection bias due to the author's field of expertise. The sample selection within the second study-company is potentially tainted, as the interviewee choice is biased and directed by the initial interviews.

One limitation comes from the fact that no industry partner in the project was able to provide access to project documentation. Significant time and efforts were spent on building connections, relationships and trust. Different possibilities for collaboration were discussed, but a number of these attempts did not succeed. Study material is potentially not well triangulated because of limitations on access to risk documentation. Please refer to Section 9.3 for more details. Furthermore, interviews with practitioners from one company had to be excluded due to the complications with signing the NDA. Furthermore, a three-hour interview with the head of the risk management team in a large construction company was also excluded as it was established rather late in the process. However, the collected data were aligned with the ones described above. This can also be seen as an informal check of the constructs and biases developed through the coding.

3.5. Discussion and summary

Considering the importance of design in engineering systems, methods to deal with risk and uncertainty are essential. By introducing and reviewing the main streams and concepts in literature, and by conducting empirical work, this chapter identified that there is still space for improvement both from the risk and uncertainty quantification and risk management process point of view. Current methods are not coping with all challenges that appear during an engineering systems lifecycle. The following chapters aim to address these gaps (Chapters 4, 5, 6 and 7 regarding quantification, and Chapter 8 regarding the whole risk management process).

Starting with the probabilistic treatments of uncertainties and by acknowledging its large merit, limitations and challenges are also provided that lead to the need for frameworks beyond probability. This highlights the need for a search for alternatives, possible improvements of risk and uncertainty representations, and a summary of the main paradigms that have to some extent been researched by different communities. As explained, current risk management practices rely on available risk management process frameworks that are based on the probabilistic view of risk. Methods that can effectively and reliably deal with uncertainty due to a lack of knowledge are still missing. A handful of alternative approaches is available, but their implementation seems fraught with difficulties.

Acknowledging risk and uncertainty assessments as decision-support tools requires that the meaning and practical interpretation of the computed quantities are presented and communicated to the decision makers in an understandable format (Aven *et al.*, 2014). There are three critical questions from a decision maker's perspective:

1. For a specific situation, which is characterized by a lack of knowledge, what options do
I have?
2. How reliable is the first answer I get, and can I use it confidently?
3. How cost-effective is a particular analysis method?

The thesis argues that non-probabilistic methods allow us to better address these three questions. It proposes to use non-probabilistic methods to be transparent when there is a lack of knowledge and to address identified issues in a more structured manner, both qualitatively and quantitatively, instead of simply ignoring the degree and quality of available knowledge. By including additional judgments, we are taking into account available information and yet clearly articulate which parts are not known. The use of non-probabilistic methods can contribute to current engineering systems design practice, with the goal to faithfully represent and express the knowledge available to best inform a decision-maker and support the decision making process.

As Polanyi, a research philosopher, said: "*We can know more than we can tell*." This is often how experienced managers' or experts' way of working is explained. The current practice needs to take a step forward from relying on “manager’s experience”, which can be seen as a simple “way out” to dealing with epistemic uncertainty rather than its management.

4. Conceptualization and collection of non-probabilistic methods

"We know more than we can tell"

– Michael Polanyi –

This chapter introduces a collection of advanced methods to risk and uncertainty quantification that promise to better cope with the challenges in current practice. These methods are introduced under the 'non-probabilistic' framework and are structured into three groups. The first group of methods is based on imprecise probabilities, the second represents a group of semi-quantitative approaches, and the third group of methods is based on exploratory modeling.

As introduced in Chapter 1, uncertainty and risk represent one of the key challenges in design-related decision making. "Newness" and lack of knowledge are characteristics of design, and yet, typical risk management methods in design rely on probability-based risk quantification methods that are heavily dependent on previously collected data (Ferson, Ginzburg, & Akcakaya, 1996). In Chapter 3, I discuss to what degree current risk management approaches are appropriate for real-world design challenges. The chapter argues that current approaches primarily focus on aleatory uncertainty (i.e. uncertainty due to the inherent randomness of the physical world) and that other methods are needed to address epistemic uncertainty (i.e. uncertainty due to lack of knowledge) and ambiguity (i.e. differing interpretation of identical factual information on uncertainty). The non-probabilistic framework is then presented and the methods are described. I illustrate the methods with application examples in other fields and discuss their relationship to the key challenges in decision making processes of designing engineering systems. The chapter concludes with a discussion of their application potential in design, as a basis for the following chapters.

In the context of this thesis, this is an important step: It provides the conceptual development of advanced risk and uncertainty quantification methods for design needs and establishes the non-probabilistic framework. This chapter describes the methods, and provides a unique and clear structure (answering research question 2 by building upon research question 1).

The chapter is structured as follows: Section 4.1 provides a unique set of advanced risk and uncertainty methods, structured in three groups. Section 4.2 describes the methods. In Section 4.3 their design context is analyzed. In Section 4.4 I acknowledge limitations and criticism related to non-probabilistic approaches, and in Section 4.5 a summary of the chapter is provided together with a link to the following chapters.

4.1. Conceptualizing the non-probabilistic framework in the context of engineering systems design

It is important to make a distinction between uncertainties that can be treated through probabilities and uncertainties that cannot. The thesis acknowledges the large merit of probability-based methods when it comes to uncertainties of stochastic nature, but also points out limitations that lead to the need for frameworks beyond probability when it comes to uncertainties due to lack of knowledge.

Non-probabilistic methods collected across different domains are here systematically presented in three groups as the non-probabilistic framework. The framework supports "beyond probabilistic" reasoning by using the non-probabilistic methods and also aligning them to the overall design and risk management needs (more details in Chapter 8). From each group of methods I briefly describe those that have the potential to better address the industrial risk management challenges discussed in Section 3.4. I further provide an overview of the fields in which these methods have been broadly discussed and used.

4.1.1. Imprecise probability

Imprecise probability (Walley, 1991) expands the possibilities of established probabilistic risk quantification to reason more reliably with limited information on actual probability distributions. The approach allows decision makers to review and discuss coherent and plausible ranges of probabilities. Given that probabilities cannot be known precisely if the modeler only has partial information at hand, imprecise probability suggests constructing probabilistic measures of interest as precisely (or imprecisely) as available data allow, in the form of intervals.

a) Coherent upper and lower probability

In coherent upper and lower probability, the major novelty is the idea to drop a central assumption of Bayesian theory, which states that uncertainty should always be measured by a single (additive) probability measure. There is a large number of arguments that support the concept of coherent upper and lower probability and why it is needed (Kozin & Petersen, 1996). Given that it does not require unjustified assumptions, which is the case with traditional approaches as argued in Section 3.4, the use of this method nicely builds on Colyvan's (2008) argumentation.

b) The Dempster-Shafer theory of evidence

The Dempster-Shafer theory of evidence originates from the work of Dempster (1967) in the context of statistical inference. It was later formalized by Shafer as the theory of evidence. In their study, Beynon, Curry, and Morgan (2000) pointed out that the Dempster-Shafer theory of evidence, as a technique for modeling reasoning under uncertain, imprecise and incomplete information, seems to have numerous advantages over the more traditional statistical methods. The main feature of the Dempster-Shafer theory of evidence is the possibility to include additional judgments in evidential reasoning. This permits the theory to measure and take into account the weight of evidence, which arguably also addresses the argument about ambiguity from the previous chapter.

4.1.2. Semi-quantitative methods

Semi-quantitative methods represent quantitative methods that are combined with additional qualitative information. From the various semi-quantitative representations that are developed in different fields (see for example Flage & Terje, 2009; Berner & Flage, 2015; Aven, 2008), the NUSAP scheme is presented here (Brocéliande team, 2015).

c) The NUSAP scheme

The NUSAP scheme (Funtowicz & Ravetz, 1990) can again be seen as an extension of established probabilistic modeling of uncertainty. It adds qualitative information to the uncertainty and risk analysis in a structured manner, informing the modeling, analysis and decision-making process by making issues such as data origin, quality and key assumptions transparent. The acronym “NUSAP” stands for Number, Unit, Spread, Assessment, and Pedigree – the five elements that constitute an information set regarding uncertainty in the method. Connected to the partial information available argument from Section 4.2, it is important to note that the NUSAP scheme makes the background knowledge, as well as assumptions, transparent. That allows clear and easier communication with parties involved in decision-making processes.

4.1.3. A family of related conceptual approaches based on Exploratory Modeling

A family of related conceptual approaches is based on exploratory modeling, which uses computational experiments to run simulations. It represents the third group of non-probabilistic methods. The underlying idea is that instead of determining the best predictive model and solving for the risk mitigation procedure that is optimal (but fragilely dependent on assumptions), it is wiser to seek among the most robust actions when dealing with uncertainty due to lack of knowledge. That is, those actions that at least lead to a satisfactory result under a large number of possible future development scenarios. Considering the argument about limitations of a rational decision maker from Section 4.1, these sets of methods represent a completely new way of thinking: instead of the traditional "predict and act" paradigm, they bring a "monitor and adapt" one.

A family of conceptually related methods for dealing with uncertainty:

- **Assumption-Based Planning** was developed at the RAND Corporation almost 30 years ago as a tool for improving the adaptability and robustness of an existing policy/plan/design (Dewar *et al.*, 1993)
- **Robust Decision Making (RDM)** uses multiple views of the future to iteratively stress test one or more candidate strategies over many scenarios and refine the strategies in light of this (Walker, Haasnoot, & Kwakkel, 2013)
- **Adaptive Policymaking** was specifically developed to support the implementation of long-term plans despite the presence of uncertainties (Haasnoot *et al.*, 2012)
- **Adaptation Tipping Points and Adaptation Pathways** both approach the timing of actions and were developed for water management (Haasnoot *et al.*, 2012)
- **Dynamic Adaptive Policy Pathways** combines the work on Adaptive Policymaking with the work on Adaptation Tipping Points and Adaptation Pathways (Haasnoot, Kwakkel, & Walker, 2013).

In this thesis, I introduce Robust Decision Making, because it is the most developed approach.

d) Robust Decision Making

Robust Decision Making (RDM) has been developed over the last 30 years, primarily by researchers associated with the RAND Corporation (Dewar *et al.*, 1993). The RDM framework uses multiple views of the future to support a thorough investigation of modeling results that helps to identify a policy/plan/design (Lempert, Popper, & Bankes, 2003; Groves & Lempert, 2007), that: (1) is robust; (2) avoids most situations in which the policy/plan/design/system would fail to meet its goals; and (3) makes clear the remaining vulnerabilities. **As Chapter 7 explores RDM in depth, a more detailed description and discussion is omitted from this chapter.**

Since its development, RDM has been applied to strategic planning problems in a variety of fields, including climate change (Lempert, Schlesinger, & Bankes, 1996), complex systems (Lempert, 2002), economic policy (Seong, Popper, & Zheng, 2005), and flood and water risk management (Herman *et al.*, 2014).

4.2. Description of the methods

4.2.1. Imprecise Probability

During the last three decades, a number of mathematical structures have been developed that relax the strong axioms of probability theory (Kolmogorov's axioms) and thus allow capturing epistemic in addition to aleatory uncertainty. This group of theories is referred to as the "theories of imprecise probabilities." Imprecise probability is a generic term for a range of mathematical models that measure chance or uncertainty without sharp numerical probabilities (e.g. "can be", "for example", interval-valued). These models include belief functions, Choquet capacities, comparative probability orderings, convex sets of probability measures, fuzzy measures, interval-valued probabilities, possibility measures, plausibility measures, and upper and lower expectations or previsions (Walley, 1991). Imprecise probability admits that probabilities cannot be known precisely if the modeler only has partial information at hand.

The major novelty in the concept is to drop a central assumption of Bayesian theory, which states that uncertainty should always be measured by a single (additive) probability measure. Unlike the Bayesian "dogma of precision", in order to characterize the uncertainty of an event with imprecise probabilities, we need both lower and upper probabilities.

There are a large number of arguments that support the concept of imprecise probabilities. The following list is taken from Kozin and Petersen (1996) and illustrates from the practical point of view why imprecision in probabilities is needed:

- to reflect the amount of information on which they are based;
- to model a state of complete ignorance, meaning a total absence of relevant information;
- to combine several sources of information;
- to combine different probabilistic judgments generating an imprecise model;
- to treat disagreement among group members over probabilities obtained by judgments in the same way as conflict between several assessments of one individual: both are sources of imprecision;
- to capture uncertainties of some problem situation more faithfully, not only due to randomness.

Football example **(Walley, 1996)**

Consider a football game whose possible outcomes are win (W), draw (D) or loss (L) for the home team. To express its uncertainty about the outcome, the user makes the judgments:

Probably not W,
W is more probable than D,
D is more probable than L.

What can we say about the probabilities of the three outcomes?

The theory of coherent imprecise probabilities allows computing interval-valued probabilities based on the above partial and imprecise statistical information that is closer to the natural language, although tied to probability. The answer to the question is: P(W) = [1/3; 1/2], P(D) = [1/4; 1/2], P(L) = [0; 1/3].

If more non-conflicting judgments are provided, the bounds for the probabilities become tighter. Many other kinds of qualitative or quantitative judgments could be added to the three we have considered, for example,

if not D then W is very likely,
W is between 1 and 2 times as probable as D,
I am willing to bet on L at odds of 4 to 1,
W has precise probability 0.4.

The theory of coherent imprecise probabilities can also accommodate different reliabilities of different sources of information, if there are grounds to assume that one source of information is more reliable than another.

4.2.2. The Dempster-Shafer Theory of Evidence

Other imprecise probability theories allow deriving interval-valued probabilities given a different type of input. One of those theories is the theory of belief functions, the Dempster-Shafer theory of evidence. In their study, Beynon, Curry, and Morgan (2000) emphasize that the theory was popularized in the literature of artificial intelligence and Expert Systems, but it has also been applied to certain extents in the fields of face recognition, statistical classification, target identification and medical diagnosis.

The main feature of the Dempster-Shafer theory is the possibility to include additional judgments in evidential reasoning. This permits the theory to measure and take into account the weight of evidence. Another key feature highlighted by Beynon *et al*. (2000) is that, unlike in possibility theory and statistical reasoning, there is no need to force our probability or belief measures to sum a unity. Hence, possibility theory can be considered a special case of Dempster-Shafer's theory.

The Dempster-Shafer theory of evidence is based on complex mathematical explanations, a discussion on which goes beyond the scope of this thesis. One study by Walley (1996), where the Dempster-Shafer theory of evidence is mathematically exhaustively explained, is followed with a set of six examples, each mathematically grounded. The authors of this thesis tried to find an example where an extent knowledge of mathematics is not necessary to follow the argumentation, but, having failed to do so, focus on one example of a key feature that is mentioned above.

Example: reliability analysis (quoted from Aven, 2014):

"To illustrate, suppose that a diagnostic model is available to indicate with reliability (i.e. the probability of providing the correct result) of 0.9 when a given system has failed. Considering a case in which the model does indeed indicate that the system has failed, this fact justifies a 0.9 degree of belief in such an event but only a 0 degree of belief (not 0.1) in the event that the system has not failed. This latter belief does not mean that it is certain that the system has failed, as a zero probability would; it merely means that the model indication provides no evidence to support the fact that the system has not failed. The pair of values {0.9; 0} constitutes a belief function on the propositions 'the system has failed' and 'the system has not failed.'"

4.2.3. NUSAP (Number, Units, Spread, Assessment, and Pedigree) tool

In contrast to the previously presented theories, where expert knowledge is required to interpret the results, a different technique was developed during the 1980s. The idea is to draw attention to the properties of numbers (which are often ignored) and to offer transparency when it comes to the quality of information. The NUSAP scheme targets a broader audience and the origin of the data plays a bigger role. Funtowicz and Ravetz (1990), alarmed by the misuse of numbers in debates about nuclear safety levels, and later by the misuse of scientific findings by climate change "skeptics" to delay climate action, constructed the NUSAP notation. With the focus on policy-related research, they proposed that nowadays tasks should not only include the management of uncertainties, but also the assessment of quality and communication with the public.

This thesis argues that high-quality decision making not necessarily requires the elimination of uncertainty, but rather its effective management, as offered by the NUSAP scheme. The NUSAP measure can capture more background features than imprecise probabilities, however, at the "cost" of being a qualitative measure. Engineering systems design risk management approaches must be based on coping with a lack of knowledge at least as much as on the application of knowledge (Funtowicz & Ravetz, 1990). The NUSAP measure has a large information content, but by being a qualitative expression, there is no strict formal way to base decision making on it.

Funtowicz and Ravetz (1990) coined the term NUSAP as an acronym for the five categories of information included in their measure: Number, Units, Spread, Assessment, and Pedigree. The essential idea is that a result of any analysis, including risk and uncertainty quantification, should not be a single number, but should be accompanied by additional information to allow decision makers to interpret its overall meaning value (here introduced through the four additional categories). The "unit" measure states whether we are talking about percentage, money, or something else. "Spread" and "Assessment" are related to uncertainty. Spread is used to express the random error, and the systematic error is expressed by Assessment. The most significant novelty comes from the "Pedigree" measure, which informs on the information feed, or in other words, the origin and quality of data analyzed. By providing detailed information to the decision maker on how data were collected, what the sample size and similar measures are, the NUSAP measure lets them judge the overall value and meaning of the presented data. It eliminates uncertainty or misinterpretation on whether for example a probability measure is just a guess or based on extensive simulation and testing.

There are guidelines for NUSAP application (Brocéliande team, 2015) and the following list is quoted according to the same source.

Typical strengths of NUSAP are:

- NUSAP identifies the different types of uncertainty in quantitative information and enables them to be displayed in a standardized and self-explanatory way. Providers and users of quantitative information then have a clear and transparent assessment of its uncertainties.

- NUSAP fosters an enhanced appreciation of the issue of quality in information. It thereby enables a more effective criticism of quantitative information by providers, clients, and, generally, users of all sorts, expert and laypersons.

- NUSAP provides a useful mean to focus research efforts on the potentially most problematic parameters by identifying those parameters, which are critical for the quality of the information.

- The diagnostic diagram, a NUSAP method, provides a convenient way in which to view each of the key parameters in terms of two crucial attributes. One is their relative contribution to the sensitivity of the output, and the other is their strength. When viewed in combination on the diagram, they provide indications of which parameters are the most critical for the quality of the result.

4.3. Applications of non-probabilistic risk and uncertainty quantification methods

4.3.1. Current applications of non-probabilistic methods

Non-probabilistic methods have so far been applied in several areas. To my knowledge, the methods have mostly been used outside the design, product development, and project management domain even though they were developed some time ago.

One well-recognized application of imprecise probabilities is in the domain of artificial intelligence. In a seminal study, Walley (1996) compares four measures that have been advocated as models for uncertainty in expert systems. The measures are additive probabilities (used in the Bayesian theory), coherent lower (or upper) previsions, belief functions (used in the Dempster-Shafer theory), and possibility measures (fuzzy logic). His work is considered as a reliable scientific background for deeper understanding, mathematical explanations and representative examples of the mentioned theories. Findings in his study demonstrate that each of the four measures is useful for particular kinds of problems. However, only lower and upper previsions, here introduced as coherent upper and lower probability, perform in a sufficiently general way to model the most common types of uncertainty.

The methods have also been introduced and applied to the following fields: 1) Civil/structural engineering (Zio, 2009; Berner & Flage, 2015); 2) Risk, resilience and vulnerability of critical infrastructures (Zio, 2007); 3) Environmental risk assessment (Guyonnet *et al.*, 2003); 4) Offshore oil and gas installations (Lavasani *et al.*, 2011); 5) Risk assessment of radioactive waste repositories (Helton, 1993).

Previous applications of coherent upper and lower probabilities: The potentials of coherent upper and lower probabilities have been analyzed in the field of reliability and safety assessments (Kozin & Petersen, 1996; Ferson & Ginzburg, 1996). Due to the important impact of safety and reliability analyses on human, environmental and economic conditions, it is essential they comprise the maximum amount of useful information. The main question in those analyses is whether a potentially dangerous technical object meets the regulation values or not. In cases of small samples of operational data, precise probabilities represent an arbitrary solution. On the other hand, the use of coherent upper and lower probabilities does not offer a simple *Yes* or *No* answer, but accompanies both answer options, meaning that they offer information on what is more probable. That is usually sufficient for judging which hypothetical events are most likely to happen. Only when the regulation value lies outside the upper and lower interval, it is possible to determine the precise *Yes* or *No* answer.

Previous applications of the Dempster-Shafer theory of evidence: The Dempster-Shafer theory has also been applied to a certain extent in the fields of facial recognition (Ip & Ng, 1994), statistical classification (Denoeux, 1995), target identification (Buede & Girardi, 1997), medical diagnosis (Yen, 1989), risk assessment and applied biomathematics (Ferson *et al.*, 2003) and climate change (Ben Abdallah, Mouhous-Voyneau, & Denœux, 2013). A full overview of the research directions is available in (Denoeux, 2016). Significant progress was made in signal processing by implementing imprecise methods thinking for reliability analysis (Kozin & Petersen, 1996).

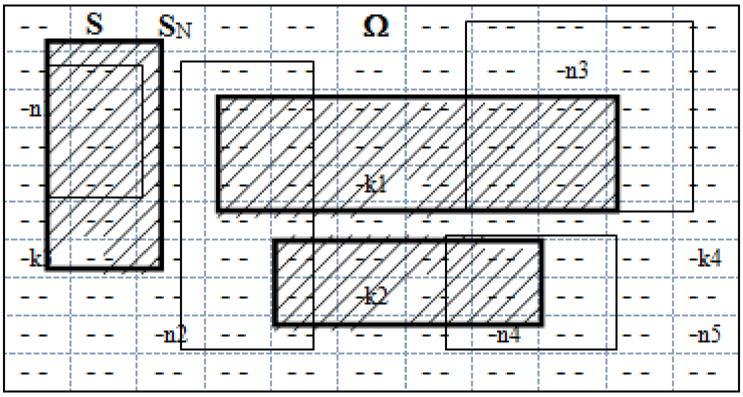


Figure 4 Graphic presentation of an electronic device consisting of components "-" and Integrated Circuits (rectangles) (Kozin & Petersen, 1996).

In this example (Figure 4), the set of all components **Ω** contains 220 components “-”. For the system S, e.g. an electronic device, consists of four integrated circuits n1, n2, n3, n4 and the rest n5. By knowing the reliabilities of subsystems in case of a system failure, we obtain probabilities of finding a failed component in each of the subsystems. By using the Dempster-Shafer theory of evidence it is possible to calculate same probabilities for other subsystems, such as k1, k2, k3. The analogy of this example with clusters-of-risks issue will be elaborated in Section 4.4.3.

To date, the author is aware of only one application of the Dempster-Shafer theory in the wider context of engineering projects (Taroun & Yang, 2013). This work shows the application of the Dempster-Shafer theory for handling the risk assessment of a construction project. To describe the methodology, the authors use an illustrative case study – a real project from a large construction company. A senior manager from a large construction company with an annual turnover of £4.3 billion participated in this study.

Taroun and Yang’s study also includes feedback on the methodology introduced from a large number of practitioners working in different construction companies. The evaluation criteria were based on four aspects: 1) analysis complexity, 2) methodological clarity, 3) time and resource consumption, and 4) quality and usefulness of the results. Overall, the feedback was positive. The understandable approach based on the use of the Microsoft Excel package was marked as simple and practical. What attracted managers’ attention was the concept of ignorance and the flexibility of providing incomplete assessments. Interestingly, the practitioners pointed out that they see the method introduced as particularly beneficial for large-scale and unique projects. However, some of them expressed their concern related to the large number of inputs. Finally, one of the conclusions was that in order to master the tool and input values it requires some time and understanding of different situations.

Previous applications of the NUSAP Method: Some of the experiences in applying the NUSAP system for environmental uncertainty assessments are summarized in the work of van der Sluijs *et al.* (2005). The authors emphasize that complex environmental problems are affected by all types of uncertainty, and that mainstream uncertainty methods such as the Monte Carlo analysis are not suitable to address all issues. They also point out that what is characteristic for this class of problems is that quantifiable uncertainties are dominated by unquantifiable ones. Therefore, they promote the use of both quantitative and qualitative assessments that is obtained through the NUSAP tool. Another line of argumentation lies in the fact that the NUSAP tool provides more transparency and a better public understanding of our actual capacities to understand and predict complex environmental risks. This is again closely connected to the role and impact of design and its call to consolidate (unite) different aspects, stakeholders and functions in a system. An example how the NUSAP method could be used in the oil and gas industry is available in the study by Berner and Flage (2016), and for uncertainty communication in environmental assessments in the study by Wardekker *et al.* (2008).

4.3.3. Discussion of non-probabilistic methods in engineering systems design

In order to avoid a loss of trust in scientific estimations of risks that can occur in the design of large-scale engineering systems, we need to carefully address two things. First, it is essential to clarify the limitations of our current capacities to avoid unrealistic expectations from science. Second, it is important to allow more transparency and more precise clarifications of what we know exactly, and of what is not certain, through a decision making process. To achieve this, and in order to achieve more concise decision making, we need to carefully address the following three aspects.

1) Quality of models: In addition to leading the research into the improvement of existing methods, we need to dedicate more attention to the models that have the explicit capability to deal with epistemic uncertainty. In particular, the thesis argues that in situations that are dominated by weak knowledge and information, stronger assumptions have to be made.

When calculating a single-valued probability those premises should not simply be dismissed, but rather elaborated. Even the probabilistic opponents Bernardo and Smith (2009) stated "*In practice, there might, in fact, be some interval of indifference* […]."

For instance, the application of imprecise probabilities in design risk management would allow computing interval-based probabilities when weak information is available, instead of unjustified simplifications that can occur during the modeling process. In the exploratory case studies, I identified situations where, in order to continue with the modeling, experts needed to "guess" the distributions since the information available was not adequate.

2) Quality of data: The quality of the data used for risk assessments plays a key role. A more accurate reflection of the actual state of knowledge used in risk assessment is required. The whole risk management process should be transparent and clear about the information available, its quality through assumptions made, types of analysis and their limitations, representation of the results produced, and finally their interpretation. Furthermore, no "impossible predictions" should be expected from the analysts when scarce knowledge has to be fitted into a probability distribution.

For instance, the application of the NUSAP scheme in design risk management would allow to inform decision makers about the background quality of the available information. Throughout the first exploratory case study, the senior project manager explained that transparency in the origin of the data collected is essential when communicating with the decision makers, and that this concept allows a systematic and easily understandable representation.

3) Use and interpretation of results: On the other hand, the interpretation of the results is equally important. Considering its importance in the decision making process, a general understanding of the produced results should be a priority. Higher "decision relevance" in the presentation of assessment results is feasible, as additional information is included. That brings additional validity to the results.

For instance, the application of Dempster-Shafer in design risk management would allow computing natural language statements. In addition, through the exploratory case study we identified clusters-of-risk issue that are not supported by current risk management tools (Kulikova, 2016). A cluster of risk is a term for the same risks that can occur at different locations. In a project it is necessary to know whether to mitigate that risk, rather than which specific location will be impacted. Therefore, like in the example of the electronic device, you are interested in the likelihood of a system component failure and not in discovering which component is most likely to fail.

4.4. Limitations and criticism of non-probabilistic methods

There are different opinions on whether non-probabilistic methods are the only right choice for such a complex issue as that of this thesis. First of all, putting into operation a new method for analysis, planning, and management requires a certain investment. In some cases that investment can be significant (Raz & Michael, 2001). The implementation of a novel approach in a general sense is usually followed by skepticism, since its usage might reveal unpredicted complications. Therefore, the practice needs a solid justification before even attempting to involve changes.

Furthermore, lack of operational meanings and interpretations are the key critique points for alternative uncertainty representations and treatment in risk assessment, as stated by Flage *et al.* (2014). In their discussion, Aven and Zio (2011) tackled some researchers' concerns – an imprecise probability result is generally considered to provide a more "complicated", i.e. harder to process, representation of uncertainty. In their study, they acknowledge arguments against imprecise probability, such as that simple representation should be favored. The use of imprecise probabilities goes against of the idea of simplicity, and for many, particularly first-of-a-kind applications, it will lead to initial confusion and difficulties. Others strongly defend the Bayesian approach and heavily criticize any other attempt to perform uncertainty analysis (Aven & Zio, 2011).

Implementation of the Dempster-Shafer theory of evidence was not readily accepted in the risk community. After several iterations it was proven as a valid method, or at least as a mathematically sound one, however, the use of the Dempster rule has been seriously criticized when significant conflict in information is encountered (Sentz & Ferson, 2002). Furthermore, as stated in the same report, other researchers have developed modified Dempster rules that attempt to represent the degree of conflict in the final result. Mathematical representation of epistemic uncertainty has proved to be challenging. Calculating Dempster-Shafer intervals can be highly computationally expensive (Swiler, Paez, & Mayes, 2009). Several studies, such as Bauer's (1997) elaborated on ways and methods to overcome this difficulty. Various approximation algorithms have been suggested that aim at reducing the number of important elements in the belief functions involved. More recently, Xiao *et al.* (2015) worked on possibilities to improve the computational efficiency of the evidence theory. One of the directions in current research is exactly toward transforming these methods into more computationally usable ones, including the development of a "computational toolbox" as it exists in various commercial and non-commercial forms for probabilistic risk assessment.

However, some methods, for example the NUSAP scheme, are at least less complex than other methods. According to the Brocéliande team (2015), the NUSAP's weaknesses are the novelty of the method, and the limited, but significant and growing, number of practitioners using it. In addition, the Pedigree scoring is to a certain degree subjective. The choice of experts to do the scoring is also a potential source of bias (Brocéliande team, 2015).

The main challenge with exploratory-modeling-based methods is their computational complexity and the fact that they provide results that are more complex, contextual and provisional (Bankes, Walker, & Kwakkel, 2013). Some of the details are further discussed in Chapter 7.

Reflections

There are certain limitations in terms of the collection of the non-probabilistic methods.

The literature review was based on searches conducted in Google Scholar, Web Of Science and internal DTUFindIt databases by using keywords. I first performed an exploratory search for publications (Rowe, 2014). In this search multiple fields were encountered (such as engineering systems design, project management, megaproject management, risk analysis, and product development). A large number of articles and books were identified through this search approach. The results from the exploratory search served as input for the later more structured search that was complemented with the input from subject matter experts within the risk field. This was done as a literature review can sometimes encounter challenges in identifying seminal works.

The literature review iterated between four phases: 1) planning and scoping, 2) conceptualizing the review, 3) searching, evaluating and selecting literature, and 4) analyzing the literature. Initially the collection was limited to the non-probabilistic methods in risk analysis and their use outside engineering systems design. Later it was expanded with adjacent topics of high relevance such as decision making and expert judgment elicitation. An exploratory and systematic search was applied to each new sub-topic, consulting subject matter experts in regard to the keywords. Key concepts within each discipline were identified and iteratively served as input for the other sub-topics as they were (to some extent) related.

The list is not exclusive; it is expandable as the non-probabilistic framework supports adding different methods that can fit into one of three groups. If or when developed, a new group can be added.

4.5. Summary

This chapter provides a collection of advanced methods for risk and uncertainty quantification by answering research question 2. These methods are introduced under the non-probabilistic framework and have been organized into three groups. The first group is based on imprecise probabilities, the second represents a group of semi-quantitative methods, and the third group is based on exploratory modeling.

As argued in this chapter, the methods introduced lack concrete application examples in order to demonstrate their full potential. That sets the basis for the next three chapters (Chapter 5, 6 and 7) in which the respective focus is on one group of methods. For that reason research question 3 was refined into the sub-questions (3.1, 3.2, 3.3), each corresponding to one group of non-probabilistic methods addressed in the corresponding chapters.

5. First group of methods: Improving expert judgment elicitation by leveraging different data formats

"An expert is a person who has made all the mistakes that can be made in a very narrow field."

— Niels Bohr —

Chapter 5 illustrates the problem of imprecision and how we can employ the first group of non-probabilistic approaches to better support decision making. In the context of the thesis it represents an important empirically-based step. The chapter summarizes a variety of useful data formats and aggregation methods and introduces probability bound analysis, Dempster-Shafer structures' mathematical kin.

Oil and gas investment risk analysis uses elicitation of subjective probabilities to predict the size and properties of hydrocarbon deposits and to quantify the relevant uncertainty. These risk assessments are a central element of the decision making support for allocating investments into designing exploration activities. However, the use of such subjective probabilities to describe epistemic uncertainty has been challenged by different scholars (Beer, Ferson, & Kreinovich, 2013). I examine the current practice at a large Danish company exploring offshore oil and gas reservoirs and identify possibilities for improvements in the use of expert elicitation of probabilities. I explore different techniques for generating and aggregating experts' judgments. First, the study introduces different data formats (points, intervals, weighted intervals, confidence-boxes, beta distributions, Burgman elicitations) and explores the impact of using them in the expert's ability to provide their assessments. The study finds that it is helpful for experts to provide their assessments in the format they are the most comfortable with. This depends on the specific situation, their state of knowledge, and personal disposition. Second, the study explores the effectiveness (in terms of decision support) of four aggregation methods (averaging, mixture, enveloping, pooling). To that end, I interview managers in the company regarding the implications of these alternative aggregation methods in their managerial practice, as well as their ability to work with the different data formats previously studied with the experts. While there are still obstacles to implementing sophisticated expert elicitation and aggregation practices, a clear need for such type of analysis is recognized, and options for step-by-step implementation are discussed. At the end the challenges are highlighted that relate to the impact of the choice of aggregation method on decision making.

5.1. Introduction

A common approach to risk analysis is based on using probability models to reflect aleatory uncertainties (i.e. variation in large populations of similar units) and using subjective probabilities to describe epistemic uncertainties about the unknown parameters of the probability models (Nilsen & Aven, 2002). However, the use of subjective probabilities has been criticized; it is argued that the approach provides results that are too precise in relation to the available information (Beer, Ferson, & Kreinovich, 2013).

Expert judgments are often unavoidable in risk analysis, decision making, and forecasting, as risk analysists do not have access to complete and accurate data (Cooke, 1991). The role of experts is important because their judgments can provide valuable information, particularly in view of the limited availability of "hard data" regarding many uncertainties in risk analysis (Renn, 1998; Zio, 2009). More formally, consulting several experts when considering risk estimates and forecasting problems has increasingly become customary practice after World War II (Clemen & Winkler, 1999). The motivation for using multiple experts is simply the desire to obtain as much information as possible. Expert judgments are provided on the probabilities of events of a certain interest. Procedures for combining expert judgments are often compartmentalized as mathematical aggregation methods or behavioral approaches (Clemen & Winkler, 1999).

In this chapter, I consider the problem of using multiple experts in oil and gas prospect risking. In this context, a "prospect" may be defined as "*a specific locality within an area where we possess or may acquire a lease or concession and which we believe to have geological or economic characteristics that may warrant testing by drilling*" (Harbaugh, Davis & Wendebourg, 1995). The main business of our case company is the exploration and commercial production of oil and gas. Exploration activities carry very significant costs, and are thus supported by detailed expert risk assessments, i.e. the so-called prospect risking. A decision to drill an exploration well with the objective to find a new oil or gas field should be based on a sound analysis of the prospect risks and its possible volume: what is the chance that a well will contain hydrocarbons, and how much could it be at what level of technical and commercial difficulty of extracting it? The company explores different locations and prospects, and its performance directly depends on the success rate of drilling. Test drillings are very expensive and represent a significant investment. To increase the success rate with regard to identifying perspective oil deposits, opinions of multiple experts are solicited. The study reported in this chapter investigated two aspects: 1) The effectiveness of using different "data formats", i.e. uncertainty representations, from the single expert's point of view; 2) the effectiveness of alternate methods of combining single experts' elicitation into an overall decision-support basis for management. Some of the wider literature discussing elicitations is also available (e.g. Burgman *et al.*, 2006).

The work presented here aims to fill a research gap in the practical application of expert elicitation in risk assessment. While there has been significant development in terms of research in aggregation methods from the mathematical and philosophical point of view (described in Section 5.4.1), it has been recognized (Mosleh, Bier, & Apostolakis, 1988; Cooke, 1991; Otway & Winterfeldt, 1992; Renn, 1998) that there are challenges in their application, but also in the effective representation of epistemic uncertainties (Ferson & Ginzburg, 1996; Dubois, 2010). I focus on finding solutions that are feasible to implement in practice, given that some of the advanced methods may be either too complicated or simply too time consuming to use. Risk analysis in this case can be seen as the basis for decisions to drill or not to drill a well, and as such form the link between subsurface evaluation and the business aspects of the petroleum industry.
industry.

5.2. Case study

Hydrocarbon exploration is a high-risk business that relies heavily upon predictions accompanied by significant uncertainty. Thus, there is a need to more thoroughly investigate objective means of estimating the outcomes of the exploration of prospects. The search for oil and gas can be seen both as a business endeavor and a scientific activity. In exploration, risks deal with the potential of loss, such as the cost of drilling a dry exploratory hole versus a compensating gain, such as the discovery and production of commercial quantities of hydrocarbons.

The alternative outcomes to the drilling of a prospect can be expressed as a discrete probability distribution. There is a probability that the hole will be dry and a complementary probability that it will be a discovery well. If it is a discovery, the distribution can be subdivided to express the probabilities attached to different volumes of oil that may be discovered at different levels of technical difficulty (and cost) of extracting it (Harbaugh, Davis & Wendebourg, 1995). If we can estimate the form of this distribution, it can be linked to the financial gains or losses corresponding to each of the alternative outcomes in the distribution. This linkage provides a succinct summary of the alternative financial gains and losses resulting from drilling a well, and weighs them according to the probability of their occurrence. The challenge in assessing risk is to obtain and use the most appropriate probability distribution for a prospect, and to use this as an input to financial analysis.

To properly formulate a decision, the experts in the company are asked to provide their opinion on five factors that are of key importance to finding oil and gas in sufficient volume and quality. Their competences are related to offshore drilling, where the costs per well-attempt are much higher than in onshore drilling and therefore represent a significant investment. In our case study, we focus on the five following factors (parameters), which are part of the company's concerns:

Source (We need a source of the oil - what is the risk that the source rock exists or does not exist in the area?)

Charge and timing (When the oil starts to migrate (bubble off from the source rock) - where does it go? Does it go into a structure? Is the area we want to drill favorable to receive the oil? When did the oil arrive there?)

Reservoir (A reservoir consists of reservoir rock that is completely different from the source rock. The porosity of this rock must be large, so it can receive the oil. What is the probability of reservoir rock presence in that location?)

Seal (What is the probability of a so-called cap rock existence? We must have a seal on top.)

Trap geometry (The concept of subsurface heat, the shape of the reservoir. The oil is trapped by the combination of seal and shape. Is the geometry of the reservoir suitable for oil to stay trapped?)

At the moment, the way of generating the associated probabilities (i.e. the data used in the subsequent analysis) is based on expert elicitation. Sessions with geologists and geophysicists are organized, where each of them provide their opinion as a single point estimate on each of the five mentioned parameters. The experts then discuss their beliefs and findings and jointly agree upon the probabilities.

Through the interviews with practitioners, several problems were identified with this way of working:

- **Single points do not reflect uncertainty around the estimate:**

Experts are supposed to provide a single point, precise estimate. Many argue that in some situations, given the available information, experts might not be able to provide such a precise estimate and/or the confidence in such estimates can vary depending on the situation or information available.

This information is also lost when the results are used as input for financial analyses and calculations on the strategic level.

- **The final result does not reflect whether there was an agreement among the experts:**
  As the first step in current practice, each of the experts provides his/her own estimates. Then they discuss their opinions, and after debating reach a common estimate. In this way the decision makers lose the information regarding the uncertainty of the results; the more disagreement at the beginning, the more uncertain is the result.
- **Dominant person:**
  The experts discuss their beliefs and are supposed to agree on the estimates. Here behavioral aspects become very important, as it has been noted that normally one dominant person takes the lead, so that the rest follow or are overpowered. Experts who have legitimate counterexamples are overruled.
- **Background knowledge not presented to the decision maker:**
  When results are presented to the decision maker, the background knowledge and information on which the judgments were based are not included. The decision maker only sees the final estimates, which do not clearly differentiate between assumptions and personal biases of the experts. This can significantly impact the decision quality.

The challenges documented above reflect the need to explore various ways of eliciting and aggregating the data that would support the decision maker in obtaining a final result in these situations. Furthermore, risk and uncertainty are not only associated with drilling operations, but also with field development and production after discoveries are made. These are important components of risk in the oil and gas industry, but are beyond the scope of this thesis which focuses solely on evaluating prospects. However, the approach of this thesis on how to generate expert elicitation data and how to aggregate different expert opinions is likely to be useful for all of the mentioned situations.

Section 5.3 suggests several ways of expressing data. Section 5.4 elaborates on aggregation methods.

5.3. The elicitation of expert opinions through different data formats

5.3.1. The motivation for reviewing different data formats to elicit expert opinions

The usefulness of the Bayesian approach is well argued and documented for drawing inferences and quantifying uncertainty when frequentist data are limited (Jaynes, 2003). Nevertheless, several limitations have been acknowledged, and the use of Bayesian probability has also been challenged (Colyvan, 2008; Coolen-Schrijner *et al.*, 2011; Aven & Anthony, 2015). Using a single probability (or a precise probability distribution) to describe uncertainty hides how much of the uncertainty is epistemic versus aleatory (Ferson, 1996; Walker, Lempert, & Kwakkel, 2013) as well as the strength of the underlying knowledge supporting that probability (Flage *et al.*, 2014).

These issues have led to a discussion regarding the theoretical and practical basis for alternative approaches to uncertainty representation in risk assessment (Aven & Anthony, 2015). The available research indicates that methods by which expert opinions are elicited can have a significant effect on the accuracy of the resulting estimates (Flage *et al.*, 2014). Therefore, method choice, and the corresponding data format, play a crucial role in the whole risk assessment. A choice of an alternative approach to uncertainty representation should be accompanied by the corresponding format of elicitation of opinions.

Literature on the use of expert opinions can roughly be categorized into two areas: 1) techniques for improving the accuracy of estimates, and 2) techniques for aggregating the opinions of multiple experts. From that perspective this work can be seen as a two-fold contribution. First the thesis reviews different data formats. Different formats and ways of providing expert judgments vary from field to field. However, the field lacks a comprehensive review that summarizes available options. The need for such a review is also confirmed through research that concludes that experts may be hesitant to assign subjective (precise) probabilities that may be perceived as unreliable or untrustworthy (Hubbard, 2009). Second, the thesis applies different aggregation methods for the particular case study in Section 5.4. For both parts, I use R as an open source programming language for data analysis and visualization.

The main streams of literature are provided in the following sections, but additional relevant work includes:

- Literature reviews on the topic such as: (Ouchi, 2004; Burgman *et al.*, 2006; Werner *et al.*, 2017), etc.
- Seminal and related work such as: (Mosleh, Bier, & Apostolakis, 1988; Otway & Winterfeldt, 1992; Cooke & Goossens, 2000; Cooke & Goossens, 2004; Bolger & Rowe, 2015).

5.3.2. Data formats to represent expert elicitation results

Based on a review of the literature, I consider six different data formats:

- Points (Otway & Winterfeldt, 1992)
- Intervals (Speirs-Bridge *et al.*, 2010)
- Weighted intervals (Cooke, 1991)
- Confidence-boxes (C-boxes) (Ferson *et al.*, 2003)
- Beta distributions (Jaynes, 2003), and
- Burgman elicitations (Burgman *et al.*, 2006).

Confidentiality concerns preclude me from presenting actual elicited data from the case company in this thesis. For the purpose of the analysis I wrote the code in the R language (Appendix 5) that randomly generated illustrative data, which however correspond to "normalized" versions of the actual, observed data. The purpose is to illustrate different situations, challenges attached to them, and how differently those challenges can be addressed. In practice, the numbers can be easily replaced by the exact estimates of the experts. As a result, the written code can be seen as a toolkit that is available for application in a range of specific scenarios.

For each of these data formats the thesis generated illustrative values for these four cases/situations, visualized in Figure 5:

- when the experts mostly agree on the estimates (**Consistent**)
- when they mostly disagree (**Diverse**)
- when there is one outlier (**Outlier**), and
- when their opinions can be divided into two groups (**Bimodal**).

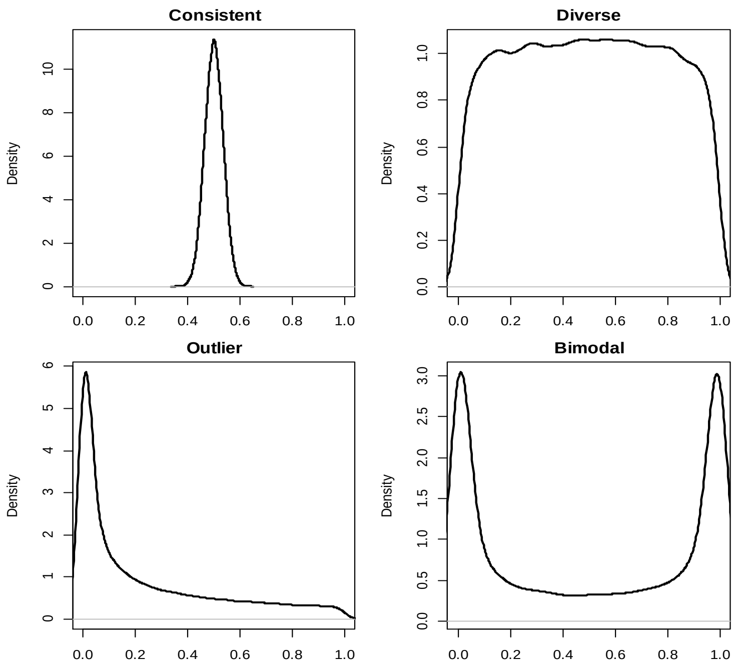


Figure 5 Four introduced situations.

The following figures (Figure 6 and Figure 7) show examples of generated data for one factor introduced in Section 5.2. In the study, this is repeated for all five factors (source, charge and timing, reservoir, seal, and trap geometry). Figure 6 illustrates different data formats generated per expert. The code is written for ten experts, however the program visualizes up to nine fields at the same time. For that reason, different data formats are plotted for only nine experts. Figure 7 illustrates these different data formats for all 10 experts.

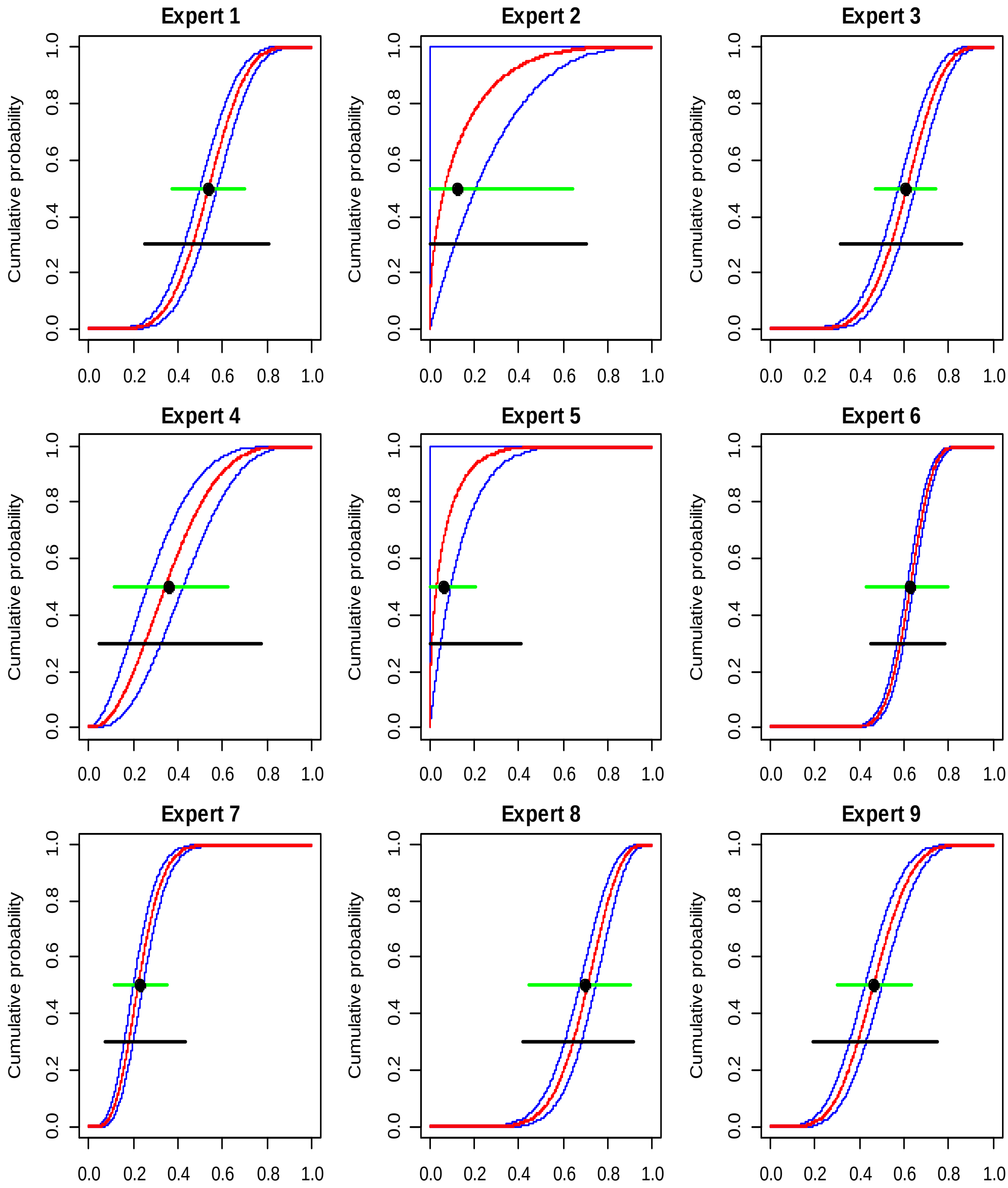


Figure 6 Generated data per each expert with all data formats.

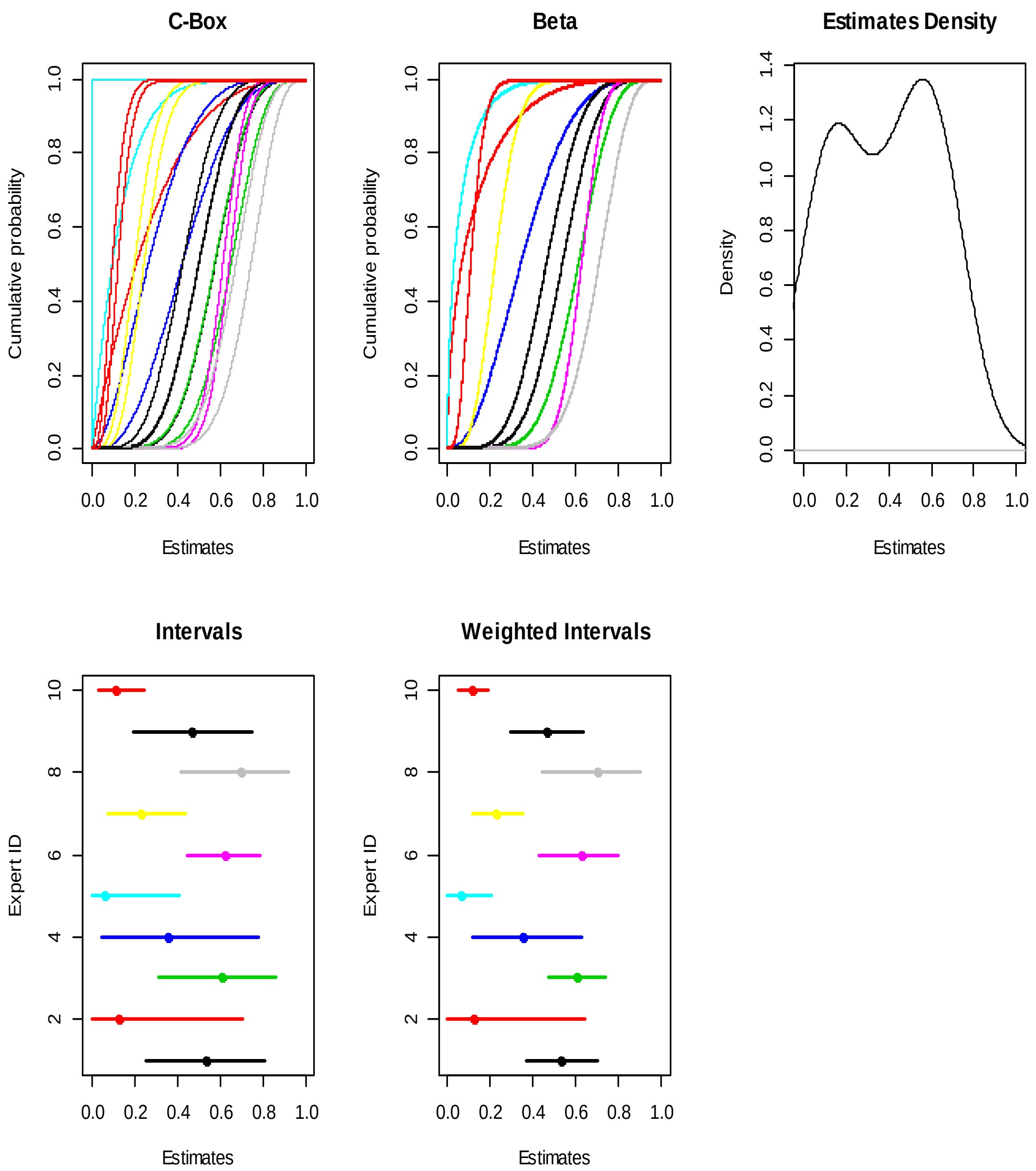


Figure 7 The controversy for each model.

5.4. Aggregation of experts' judgments

5.4.1. State-of-the-art

Forecast accuracy can be substantially improved through the combination of multiple individual forecasts (Paté-Cornell, 1996). Development of the combination of individual predictions has taken place within econometric theory (Sharpe, 1964). However, as forecasts concern future events that are uncertain, another area of expansion has been the combination of

probabilities, probability distributions, and probabilistic quantities (Clemen, 1989). During the last three decades, yet another trend was noticed in the aggregation of probabilistic judgments. It is related to the recognition that at least two types of uncertainty can be identified: aleatory and epistemic. If individual forecasts are provided in a form that admits epistemic uncertainty, a set of specialized combination rules has been developed, which is discussed in the following.

Clemen and Winkler (1999) distinguish between behavioral and mechanical ways of aggregating the judgments of experts. In behavioral aggregation, the experts themselves produce a combined or consensus view. To do so, interactive, structured discussion methods can be employed, or non-interactive consensus-building techniques be used such as the Delphi method. The so-called mechanical aggregation methods combine individual opinions by means of mathematical formulas; this process is entirely impersonal as soon as an aggregation algorithm is chosen. The simple average of point forecasts is a straightforward approach that is computationally efficient. A simple alternative is the geometric average that is defined as the n-th root of the product of n numbers (Jaynes, 2003). The conceptual drawbacks of these methods are that the aggregates are not sensitive to differential expert information, quality or dependence (Clemen, 1989).

A method compensating for a lack of individual sensitivity is to multiply each opinion by a weight that can reflect, for example, the performance of an expert in generating "correct" past judgments, or his or her background knowledge on which the judgment is based. The weights can be derived from expert calibration sessions (Cooke, 1991) and then a weighted average serve as the combined value. This method is rather difficult to implement in practice, as it is organizationally complex and time demanding. Furthermore, it is based on a rather strong assumption that experts' performance in new elicitation sessions is as good as in the calibration sessions.

All methods mentioned above have a common drawback: the degree of disagreement among the experts is lost as soon as their opinions are combined. Preserving the disagreement may have an important impact on how the decision maker uses the information and what conclusions he or she can draw. A way to preserve the level of disagreement is to consider the elicited individual point-valued probabilities as the realizations of a random variable, and to fit a probability distribution into this sample. This will yield a second-order uncertainty model that is uncertainty about the probability of interest (Jaynes, 2003). The beta distribution defined by the interval [0, 1] is a natural candidate to model the second-order uncertainty (Goodman & Nguyen, 1999).

Although there are several aggregation methods that can accommodate different analysts' dispositions (as discussed above), there are still two points that motivate a search for different approaches. One point is that experts may feel uneasy in providing precise (point-valued) probabilities on uncertain future events, and may consider that the state of their knowledge allows only interval-valued, comparative or other imprecise judgments to manifest their degree of epistemic uncertainty on the issue in question. The other point is that subjective probabilities are dependent on the subject's ability to process available probability-related data and background information. An inability to address these two points by the "mechanical methods" discussed above may lead to questions about the validity of the elicited probabilities.

Methods that allow for providing interval-valued probability judgments together with accounting for knowledge support neither fall in the behavioral nor the mechanical category. These methods suggest extracting probability-related knowledge in a predefined format from experts and then, by applying mathematical algorithms, producing probabilities. The externalized knowledge can be explicitly revised by other parties, and improved and corrected if needed. This is conceptually a very different way of deriving probabilities that is transparent, traceable and repeatable, and that contributes positively to trust in the results. Let us refer to this approach to probability elicitation as post-probabilistic. Figure 8 and Figure 9 illustrate the two approaches to deriving subjective probabilities.

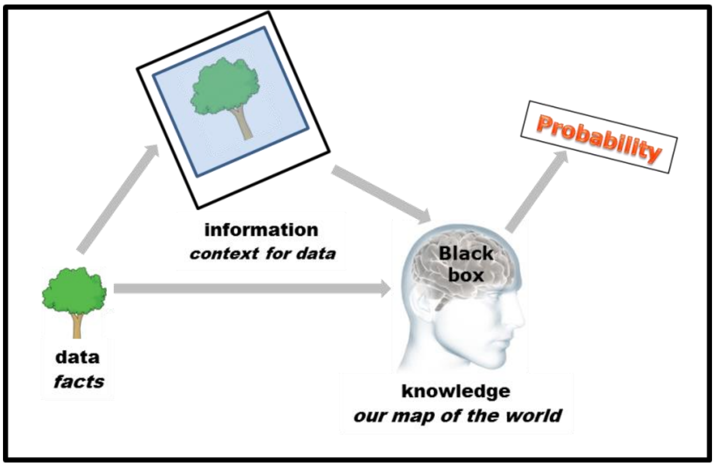


Figure 8 Direct (conventional) way of eliciting subjective probabilities.

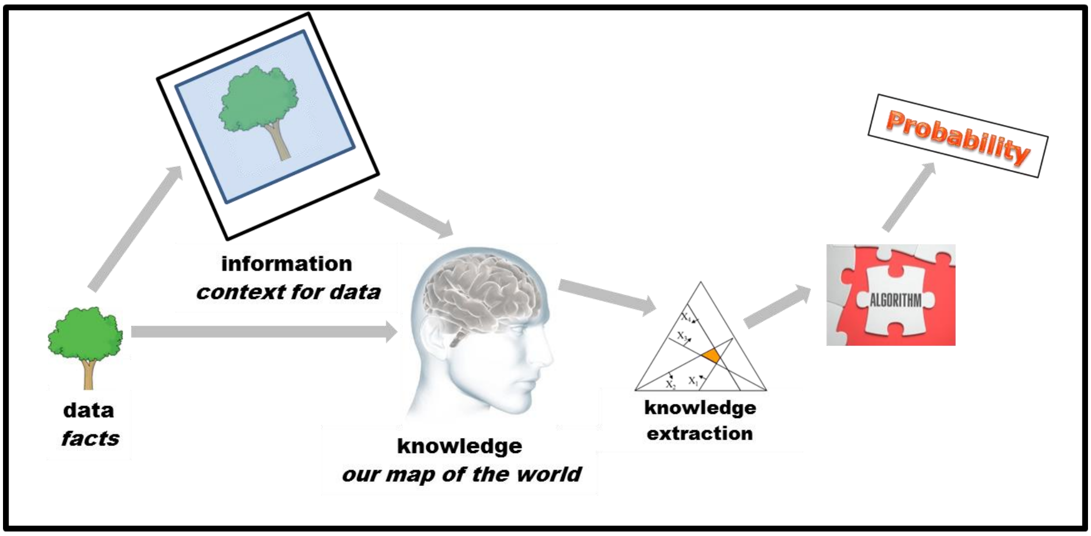


Figure 9 Non-probabilistic way of deriving subjective probabilities.

A format for extracting expert knowledge can be as comparative judgments, which are rooted in the hypothesis of comparable uncertainty (Bourgeois *et al.*, 2016). If a baseline probability for comparison is known, the result is an interval-valued probability. For example, it may be sensible to assume that the probability of failure of a gas valve being used outdoors is equal or greater than the probability of failure of a similar valve deployed indoors. Given the latter probability is known, (let denote it by Pind) the interval [Pind, 1] is the extracted knowledge. Given the probability of failure of a similar valve working in more aggressive environments, say on oil and gas offshore platforms (Poff), a narrower interval can be derived, [Pind, Poff]. This type of extracted knowledge obtained from different subjects can be combined and, as an outcome, an aggregated probability interval can be algorithmically derived. As a more nuanced approach, confidence or reliability factors can be attributed to each extracted interval or comparative judgment; a combined estimate can then be derived by applying some algorithmic rules.

A set of combination rules is available to accommodate a rather wide variety of interval-valued kinds of information provided by multiple experts, along with different assumptions with regard to the confidence level to those judgments (Kozine & Utkin, 2002; Kozine & Utkin, 2005). For example, by assuming 100% confidence in all provided interval-valued judgments, the conjunction rule should be used for combination. By admitting complete ignorance about the reliability of the elicited interval-valued judgments, the unanimity rule should be applied. Another combination rule should be used if confidence levels are different for all or for some judgments (Kozine & Utkin, 2005). Other simple rules are described in (Kozine & Utkin, 2002) and can be applied to some special cases possibly encountered in practice. Such special cases can for example be nested or adjacent intervals derived from experts.

In the following section I describe the combination of probabilities elicited from multiple experts in different formats.

5.4.2. Aggregation for the case study

Systematic comparisons between different approaches can serve to address some of the previously raised challenges and provide new insights for practitioners and guide future research directions. However, relatively few comparisons between different methods exist. For instance, Dubois and Prade (2009) compare Bayesian probabilities, belief functions and possibility theory.

Sandia reports (Sentz and Ferson, 2002; Ferson *et al.*, 2003) provide a general methods introduction and examples. A purely mathematical viewpoint could consider arbitrary operations for combining estimates involving uncertainty. However, the goal is to combine different estimates in a sensible and meaningful way. Therefore, we focus on a more specific – case – application that is of high relevance for the field and apply the four aggregation methods (averaging, mixture, enveloping and pooling).

The following aggregation methods were considered in this study:

Averaging

Introduced in Section 5.4.1.

Mixture

The idea behind mixture is that there are multiple values of a quantity that are expressed at different times, or in different places or under different situations. (Ferson *et al.*, 2003) exemplify mixture through the story of blind men who encountered an elephant and all recounted very different stories. One, feeling the trunk, said the elephant was like a snake. One, feeling the elephant's leg, insisted the animal was like a tree. A third, feeling the animal's side, asserted that an elephant was like a wall. The point of the statements is that all of these things are true at the same time. Stochastic mixture offers a perspective that can see how a quantity, like an elephant, can manifest different or conflicting values.

Pooling

Pooling can be seen as a weighted linear combination of the expert's opinions (Clemen & Winkler, 1999).

Enveloping

Enveloping should be used to aggregate the estimates into one reliable characterization when the reliability of individual estimates is uncertain. Enveloping can be used as a strategy that allows a risk analysis to proceed even though the eliminations could not be taken to completion to identify a single scenario. For instance, a police officer getting conflicting statements from the witnesses while investigating a crime might choose enveloping as a strategy (i.e. arresting everyone mentioned as suspicions) (Ferson *et al*., 2003).

Discussion on the results:

The empirical findings are in line with the literature: there is no single all-purpose aggregation method for expert opinion (Speirs-Bridge *et al*., 2010).

In general, initial results report the following. For Consistent situation (introduced in Section 5.3.2), averaging can be used. On the other hand, averaging should not be the first choice for Bimodal situation, as the information regarding the two groups of opinions would be lost. Instead envelope can be preferred. Pooling on the other hand has the highest potential when the estimates are provided in the form of distributions. And mixture has its advantages for the situations like Outlier. Figure 10 illustrates introduced aggregation methods for generated data in this study.

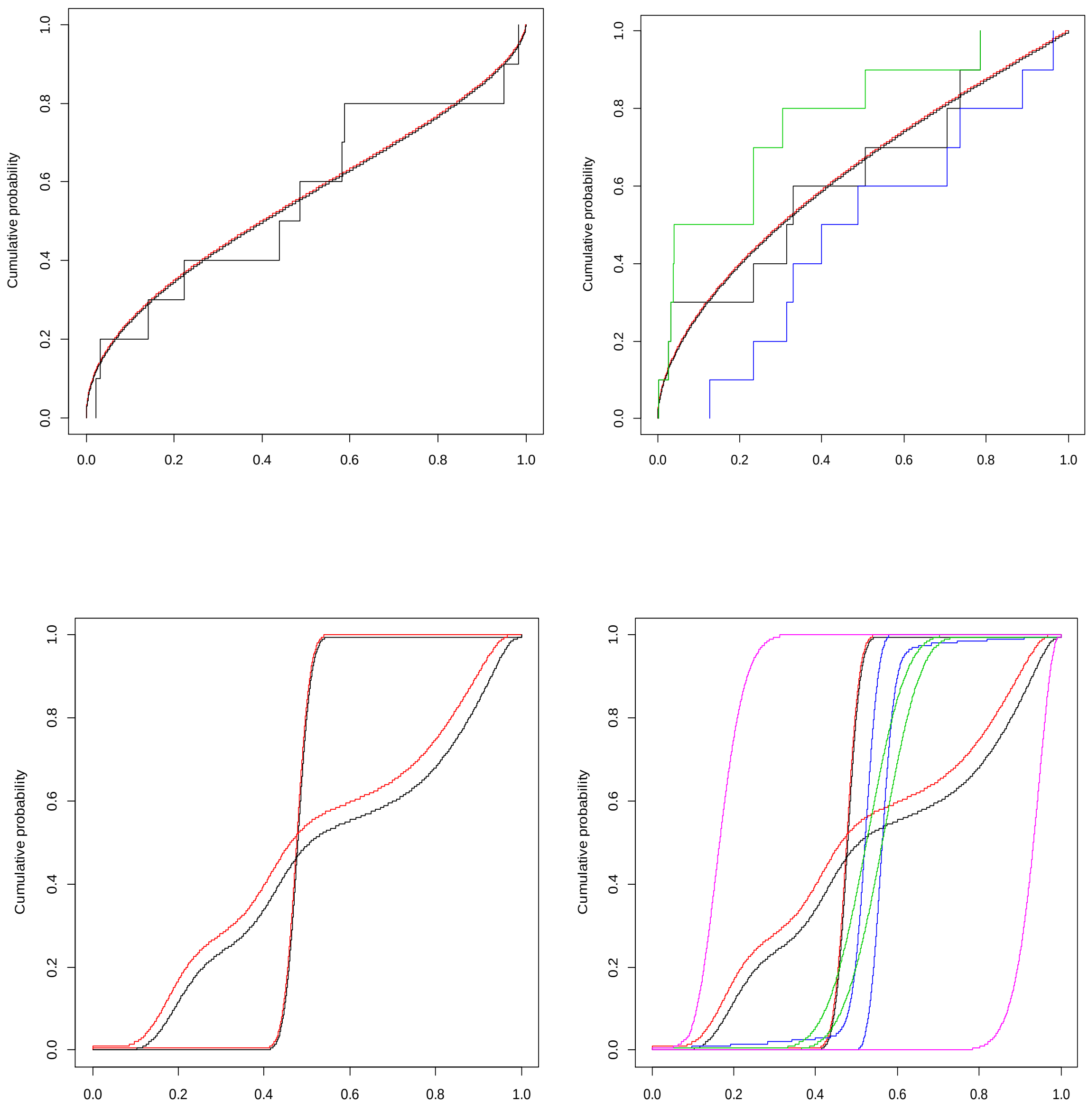


Figure 10 Example of different aggregations (averaging, mixture, pooling, enveloping) for the generated data.

5.5. Feedback from the company on different data formats and alternative aggregation methods

I discussed our preliminary results with the experts in the company, and interviewed the chief analyst about the respective potential of the introduced methods. Furthermore, I asked for feedback on both opportunities and challenges that can be noticed at this stage. To summarize the insights:

- **Highly needed:** there is a need to further develop the expert judgment elicitation process in the company. By introducing "blind risking" and more mathematically-based ways of aggregating the estimates, several challenges in the current practice can be addressed (dominant person, consensus among the experts, background knowledge, etc.).
- **Geologists' profession is hierarchical and senior driven:** seniority plays a role when assigning weighs to the estimates, while this is not necessarily justified. Aggregation methods that can overcome this limitation are greatly needed in the field.
- **Flexibility for the experts:** The ability to use different data formats represents a major contribution that allows the experts to choose the format they feel the most comfortable with. Furthermore, this allows them to express uncertainty in their estimates more precisely (for instance, by choosing broader ranges for intervals or lower confidence level in weighted intervals). Also, they are then able to make a comparison with previous cases/projects and elaborate on whether and why their estimates differ from case to case.

The interviews also reveal several challenges with the introduced alternative approaches.

At the moment, there is a gap between the required mathematical skill set and the educational background of the employees. Only a small number of employees are professionally trained in uncertainty quantification methodologies (two to three persons). The introduction of more sophisticated mathematical analyses may therefore present a problem and meet with resistance.

Furthermore, communicating the results and different data formats to other departments in the company can be complicated. The reservoir analyses and their output need to be aligned with the other analyses on the strategic level, such as Expected Monetary Value (EMV).

Some research that confirms this problem is available in the study by Caron and Ruggeri (2016). Introducing big changes that would affect many of the company's procedures is very demanding, as changes in data formats would change the overall calculations. Instead, a step-by-step implementation with examples from smaller projects can lead to further developments. What the practitioners are most interested in is whether, and if so how, the decision (to drill or not) can change. If estimates are more accurate when using the alternative approaches, it is possible to expect other initiatives. This work is one step toward achieving that goal.

5.6. Conclusions and future research

Expert judgment elicitation is an important part of oil and gas risk analysis. A number of challenges in combining experts' judgments have previously been addressed. Mostly, the research attention was on foundational and mathematical issues in order to further develop methods. However, the practical application of more advanced methods for generating and aggregating experts' elicitation judgments still requires more research (understanding) and in-depth studies. This chapter emphasizes the need to evaluate and more thoroughly explore alternative approaches to generating and aggregating expert judgment elicitations.

Research on this topic is common for multiple domains. Different engineering practices use experts in their processes; forecasting (e.g. weather or market analysis) relies heavily on combining experts' judgments. Understanding specific processes and related challenges helps developing and applying the methods in the best possible way. The thesis argues that depending on the data available, knowledge base and experts' expertise, different methods and techniques should be taken into account. There is no one-method solution for all challenges and practices. Furthermore, the choice of an aggregation method can greatly impact decision making. Therefore, a careful analysis of the available options is needed, as well as a good understanding of the benefits of using each of the methods.

The main contribution of this chapter is to collate different data formats used in expert judgment elicitation. For the particular needs of the case study, I provide options regarding aggregation methods. Furthermore, the thesis describes the obstacles to the practical application of some of the methods. However, the thesis first raises awareness of the lost information regarding uncertainty about the numbers produced by experts. Second, the thesis provides freedom for the experts to express their opinions in the format they feel most comfortable with. Third, the thesis offers a useful toolkit for conducting analyses, not only for petroleum exploration practices, but for risk analysis in general.

There are several options to further improve current practices. An interesting area of research could be to investigate techniques for the evaluation of aggregation methods. Another interesting direction can be seen in decision analysis based on different aggregation methods. Moreover, the need is also recognized to explore ways of visualizing uncertainty that are sufficiently understandable and useful for communication on all managerial levels.

Finally, decision makers should be encouraged to fully take into account the results of the analyses. Transparency in the level of uncertainty surrounding the results would help in attaining the necessary level of trust in the calculations.

Reflections

The coding of the collected data was checked in multiple ways. Different subject matter experts (researchers) were shown parts of the raw data and the analyses to discuss the method of coding as well as the analyses themselves. A professor and two PhD students completely independent of the project scrutinized the mathematical approach to aggregation in order to verify that the approach is in line with the intent. Even though risk researchers, mathematicians and data scientists reviewed the applicability of the method and the application process and analysis, there may have been a bias toward certain types of aggregation models that I favored, partially due to the extent to which they were discussed and due to my background/skills in mathematics.

The case/sample selection was chosen as a representative case (Flyvbjerg, 2006a; Eisenhardt & Graebner, 2007) due to the fact that the company has a mature risk management with experienced people and available resources. Many corporations of this size engage in purely quantitative risk management that serves as input for decision-making. Case selection was performed carefully as the selection process is crucial (Flyvbjerg, 2006a; Gibbert, Ruigrok, & Wicki, 2008). The selected company employs a separate risk function and has experience with aggregation methods. Expert judgment elicitation is relevant for its decisions.

The field of oil and gas is highly reliant on expert judgments, making it a very relevant field to study. As many other fields employ similar, even identical tactics in equally volatile markets, the learnings from this case study may inform other contexts. To strive for improving rigor in relation to external validity, an overall nested case study approach was applied (Gibbert, Ruigrok, & Wicki, 2008). Multiple case studies were applied within one company (please see Chapter 6), which is known as a nested case study. This improves rigor as the studies can inform each other and the researcher obtains more complete knowledge (Gibbert, Ruigrok, & Wicki, 2008).

6. Second group of methods: Qualifying risk quantifications through the application of the NUSAP tool and de-biasing of expert judgment

"Assumptions are made and most assumptions are wrong" – Albert Einstein

"Assumptions are dangerous things" – Agatha Christie

This chapter focuses on the communication, visualization and representation of the background knowledge in risk assessment. Based on a case study, a set of methods from the second group of non-probabilistic approaches is applied to visualize uncertainty surrounding data and risk quantification results. The specific method chosen is NUSAP, as previously introduced in Chapter 4. In addition to the general application of the method, the research interests developed during the case study to include a deep focus on two aspects of the NUSAP method application: Part of the qualification of risk quantification results (as through the NUSAP approach applied in this chapter) is to evaluate the degree of expert judgment biases in risk quantification. This resulted in the development of an approach to quantify and correct biases in expert judgments in risk assessments. Similarly, a key (arguably, the key) aspect of NUSAP-type characterization approaches is to inform and to increase decision-makers' trust in risk quantifications. That area will be the second "deep dive" in this chapter, in addition to the general NUSAP method application.

The synthetic case study was developed in collaboration with the company in the oil and gas sector, as introduced in Chapter 5, where the company's risk management and expert judgment elicitation processes are described in detail. Here the focus is on the quality of the background knowledge available for the assessments, and ways to communicate it when presenting the results. For this purpose, first the NUSAP tool is employed. The findings are described and documented in Section 6.2, together with feedback for practitioners. This fruitful collaboration led to opening two related "deep dive" questions: First, how can we facilitate the communication among different stakeholders (both internally and externally) to clearly articulate levels of trust. For this purpose, I introduced them to the RiskImaging tool and demonstrated how their practice could benefit from such a type of analysis (Section 6.3). Second, we developed a tool for calculating and correcting biases identified in their expert judgments as presented in 6.4. The chapter summary and conclusions are described in 6.5.

I argue that the approaches introduced in this chapter, along with the elaborated findings, are applicable to different engineering practices with a similar risk management process and/or challenges – they do not depend on the specific application area, but on the set of risk management (and risk quantification) methods that are employed by an organization. The cases serve to provide guiding examples to practitioners.

As a general note, the discussions in this chapter are based on representative data generated in collaboration with the case study partners. Due to the confidentiality issues, the actual data cannot be used.

6.1. The importance of background knowledge and its representation

The background knowledge for an analysis refers to the available information and data, prior experience and knowledge of the managers and analysts, and the understanding of a phenomenon (such as a particular process, technology or system components and their interconnection, etc.). The number and strength of assumptions accompanying such analyses can vary significantly, and thus impact the quality and trustworthiness of the final risk assessment result (Apostolakis, 2004). The treatment and management of uncertain assumptions in quantitative risk assessments and during the subsequent processes of risk management and decision making have recently attracted attention from a number of researchers (Aven, 2013b). One of the reasons for the classification of uncertain assumptions is because it can be useful for determining how uncertainty can be treated in a risk assessment. However, the benefit of such a classification and clarification can be further improved if the critical assumptions can be effectively communicated to the decision makers in order to be able to decide how to manage the risk/uncertainty in question, and what the accuracy and overall "trustworthiness" of the results is. It is also important that other stakeholders that are not directly involved in the implementation of a risk assessment can get an overview of which assumptions have been identified as critical. The involvement of other stakeholders (for instance team leaders) can in some cases be essential as they can be closer to the operation (the execution of an activity) than the decision makers and/or risk managers, making them better suited to point out potential deviations from critical assumptions.

Another underlying premise is that there should be a balance between the resources used to treat an assumption and the criticality of the assumption in question. The criticality of an assumption can be considered to depend on the assumption setting to which it belongs. Some research on this topic was previously done by (Flage *et al.*, 2014; Berner & Flage, 2016). The criticality of assumptions in the different settings ranges from low to high. A setting with high criticality has weak background knowledge, a moderate/high belief in a deviation from the original assumption, and a moderate/high impact of this deviation on the risk index. An example of a range of different settings is provided in Table 5.

Table 5 Settings faced when making assumptions in risk assessments (based on Berner & Flage, 2015)

Belief in deviation from original assumption	**Sensitivity of risk index**	**Strength of knowledge**	
		Strong	Moderate/Weak
Low	Low	Setting I	Setting II
	Moderate/High	Setting III	Setting IV
Moderate/High	Low		
	Moderate/High	Setting V	Setting VI

For a critical discussion, I would draw attention to the Setting I as one way to characterize black swans. A black swan is described as "*a surprising extreme event relative to present knowledge and beliefs*" (Aven, 2013a). When an assumption is recognized as Setting I, it does not mean that the assumption absolutely cannot deviate – it would "just" surprise us, as black swans do. Furthermore, in some cases the sensitivity evaluation can be misleading. Further details on the strategies related to assumption setting and their impact when deciding the overall risk management strategies are e.g. elaborated by Berner (2016).

On the other hand, when involving, informing and/or reporting to the stakeholders on a higher hierarchical level, such as corporate portfolio and strategy level, it is of great importance to have the ability (in the form of a tool) to communicate the quality of background knowledge on which the estimates are made, as it may impact the overall corporate direction or lead to additional research. An effective, easy to comprehend and sufficiently informative approach is desirable. For these reasons, the chapter explores the NUSAP notational scheme that was developed in order to address new types of policy problems referred to as problems "*where facts are uncertain, values in dispute, stakes high and decisions urgent*" (Funtowicz & Ravetz, 1990).

6.2. Visualizing and communicating uncertainty around data and analysis' results

Efforts to develop approaches to represent uncertainty in risk assessments follow both quantitative and semi-quantitative lines, where semi-quantitative is to be understood as quantitative representation supplemented with qualitative assessments of aspects not sufficiently and appropriately captured by the produced numbers. The latter type of approach can be referred to as extended quantitative risk assessment, and has parallels with the so-called NUSAP notational scheme of uncertainty and quality in science for policy.

Working in the field of policy-related research, Funtowitcz and Ravetz developed a novel approach for dealing with the uncertainty and quality of available information. The acronym "NUSAP" stands for Number, Unit, Spread, Assessment and Pedigree, the five elements that constitute an information set regarding uncertainty in their method. The underlying idea is that a single number does not inform sufficiently, therefore properties of numbers should not be ignored. Moreover, the developers' view on certain uncertainties associated with problem framings and assumptions can only be described through a qualitative connotation, since those uncertainties cannot be quantified.

The NUSAP tool allows results of a risk and uncertainty analysis to be represented as a "Number" accompanied by additional information to allow decision makers to interpret the overall meaning of the value. Here it is introduced through the four additional categories. "Unit", which may also be a conventional kind, expresses whether we are talking about percentage, money or something else. Uncertainty is in this case addressed by "Spread" and "Assessment". Spread is characterized by random error or a variance of statistics. Those values are obtained through sensitivity analysis, Monte Carlo simulations or in combination with experts' judgments. On the other hand, Assessment expresses the systematic error, which for statistical test might be the level of significance or for estimates just the qualifier pessimistic vs optimistic. Finally, the novelty of the tool comes with "Pedigree". This category informs on the information feed, or in other words, the origin and quality of data used for the analysis. By providing detailed information to the decision makers on how data were collected, what the sample size and similar measures are, the NUSAP measure allows them to judge the overall value and meaning of the presented results. In order to minimize subjectivity and arbitrariness, it eliminates uncertainty or misinterpretation on whether, for example, a probability measure is just a guess or based on extensive simulation and testing. It is given in the form of a matrix, where qualitative information for one pedigree criterion is structured. Different Pedigree matrixes can be obtained for different sorts of information. Thus, the NUSAP scheme provides qualitative information on the degree of aleatory uncertainty.

One of the biggest strengths of the NUSAP is that it can be combined with already existing practices (and methods). The utilization of visualization tools is considered essential, as it has been argued that multiple representation techniques are beneficial for learning (for instance for remembering) (Cheng, Lowe, & Scaife 2001; Ainsworth & Loizou, 2003). The combination of pedigree scores (a sequential representation technique) and radar diagrams (a visual representation technique) is an example of a combination of multiple representation techniques. In addition, diagrams enable faster scanning (search) through information (Larkin & Simon, 1987).

The case study and data generating process

The synthetic case development consisted of the following steps: 1) meetings and interviews with practitioners, 2) formulating the characters and their personalities, 3) discussing the possibilities for the application of the proposed approaches, 4) sampling the data needed for the application and evaluation of the proposed approach for communicating the uncertainty surrounding the results, which led to 5) identifying the potential experts' biases and other behavioral and organizational aspects that should be taken into account and are of interest to a particular company (this would also serve as the input for the additional analyses described in the following Sections 6.3 and 6.4).

The company's risk management process: As described in the previous chapter, it organizes expert elicitation sessions during which the experts are requested to provide their estimates as single number inputs. They provide their opinions in the form of estimates for the five different factors regarding the company's oil and gas exploration (Source, Charge and timing, Reservoir, Seal and Trap geometry). After that, they need to find a consensus on the joint estimates, which is then presented and used for decision making (Figure 10). This input is used for deciding on whether or not to invest in drilling in a certain location, or when choosing between different locations. The higher need for accuracy in the estimates comes from the fact that the market is increasingly competitive, most resources of the North Sea (that is the main location focus for some of the projects) are already found and it takes a while to design and then put into operation the whole system once oil and gas are found.

A number of challenges were identified (during several brainstorming sessions) and described in the previous chapter. For instance, it was noted that there is often a dominant personality among the experts that takes over the discussion and influences the others' opinion. In addition, the input from younger/less experienced colleagues is taken (valued) notably differently from that of the more experienced ones, and so on. Currently, they do not reflect on the information/data on which the opinions are based, nor on if there was agreement or disagreement among the experts in the first place.

I generated the data illustrated in Figure 11 based on their current process. After each of the experts provides the estimates for the five different factors (Source, Charge and timing, Reservoir, Seal, Trap geometry), the total estimates are calculated based on the geometric mean for the concrete prospect. Paleocene Prospect represents an area (location) that is well explored and understood, but most probably has small quantities of lower-quality oil and gas. On the other hand, Triassic Prospect represents an area (location) that has not been explored much, maybe in a few test drills, but has the potential for containing large quantities of high-quality oil and gas.

Paleocene Prospect								
Risk Element	**Committee member Pi** 							**Consensus**
Source	0.8	0.8	0.8	0.8	0.85	0.8	0.8	0.8
Charge and timing	0.7	0.5	0.7	0.7	0.65	0.7	0.6	0.6
Reservoir	0.9	0.9	0.9	1.0	0.95	0.9	0.9	0.9
Seal	0.5	0.5	0.5	0.5	0.55	0.5	0.5	0.5
Trap geometry	0.8	0.8	0.8	0.8	0.85	0.8	0.8	0.8

Triassic Prospect								
Risk Element	**Committee member Pi** 							**Consensus**
Source	0.1	0.1	0.1	0.2	0.15	0.1	0.1	0.1
Charge and timing	0.7	0.5	0.7	0.7	0.65	0.7	0.6	0.6
Reservoir	0.8	0.8	0.8	0.8	0.85	0.8	0.8	0.8
Seal	0.5	0.5	0.5	0.5	0.55	0.5	0.5	0.5
Trap geometry	0.8	0.8	0.8	0.8	0.85	0.8	0.8	0.8

Figure 11 The synthetic risk data.

The application of the proposed approach

The application of the proposed approach consists of the combination that includes the NUSAP matrix (Pedigree) and radar diagrams. They are introduced as the extension of the current process. As the first addition to providing the regular data/analysis, the expert team is asked to deliver the additional two tables. Second, the expert team works on providing radar diagrams.

Step 1:

In order to help experts to communicate the two different situations (Paleocene and Triassic Prospect), the use of the NUSAP matrix (Pedigree) is proposed. For both cases the differences are highlighted (Table 6 and 7).

Table 6 Pedigree scores for Paleocene Prospect data

Level	Proxy	Empirical	Theoretical basis	Method	Validation
4	Exact measure	Large sample direct measurements	Well established theory	Best available Practice	Compared with independent measurements of same variable
3	Good fit or measure	Small sample direct measurements	Accepted theory partial in nature	Reliable method commonly accepted	Compared with independent measurements of closely related variable
2	Well correlated	Modeled/ derived data	Partial theory limited consensus on reliability	Acceptable method limited consensus on reliability	Compared with measurements not independent
1	Weak Correlation	Educated guesses / rule of thumb estimate	Preliminary theory	Preliminary methods unknown reliability	Weak / indirect validation
0	Not clearly related	Crude speculation	Crude speculation	No discernible rigor	No validation

As illustrated above, the first table corresponds to the Paleocene Prospect for which the committee agrees on the overall scoring Level 3. That means that there are ways to measure the area (ground) and a number of test results are available (sample size is representative). The theory and method used to acquire the data are well-established in the field and used frequently. The committee pointed out that for some Paleocene Prospects, scoring can even go to Level 4 for sample size and theory and method applied. In total, the available background knowledge represents a sound basis for conducting risk analysis and supporting related decision making.

Table 7 Pedigree scores for Triassic Prospect data

Level	Proxy	Empirical	Theoretical basis	Method	Validation
4	Exact measure	Large sample direct measurements	Well established theory	Best available practice	Compared with independent measurements of same variable
3	Good fit or measure	Small sample direct measurements	Accepted theory partial in nature	Reliable method commonly accepted	Compared with independent measurements of closely related variable
2	Well correlated	Modeled/ derived data	Partial theory limited consensus on reliability	Acceptable method limited consensus on reliability	Compared with measurements not independent
1	Weak correlation	Educated guesses / rule of thumb estimate	Preliminary theory	Preliminary methods unknown reliability	Weak / indirect validation
0	Not clearly related	Crude speculation	Crude speculation	No discernible rigor	No validation

In case of the Triassic Prospect, the scoring is significantly lower (Table 7). Hardly any correlation is available, data are based on educated guesses, and there is just an initial understanding of the theoretical and methodological basis for the assessment. In total, the available background knowledge is significantly weaker than in the previous case. That is why further analyses and knowledge gathering are recommended.

Overall, the practitioners consider this step easy to understand and easy to present. It provides the means to screen the main information quickly and takes into account valuable information (that would otherwise be lost). Furthermore, it is recognized that once the decision-makers are familiar with this step (table content and format), they can comprehend the information faster. That is a big plus as it means that it would not impact the length of the meetings needed for reporting, which is of high importance when having limited time available with managers.

Step 2:

The already presented work can be further supplemented with the following information represented in the form of radar diagrams (Figure 12). In the first place, we inform decision makers (through a visual form as well) about 1) the *Availability of geological* studies on which assessments are based. This refers to the fact that depending on the project and the potential drilling location, a different number of studies can be available, as well as their details (quality). For instance, in some areas the company invested in a number of test drills, where in other cases there was less testing. 2) *Time for assessment* brings the information regarding the process for the assessment. If the committee has been requested to provide the opinions under time pressure (either due to business/market criticality or organizational reasons) it should be noted that such an assessment can lack a detailed analysis.

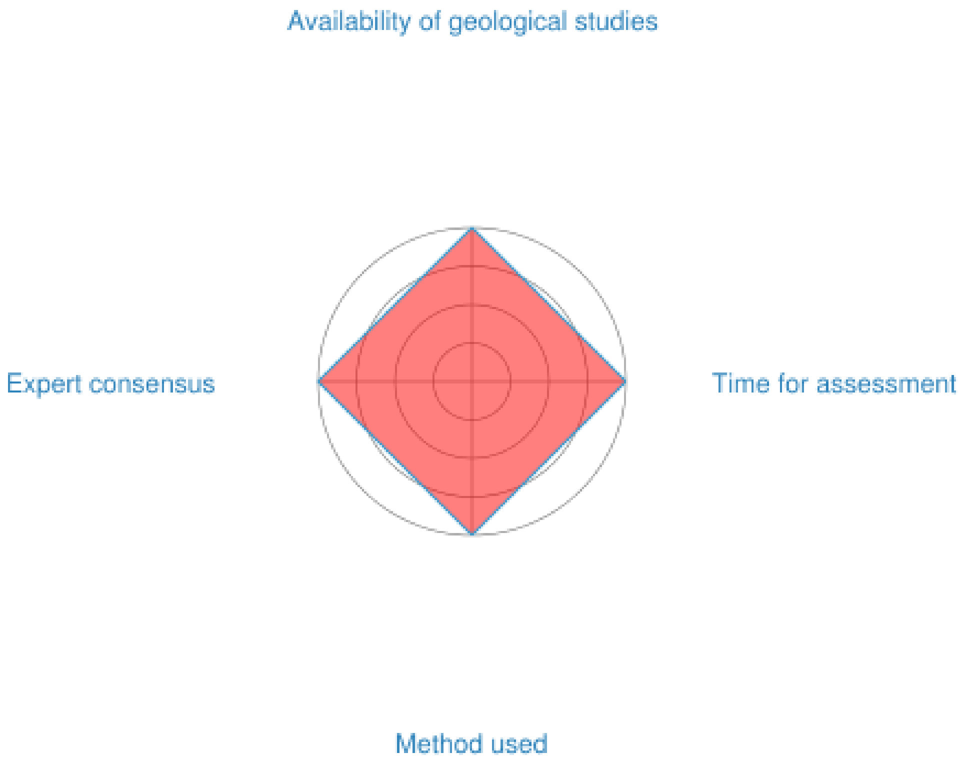


Figure 12 An example of a good quality of inputs for the assessment.

Furthermore, 3) *Method used* refers to information regarding the method used to collect, analyze and interpret data – if it is a common, well-established approach in the field that all experts are aware of, if it is a trial version of a novel approach, a not commonly used one, or if only some experts had a previous opportunity to use it. Finally, the 4) *Expert consensus* provides input on whether or not consensus was reached among the experts (the level of disagreement among practitioners). Because they need to provide a joint single number estimate, the information regarding the level of uncertainty surrounding that single estimate is lost. There can be a number of reasons for disagreement among the experts, which should not be ignored. On the other hand, if there is one fully shared view on a particular situation, it strengthens the argument for a certain decision.

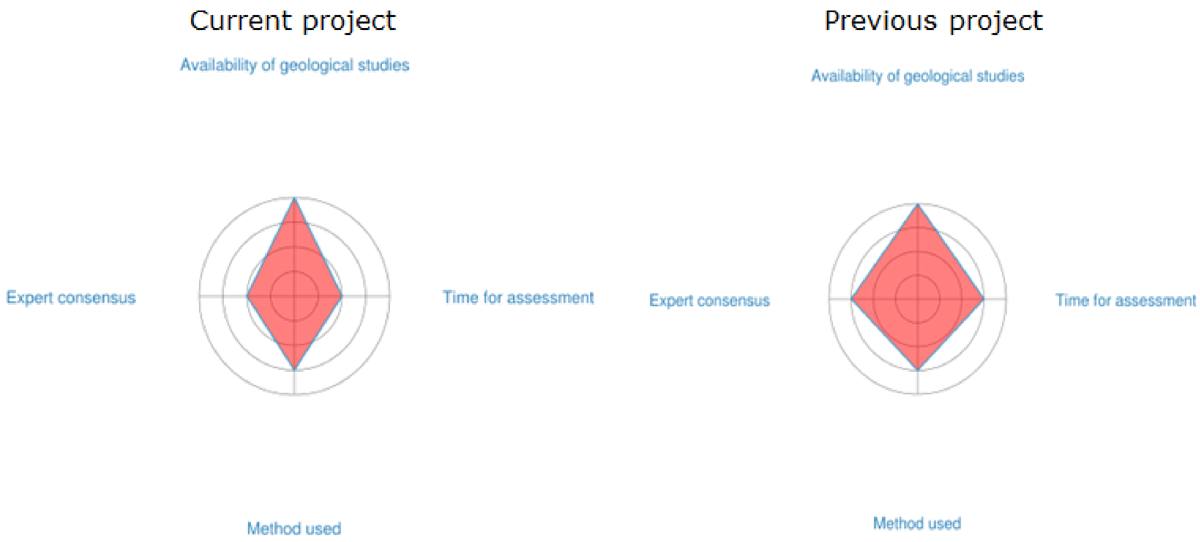


Figure 13 The comparison with another project.

Also, a visual comparison between different projects (Figure 13) allows argumentation to the decision makers if 1) more data/time is needed, 2) more research/resource is needed, or 3) simply elaborating on some detail that in the experts' view can be of significant value.

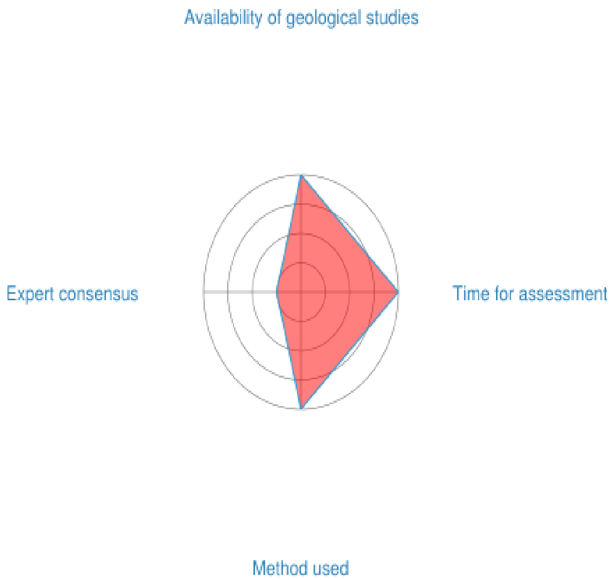


Figure 14 An example of a lack of consensus.

Finally, we identified the situation of a particular interest that requires further details. In case there is hardly any agreement among the experts, the question is what can we do other than just inform the decision makers? This Step 2 supports this need in the form of an additional feature to the already available results. The radar diagram (see Figure 14) enables the visualization of this uncertainty and can be used to facilitate the discussion around the need to understand where this disagreement comes from and if more information should be collected.

6.3. Behavioral and organizational aspects

Step 3:

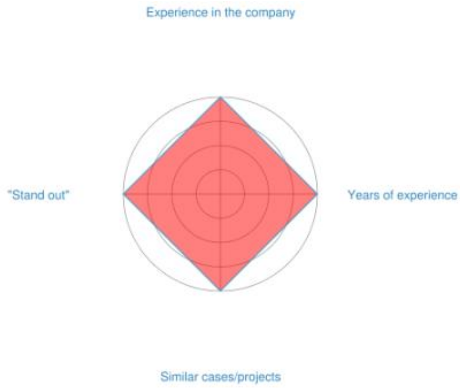
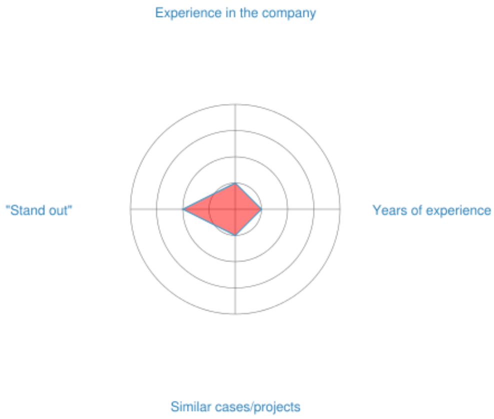
A better understanding of the team working on the project is seen as one way to further understand the above-mentioned type of uncertainty. For each of its projects the company engages five to ten of its experts (depending on present availability, size of the project, priority, etc.). The "personalities" identified are described below (Table 8).

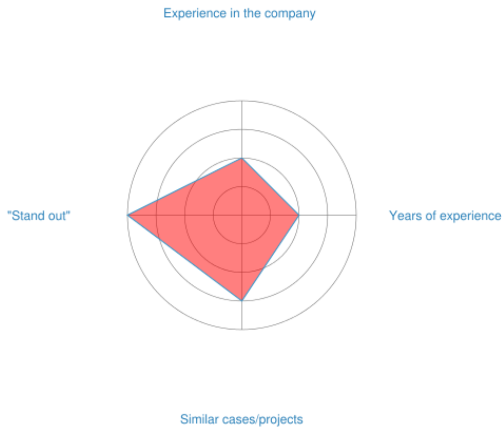
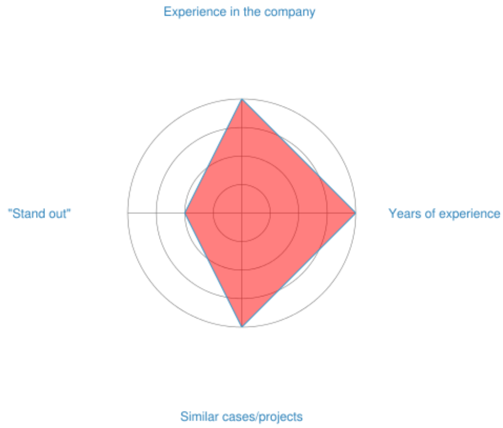
For each of the characters/personalities a specific radar diagram is formed. Again, four aspects are considered according to the company's needs:

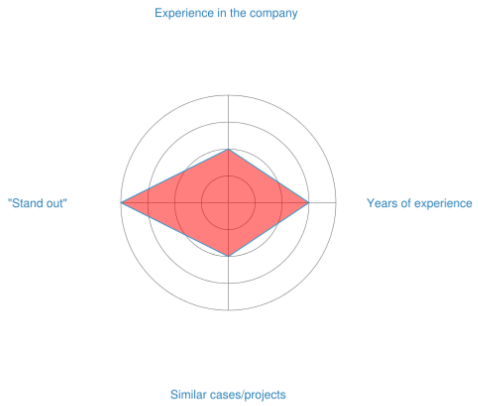
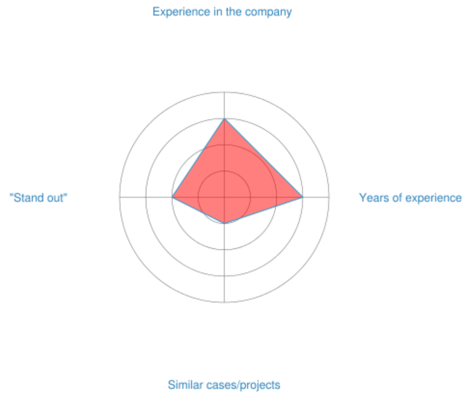
1. Experience in the company (knowing the processes, techniques, work environment)
2. Years of experience (in the field, could be working in different companies)
3. Similar cases/projects (as projects/locations can vary)
4. The capability and willingness to "stand out" during a meeting (not everyone is ready to confront the others and defend their initial estimate).

Figures illustrating personalities are provided next to the diagrams, for the sake of more "commercial" visualization and utilization inside the company (Table 8). The inspiration for the characters comes from the children's stories of Mr. Men and Little Miss (Hargreaves, 1971).

Table 8 Meet the team: Experts' characters and corresponding radar diagrams

Meet the team		
Character	**Character description**	**Character radar diagram**
Mr. Clever (correct prospect risk – opposite to Mr. Wrong) 	With 30 years in the industry but with no international experience (due to family reasons) this is the company's best and most experienced North Sea geologist. He always undertakes forensic levels of technical analysis, with an eye for detail and loves prospect characterization. He has drilled, seen a huge amount of the company's portfolio, and has developed a 6th sense for risk calibration in the North Sea.	
Little Miss Sunshine (wrong risk – opposite to Mr. Grumpy) 	Only 3 years in the industry, very keen and enthusiastic. Not enough experience to be jaded, skeptical or even realistic yet. After being complimented in her first appraisal for being really positive she has since made a virtue of pathological optimism. Rather than work the prospect in any detail, she has anchored on the play risk as she thinks this will provide a hiding place in the meeting!	

Mr. Brave (right risk in the end but made up of extremes) 	He gets off the fence early and either loves or hates each of the petroleum systems' elements. He was on a training course that emphasized polarization of risk and he emphasizes this all the time now. He does not do anything by halves!	
Mr. Grumpy (wrong risk with a pessimistic bias) 	Thirty years with the company, and close to retirement. He is long in the tooth and an old skeptic. Thinks the basin is old and tired and wants our prospecting fortunes to reflect his basin scale bias. He thinks all new concepts have been tested, and if it works he thinks it will be small benefit. A classic line in the risk meeting was that "he worked these prospects in 1985 and they were rubbish then and still rubbish now."	

Mr. Fussy (roughly the right risk but lots of false precision) 	Tends to work things to the nth degree. Cannot see the wood for the trees and always thinks that we know more than we actually do. Will argue the toss about a change from 0.95 to 0.97 for the sake of looking like he was "right" and won an argument. Loves to de-rail the meeting and twist off about tiny changes. Has been seen cutting his grass three times on the weekend!	
Mr. Topsy-Turvy (right risk but wrongly attributed – seal and trap mixed up and source and charge mixed up) 	Mr Topsy-Turvy (right risk but wrongly attributed – seal and trap mixed up and source and charge mixed up)	

Mr. Wrong (not a clue! - random number generation) 	Five years with the company but has worked for 10 different firms already in his career (all with different risk methodologies). In the middle of a messy divorce and often not fully concentrating while at work. Did not open the pre-read and has no idea about these prospects and is hung-over and off the pace today, claiming he had a dodgy kebab last night!	

The addressed behavioral aspects support the better understanding of uncertainty coming from background knowledge and information. For instance, in the case of an outlier (situation Outlier in Chapter 5), the focus could be on that character, looking for the reasons why he/she stands out from the group. It should be discussed whether the opinion should be ignored in the particular case or accepted/ further analyzed. In the case of a bimodal situation (situation Bimodal described in Chapter 5, representing two groups of opinions inside the expert committee), the diagrams can be grouped and visually evaluated if the groups are based on a certain pattern (i.e. more experienced team members versus less experienced) or not.

6.4. Analysis of different concepts

Step 4:

When considering ways to better understand behavioral and organizational aspects, it is essential to explore methods that can support the communication of different perspectives. After understanding different personalities (and stakeholders) it is also necessary to facilitate a discussion about different views. For that reason, I investigated the RiskImaging tool, which was developed in the USA and has been used in the pharmaceutical industry to address similar needs (Ramas, 2016).

Due to confidentiality reasons, I describe here what the tool enables and how I suggested (exemplified for the company’s needs) to use it in its processes/meetings/workshops. The goal of using the tool is to predict/document/visualize how risk is perceived by different stakeholders (depending on if we use historical data, current measures, or experts' judgments). The tool allows creating risk profiles – visualizing combined frequency and adversity intervals (see Figure 15).

Risk profiles visualize uncertainty in terms of estimated frequency and adversity of risk (Goerlandt & Reniers, 2015). In case of drilling success, the frequency is estimated for the success rate, whereas adversity refers to the amount and quality of oil. Uncertainty (expressed in the form of the interval) is based on the collection of experts' judgments – the more coherent and similar their estimates, the tighter the interval, and vice versa. The first risk profile (on the left) corresponds to the Paleocene Prospect (higher frequency and lower adversity). The second risk profile (on the right) corresponds to the Triassic Prospect (lower frequency of success and higher adversity).

The illustrated risk profiles can be compared with others from different projects, and the arguments for deriving decisions can be provided. In case of a large overlap of risk profiles, there is a strong argument for deriving the same drilling decision as in the compared project, whereas in case of a lack of similarity in risk profiles it becomes evident that either additional research is needed, or the decision should be the opposite to that made in the compared project.

Emotions and attitudes affect perception, which also varies across interest groups. Without further discussion on the neuroscience approach of risk perception, the thesis provides insights for attitude parameters – burden of proof, dispute tolerance, and uncertainty in adversity and frequency. The tool further allows visualizing the estimates of a single expert inside the risk profile – enabling discussion on a certain bias of the specific expert. Furthermore, the tool allows adding grouped stakeholder views – for example, differentiating between managers', scientists' or economists' groups of opinions.

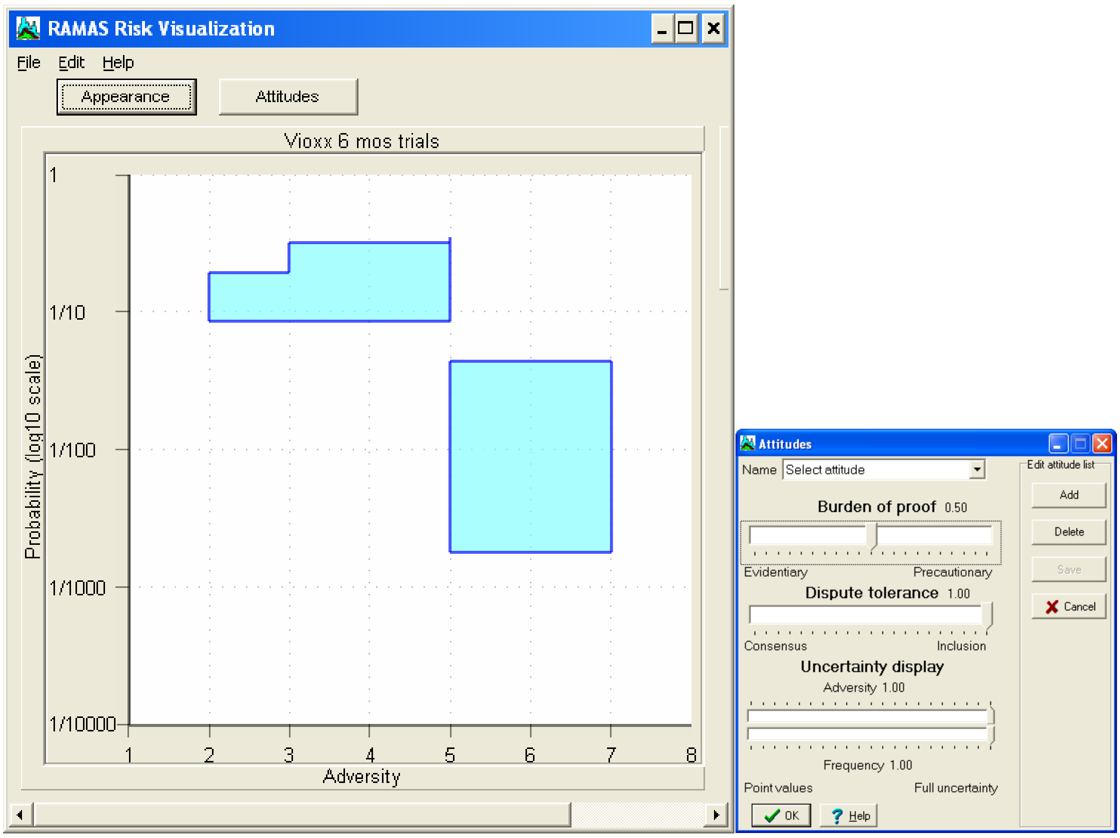


Figure 15 Preview of Risk Profiles in Risk Imaging.

Step 5:

Although simple scaling, shifting, or inflating corrections are widely used to account for biases and overconfidence, much better distributional information is usually available to the analyst. Fully using this information can yield corrected estimates that properly express uncertainty and make them more suitable for use in risk analysis and decision making. These advanced corrections express biases as distributions rather than as simple scalar values.

Such methods (and calculations) are necessary, because in many cases estimates in risk analysis have been documented to be biased in a number of different fields (Cooke, 1991). For instance, some biases are such that analyses based on these point estimates are guaranteed to result in average net losses over time, having exactly the opposite effect of their intended purpose. Some of the biases in estimates are just based on negligence concerning model assumptions, which would be relatively simple to fix with more correct assumptions, or more honest contractors. For example, ignoring the fact that costs of materials increase significantly over the time that is required to complete a large public sector project has been found to explain

20–25% of bias in estimated costs of these projects (Morris, 1990). Most of the bias, however, tends to be related to much more complicated psychological and sociological phenomena, such as self-interest bias (or a lack of incentives for accurate estimates, e.g. Flyvbjerg, Holm, & Buhl, 2007), undue optimism and risk aversion (Lovallo & Kahneman, 2003), poor management, poor communication, bureaucratic fecklessness (Morris, 1990), and many other reasons (e.g. Cantarelli *et al*., 2010).

First, predictions can be plotted and convolved with an empirical distribution in the evidence space of observed errors (from data quality or validation studies) to add uncertainty about predictions associated with a model error. Second, predictions can be shifted/adjusted to remove some of the uncertainty associated with the measurement protocol. In both of these cases, the structure of errors can be characterized as a distribution and transferred into the evidence space with arbitrary complexity. The thesis illustrates the requisite calculations to make these corrections with numerical examples from the case company.

The calculations performed during this thesis (Figure 16) propose the correction in the evidence space. The main reason for that is the ability to move distributions without exceeding 0-1 probability limitation. Second, the deviation from the “true value” is easier to spot – as it is represented as a distance from the true value on the X axis. The identified personalities from Section 6.3 thus have a formula that assesses the deviation.

The use of such a tool is seen as potentially very useful in practices that have good and large historical data sets. They can use the calculations to improve their future estimates by correcting the experts’ opinions for the patterns identified in the historical data. The tool has great potential for future research.

				Palaeocene							Triassic					
				Source	Charge & Timing	Reservoir	Seal	Trap	Total		Source	Charge & Timing	Reservoir	Seal	Trap	Total
	Play statistics	Base rate based on play statistics		0,95	0,9	0,9	0,8	0,8	0,49		0,9	0,7	0,6	0,8	0,5	0,15
	Mr Clever	Faithful representation of prospect		0,8	0,7	0,9	0,8	0,95	0,38		0,8	0,7	0,5	0,7	0,6	0,12
	Mr Wrong	No clue. Random.		0,9	0,8	1,0	0,9	0,7	0,43		0,3	0,9	0,9	1,0	0,5	0,13
	Little Miss Sunshine	Eternally optimistic		0,9	0,8	0,9	0,9	1,0	0,60		0,9	0,8	0,7	0,8	0,7	0,30
	Mr Grumpy	Eternally pessimistic		0,7	0,5	0,8	0,7	0,9	0,18		0,7	0,5	0,3	0,5	0,4	0,03
	Mr Brave	Polarized		0,6	0,5	0,9	0,8	1,0	0,25		0,7	0,7	0,4	0,6	0,6	0,09
	Little Miss Shy	Sitting on the play fence		0,9	0,8	0,9	0,8	0,9	0,45		0,8	0,7	0,5	0,7	0,6	0,13
	Mr Topsy Turvey	risk risk but not attributed correctly		1	0,6	0,9	0,5	1	0,27		1	0,6	0,5	0,5	0,7	0,11
	Mr Fussy	roughly right but lots of false precision		0,86	0,62	0,93	0,82	0,97	0,39		0,83	0,68	0,53	0,67	0,56	0,11
				0,78	0,69	0,91	0,81	0,90	0,35	0,35						
				0,79	0,70	0,91	0,82	0,90	0,38	0,37						
	Play			12,8	9,5	9,5	6,0	6,0			9,5	3,7	1,8	6,0	0,0	
Evidence	Mr Clever	No transformation		6,0	3,7	9,5	6,0	12,8			6,0	3,7	0,0	3,7	1,8	
8,3	Mr Wrong	Normally distributed about .5^1/5	6	8,0	5,8	16,1	12,3	3,5			-3,0	9,8	9,3	16,0	-0,1	
	Little Miss Sunshine	Mr Clever + bias	3	9,0	6,7	12,5	9,0	15,8			9,0	6,7	3,0	6,7	4,8	
	Mr Grumpy	Mr Clever - bias	3	3,0	0,7	6,5	3,0	9,8			3,0	0,7	-3,0	0,7	-1,2	
	Mr Brave	play + factor*(clever-play)	1,5	2,6	0,7	9,5	6,0	16,2			4,3	3,7	-0,9	2,5	2,6	
	Little Miss Shy	play + factor*(clever-play)	0,7	8,3	5,6	9,5	6,0	10,6			7,2	3,7	0,6	4,5	1,2	

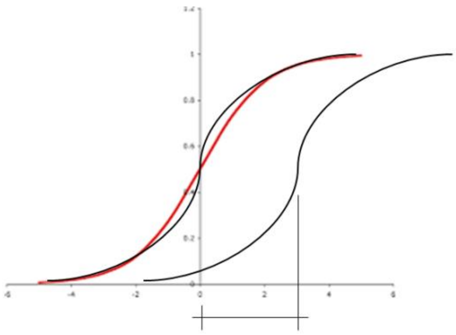


Figure 16 De-biasing of expert judgment in Evidence space.

6.5. Summary of the findings - Chapter conclusions

In this chapter the answer for research question 3.2 is provided. It focuses on the representation of background knowledge in risk and uncertainty assessment. Based on a case study, a set of methods from the second group of non-probabilistic approaches (semi-quantitative approaches) are applied to visualize uncertainty that surrounds data and results. In addition, calculations to quantify and correct biases in expert judgment in risk assessments are developed, as well as qualitative approaches to inform levels of trust of decision makers into risk quantifications.

In summary, the key insights obtained through the methods application were: The NUSAP tool can successfully represent background knowledge, it is easy to understand, and helps the experts to communicate the two different situations (Paleocene and Triassic Prospect), which was one of their main problems. Furthermore, it can be incorporated as an additional feature to their current practice, without affecting (substantially changing) their current process. The development of characters and inclusion of behavioral and organizational aspects was greatly appreciated by the practitioners. Risk Imaging allowed visualizing different stakeholders' views, which is important on the individual (different personalities and biases), group (financial, legal, procurement, etc. departments have different priorities thus different risk profiles), and external levels (the company, competitors, regulatory bodies). The tool helps facilitating the conversation about these differences. A great potential is seen in de-biasing expert judgment calculations, as it could provide valuable insights from the historical data. It is also a step forward toward decreasing the subjective impact of practitioners.

The categories, aspects, and steps developed in this case study can be updated, changed or added depending on a company's needs. Here, the thesis offers an example of the proposed approach (consisting of five separate steps). It can also be seen as a toolkit of methods that allow support for different needs (representing and communicating background knowledge, visualizing uncertainty surrounding results, better understanding of the behavioral aspects, different stakeholders' perspectives). A company and its practitioners can decide (also on a project level), how many steps they want to take – if deep-dives are needed.

The initial evaluation and feedback from practitioners states that some of the limitations are recognized as changing and adding more steps into one type of analysis may lead to additional changes in other processes and analyses. All these additional steps should be compatible with other processes and e.g. financial analyses for further aggregation inside the company – for portfolio and strategy levels. To properly integrate and add the new steps to all departments may be a rather long process. Nevertheless, having a handy way is seen as the first step in that direction.

Reflections

The reflections from the previous chapter are also relevant here (Section 5.6). Furthermore, when studying experts' role in risk management, a number of steps were taken to mitigate biases. When discussing experts there is a tendency to favor their duration of expertise, but time does not equate quality, so other measures such as the quality of expertise must be taken into account. Even when experts themselves try to describe their expertise or the rationale of their choices, the really highly skilled may not even themselves understand the rationale behind their expertise, as the ‘processing’ is hidden in their subconscious (Malcolm, 2005), which we are not able to document. There may be biases in the way we understand what an expert is and the way we judge the pedigree as well as other parameters included. It is also an important consideration if biases are stationary, or if and how much they will change over time, or due to events and new experiences.

The study places relevant focus on the qualitative and human aspects of the quantification process. There are potentially contextual limitations to the case study, as the methods were customized to the requirements of the company and real-life situations. Therefore, the exact methods may not apply in other contexts; they might need additional customization for a specific context. This realization is also in line with the topic of tailoring risk management (Chapter 8), where I point out a gap as well as a need to adequately choose and use methods. It is a limitation that most methods face; they must strive to satisfy contextual applicability and not be used ‘right off the shelf’.

7. Third group of methods: Exploring approaches for coping with deep uncertainty and introducing Robust Decision Making

"When you become comfortable with uncertainty, infinite possibilities open up"

- Eckhart Tolle -

Acknowledgement: The content of this chapter is based on the previously published article (Tegeltija et al., 2018a). The chapter quotes the article with only minor edits to harmonize the language with the remainder of the thesis. As the first author, I have initiated the research question, collaboration with a scholar from another University and collaboration with industry. I have carried out the empirical work and drafted the first version of the manuscript. Lastly, I have consolidated the comments from other authors and polished the final version of the manuscript.

Uncertainty assessment and management, and the associated decision making, are increasingly important in a variety of scientific fields. While uncertainty analysis has a long tradition, meeting sustainable development goals through decision making in long-term engineering system design demands the addressing of "deep uncertainty" (Walker, Lempert, & Kwakkel, 2013). Deep uncertainty characterizes situations where there is no agreement on exact causal structures, let alone probabilities. In this case, traditional, probability-based approaches cannot produce reliable results, as there is a lack of information and experts are unlikely to agree upon probabilities. Due to the nature of large-scale engineering systems, this chapter argues that methods to better cope with deep uncertainty can make a significant contribution to the management of engineering systems design. I introduce a set of methods that use computational experiments to analyze deep uncertainty and that have been successfully applied in other fields. I describe Robust Decision Making (RDM) as the most promising approach for addressing deep uncertainty challenges in engineering systems design. I then illustrate the difference between applying traditional risk management approaches and RDM through an example, and complement this investigation with findings from an interview with a company that puts RDM into practice. I conclude with a discussion on future research directions.

In the context of this thesis, this is the final step in investigating applications of the non-probabilistic methods (by answering research question 3.3). In comparison to the previous two chapters, this chapter is more literature-based and conceptual, as it does not have the same amount of empirical content. This is due to the nature of the methods, which, for an extensive assessment and proper case development, require long-term industry collaboration, including access to confidential information, which is outside the scope of this thesis. However, an initial empirical evaluation is here included and documented in the form of discussion (interviews) with the industry (see Section 7.5) to set the basis for further research and application of the methods introduced here.

This chapter discusses the need to go beyond probability-based tools in order to better address challenges in engineering systems design, and introduces the notion of deep uncertainty and its representations. It is structured into six parts. After the introduction in Section 7.1, the notion of deep uncertainty is explained in Section 7.2. An overview of the methods used to analyze deep uncertainty is provided in Section 7.3. I then describe one of the methods, RDM, in more detail in Section 7.4. The next section, Section 7.5, is a conceptual discussion, where I elaborate on RDM in contrast to traditional approaches in the context of engineering systems design challenges through an example of water resource management. Moreover, I interviewed the head of the risk management department in a large-scale engineering company (introduced as Company 1 in Chapter 2) on their experiences with RDM, life cycle engineering and deep uncertainty management. The final section, Section 7.6, presents conclusions and a discussion of future research directions.

7.1. Introduction

Over the last few decades, the life cycle engineering (LCE) research field has grown significantly. Achieving sustainable design and product development remains one of the central issues for the manufacturing industry (Takata & Umeda, 2007), but also for other domains where the LCE concept has been disseminated, that is, the food, building and textile industries (Alting & Legarth, 1995). Additionally, these industries have dealt with a paradigm shift from a product-centric to a service paradigm, which assists customers with accompanying services and systems for the products produced (Beuren *et al.*, 2013).

This transition to a service paradigm brought the need to enable a bigger-picture view and management practice that corresponds to such integrated systems and services. In particular, this is the case when the focus is on the sustainability, environmental impacts, and life cycle aspects of these solutions/systems. The current trend towards achieving desirable life cycle properties (i.e., "-ilities") of the systems can be carried out through LCE (Alting & Legarth, 1995). LCE enables a systemic perspective for achieving sustainability goals in engineering systems and their design. That is why I further discuss the introduced methods in the context of LCE, as the challenges raised are mostly related to the life cycle aspects of engineering systems.

Both researchers and practitioners have suggested that the development of LCE, and in particular life cycle assessment (LCA), should keep pace with the complex and changing product development systems (Chang, Lee, & Chen, 2014). LCA is an important tool for assessing the environmental impacts of product and service designs to support the achievement of sustainability. These changes lead to the increased importance of addressing uncertainty throughout the whole life cycle of a product or service. Uncertainty considerations are particularly relevant for the accuracy of LCA (Hellweg & Canals, 2014) and, therefore, research in that direction is of great significance for the field.

As discussed in Chapters 3 and 4, design and engineering activities often bring novelty, uniqueness, and first-of-a-kind solutions to an engineering problem (Gidel, Gautier, & Duchamp, 2005). The most important decision making situations in such cases are dominated by so-called "deep uncertainty": uncertainties for which experts do not agree upon models to describe interactions among a system's components, and subsequently do not agree upon corresponding probabilities and possible outcomes (Lempert, Popper, & Bankes, 2003). This leads to limited applicability of traditional risk and uncertainty management approaches and an increased need for developing novel approaches. While there is no consensus among researchers on a single approach for coping with deep uncertainty, there is an agreement about the need to model it differently. However, the tendency in practice is still to employ traditional, probability-based approaches. The increasing societal and business criticality of product development projects raises the need to explore more thoroughly the various fundamental approaches to describing and quantifying deep uncertainty as part of LCE and, correspondingly, overall engineering systems design.

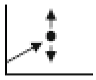
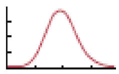
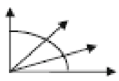
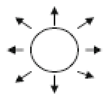
7.2. Deep uncertainty and its representations

It is important to distinguish between uncertainties that can be treated through probabilities and uncertainties that cannot. Different taxonomies and representations of uncertainty have been developed. An uncertainty matrix is proposed by Walker *et al.* (2003), which synthesizes various taxonomies, frameworks, and typologies of uncertainties from different fields. The taxonomy has been further extended by Kwakkel, Walker and Marchau (2010) – see Table 9. The goal of this synthesized overview is to support modelers in identifying uncertainties and communicating these uncertainties to decision makers. The typology of Walker *et al.* (2003) conceptualizes uncertainty as a three-dimensional concept. These three dimensions are 1) the level dimension, 2) the location dimension, and 3) the nature dimension. Of these, the level dimension tries to capture differences in the types of scales that are used in practice when assigning likelihood to events (Kwakkel, Walker, & Marchau 2010). Within this taxonomy, deep uncertainty is understood as Level 4 and Level 5. This understanding is broadly consistent with the work of Lempert, Popper, and Bankes (2003), who define deep uncertainty as "*the condition in which analysts do not know or the parties to a decision cannot agree upon 1) the appropriate models to describe interactions among a system's variables, 2) the probability distributions to represent uncertainty about key parameters in the models, and/or 3) how to value the desirability of alternative outcomes.*"

In their work, Walker, Lempert, and Kwakkel (2013) further explain and categorize each level of uncertainty. Most of the LCE problems faced by decision makers are characterized by higher levels of uncertainty. The following include several considerations involved in designing a bridge or a tunnel with a 100-year life span: estimating traffic intensity for the next hundred years, allowing the chosen design to adapt to the addition of any new installations and technologies that can/should be added to the system, estimating changes in the sea level, etc. As there is a wide range of outcomes for the alternatives that could take place, the question is how to best prepare for any combination of alternatives that may happen. The evolving, iterative, social and complex nature of LCE corresponds to a multiplicity of plausible futures, several variants for system models, a range of outcomes and associated weights or preferences regarding the various outcomes (corresponding to Level 4 or 5 uncertainty, as described in Table 9).

While many of the traditional analytical quantitative approaches are designed to deal with Level 1, Level 2 and Level 3 uncertainties (Paté-Cornell, 1996; Walker, Lempert, & Kwakkel, 2013), it has been proven that those methods face challenges when dealing with higher level uncertainty, that is, deep uncertainty (Walker, Lempert, & Kwakkel, 2013). It can be argued that deep uncertainty may sometimes be reduced by further research and information gathering. However, this may lead to additional and hidden costs and delays, thus making it infeasible. This leads to "real life" situations in LCE, where actions have to be taken now, that we know are based on incomplete information and have significant impact on following processes and outcomes. The thesis argues that deep uncertainty-based approaches can offer relevant support to these types of decision situations.

Table 9 Synthesized uncertainty matrix by Kwakkel, Walker and Marchau (2010) and the progressive transition of levels of uncertainty from complete certainty to complete ignorance by Walker, Lempert and Kwakkel (2013)

Location	**Level**				
	Level 1	**Level 2**	**Level 3**	**Level 4**	**Level 5**
Context	A clear enough future 	Alternate futures (with probabilities) 	Alternate futures with ranking 	A multi-plicity of plausible futures 	An unknown future 
System model	A single (determinis-tic) system model	A single (stochastic) system model	Several system models, one of which is most likely	Several system models, with different structures	Unknown system model; we know we don't know
System outcomes	A point estimate for each outcome	A confidence interval for each outcome	Several sets of point estimates, ranked according to their per-ceived likelihood	A known range of outcomes	Unknown outcomes; we know we don't know
Weights on outcomes	A single set of weights	Several sets of weights, with a probability attached to each set	Several sets of weights, ranked according to their per-ceived likelihood	A known range of weights	Unknown weights; we don't know we don't know

A range of traditional uncertainty and risk management methods has been applied to Level 4 and Level 5 problems. Group processes, such as the Delphi technique (Rowe & Wright, 1999), have helped large groups of experts to combine their expertise into narratives of the future. This can be understood as a "Level 4" method, where plausible future scenarios are developed without necessarily quantifying the associated uncertainties. In their work, Ferson and Ginzburg (1996) illustrate examples in risk analysis for which classical Monte Carlo methods yield incorrect answers when used to quantify higher levels of uncertainty. On the one hand, the development of Information Technology (IT) generated statistical and computer simulation modeling that allows the capturing of quantitative information about the extrapolation of current trends and the implications of new driving forces. Formal decision analysis can systematically assess the consequences of such information. Some more recently developed approaches, such as scenario planning, help individuals and groups to accept the fundamental uncertainty surrounding the long-term future and consider a range of potential paths, including those that may be inconvenient or disturbing for organizational, ideological, or political reasons (Schoemaker, 1995).

However, despite this rich legacy, one key aspect remains a problem. The traditional methods briefly outlined above face challenges when dealing with the long-term multiplicity of plausible futures, unknown causal structures, probabilities and difficulty in identifying preferred solutions. In the following section, I introduce a family of conceptually related approaches that are used to cope with such situations, that is, deep uncertainty.

7.3. A family of related conceptual approaches for coping with deep uncertainty

The deep uncertainty literature rests on three key concepts:

1) Exploratory modeling: In the face of deep uncertainty, one should explore the consequences of the various presently practically irreducible uncertainties for decision making (Lempert *et al.*, 2006; Weaver *et al.*, 2013). This exploration uses computational scenario-based techniques for the systematic exploration of a very large ensemble of plausible futures (Bankes, 2002; van Asselt and Rotmans, 2002; Bankes, Walker, & Kwakkel, 2013).

2) Adaptive planning: Decision robustness can be achieved through plans that can be adapted over time in response to how the future actually unfolds (Kwakkel, Walker, & Marchau 2010; Wilby & Dessai, 2010; Haasnoot, Kwakkel, & Walker, 2013).

3) Decision support: The aim of decision advice is to facilitate learning about a problem and potential courses of action, not to dictate the right solution. This entails a shift from *a priori* to *a posterior*i decision analysis (Tsoukiàs, 2008).

One method of decision making would be to determine the best predictive model and solve for the optimal uncertainty mitigation procedure. However, this method is fragile, depending on assumptions. In conditions of deep uncertainty it is better to seek, among the alternative decision options, those actions that are most robust – that achieve a given level of goodness across the multitude of models and assumptions consistent with known facts (Walker, Hassnoot, & Kwakkel, 2013). From an analyst's and a manager's point of view, this means that the aim is no longer to answer the question of "What will happen?" but rather "Given the agreement that one cannot predict everything, which actions available today are likely to best serve me in the future and keep my options open?"

A family of approaches exists for dealing with deep uncertainty:

Assumption-Based Planning was developed at the RAND Corporation almost 30 years ago as a tool for improving the adaptability and robustness of an existing policy/plan/design (Dewar *et al.*, 1993).

Robust Decision Making (RDM) uses multiple views of the future to iteratively stress test one or more candidate strategies over many scenarios, and refine the strategies in light of this (Walker, Haasnoot, & Kwakkel, 2013).

Adaptive Policymaking was specifically developed to support the implementation of long-term plans despite the presence of uncertainties (Haasnoot *et al.*, 2012).

Adaptation Tipping Points and Adaptation Pathways are both approaches that consider the timing of actions and were developed for water management (Haasnoot *et al.*, 2012).

Dynamic Adaptive Policy Pathways combines the work on Adaptive Policymaking with the work on Adaptation Tipping Points and Adaptation Pathways (Haasnoot, Kwakkel, & Walker, 2013).

RDM is a promising approach to address the challenges in LCE, as it offers a structured method for planning under deep uncertainty and it is the best-known deep uncertainty approach. Simulation models are used to evaluate different designs over a wide variety of different conditions. Next, using scenario discovery (Bryant & Lempert, 2010; Kwakkel & Jaxa-Rozen, 2016), the analyst can discover conditions under which designs fail. In light of this, designs can be improved. RDM, together with a set of model-based tools, can support decision making under deep uncertainty in LCE by providing recommendations that enable managers to choose and improve a design that produces satisfying outcomes across a broad range of possible future conditions.

7.4. Robust Decision Making to manage deep uncertainty

RDM has been developed over the last 30 years, primarily by researchers related to the RAND Corporation (Dewar *et al.*, 1993). The RDM framework helps decision makers to use multiple views of the future in support of a thorough investigation of modeling results that, according to Lempert, Popper, and Bankes (2003) and Groves and Lempert (2007), helps to identify a design that:

1) is robust (i.e., it performs "well enough" across a broad range of plausible futures, but may not perform optimally in any single future; it also has little regret),
2) avoids most situations in which the design would fail to meet its goals, and
3) makes clear the remaining vulnerabilities (i.e., conditions under which the design would fail to meet its goals).

According to Walker, Haasnoot, and Kwakkel (2013), RDM includes the following five steps:

1) Scoping – determine the scope of the analysis by identifying exogenous uncertainties, modeling options, key relationships, and performance metrics; construct a simulation model that relates actions to consequences.
2) Simulation – identify a candidate model to evaluate and run it against an ensemble of scenarios.
3) Scenario discovery – identify vulnerabilities of the candidate model (i.e., what combinations of exogenous uncertainties, and in which ranges, cause the design to fail to meet the goals?).
4) Adaptation – identify hedging actions (modifying existing models or defining new ones) to address these vulnerabilities. Repeat Steps 2 and 3 for additional candidate models.

5) Display – plot expected outcomes of all models over probabilities of vulnerable scenarios, and choose the most robust option for implementation.

Over the years, RDM has been employed to provide support in strategic planning problems in a variety of fields, including climate change (Lempert, Schlesinger, & Bankes, 1996), complex systems (Lempert, 2002), economic policy (Seong, Popper, & Zheng, 2005), and flood and water risk management (Lempert, Sriver, & Keller, 2012; Herman *et al.*, 2014).

7.5. Discussion of uncertainty quantification in LCE

Given the importance of decision support in LCE, it is essential to explore approaches for dealing with deep uncertainty. Some of the non-probabilistic methods introduced by Aven *et al.* (2014) try to resolve the problem within the *"predict and act"* paradigm in risk and uncertainty management, by introducing methods that are less reliant on probabilistic data. This set of methods corresponds more to the improvement of LCA by allowing better, more accurate estimates. In addition, these methods allow the experts to provide information in data formats that they feel comfortable with (points, intervals, and ratios as well as their combination), depending on the confidence level. Some studies further enhance the usage of non-probabilistic methods through comparative analyses with probabilistic approaches (André & Lopes, 2012).

The approaches discussed in this chapter, on the other hand, drop the *"predict and act"* thinking altogether and introduce a *"monitor and adapt"* paradigm to replace it. These approaches change modeling more fundamentally and have produced reliable results in fields such as water management (Herman *et al.*, 2014), climate change (Walker, Hassnoot, & Kwakkel, 2013), and policy-related research (Hamarat, Kwakkel, & Pruyt, 2013). Once crucial decisions under deep uncertainty have been made and additional information and knowledge have been collected, traditional approaches can be employed to continue the uncertainty management in LCE.

Arguably, the challenges that practitioners face in other fields are in many ways close to the ones that are often seen in LCE. For instance, such situations are characterized by a large number of stakeholders, weak available information, significant impact on the further process and a notable societal impact. I focus on the uncertainty quantification and how these methods work and the kinds of insights they produce in the context of LCE through the lens of RDM.

Traditionally in engineering, when dealing with a lack of hard data, uncertainty analysis is based on expert judgment. Experts are asked to provide precise estimates on different activities and these estimates are the input for probabilistic analyses (described in more detail in Chapter 6). The models used in these analyses need to have all the activities and correlations predetermined upfront. For a number of reasons these correlations are not always obvious or visible to the modelers.

With the latest developments in the manufacturing industry it is often not feasible to find solid ground for estimating probabilities. Moreover, subjectivity in expert judgment remains a challenge (Ellsberg, 1961; Cooke, 1991). Furthermore, the results do not reflect the availability and quality of background information, or a number of assumptions behind the calculations.

The current trend towards achieving desirable life cycle properties (i.e., "-ilities") further challenges the applicability of deterministic models (de Weck, Ross, & Rhodes, 2010). As stated by Ricci *et al.* (2014), a survivable, flexible, or evolvable system should be able to sustain value delivery over time by responding to exogenous changes in the operational environment. To achieve this, we need to allow adaptivity and imprecision throughout the life cycle, and explicitly design for this.

One way to do this is to employ RDM in LCE: a large number of futures are generated based on performance criteria. First, RDM is used to sample a wider range of futures, which are subsequently assessed to see whether they are dire, benign, or opportunistic. Second, it offers a holistic assessment of the performance of generated options over the wider range of futures. The idea is that a design solution should work satisfactorily over a broad range of these possible futures. RDM also identifies what combinations of uncertain future stresses lead to system vulnerabilities through "scenario discovery" (Matrosov, Woods, & Harou, 2013; Thissen, Kwakkel, Mens, 2017). The five-step RDM process (see above) is then repeated iteratively until a suitably robust solution is found. RDM aims to assist in the development of a solution whose performance is good enough over a wide range of futures (i.e., it is robust) rather than an optimal solution for a single specific future.

An example is presented by Matrosov, Woods, and Harou (2013), in which the authors applied RDM for a water management problem when statistical distributions of future events were poorly known, and they followed the five described steps. In terms of LCE, RDM differs from traditional approaches, for instance Scenario Planning, by sampling a larger number of possible scenarios that are further evaluated. In the mentioned example, they generated 311080 possible simulations/scenarios, whereas Scenario Planning typically involves the evaluation of only a few identified scenarios. This provides more thorough analyses (e.g., minimizing life cycle regret) that improve the quality of decision making in LCE, which impacts the quality of products and systems produced.

It is worth noting that such analyses are now feasible given the advances in computational methods. Moreover, RDM is complementary to other approaches (Matrosov, Woods, & Harou, 2013) that provide additional information to the decision makers when managing deep uncertainty.

This kind of modeling does not require unjustified assumptions and provides a structured framework for the iterative refinement of future plans.

A brief discussion with practitioners on RDM in the LCE context

The case company is a large Danish company with extensive experience in designing and managing large engineering projects, including assessing cost and operational life cycle properties of complex, long life cycle infrastructure systems. I interviewed the head of the risk management department on his experience with RDM. The interview was conducted after the interviews presented in Section 8.4.2.1. Details regarding their risk management team and process can be found there.

In their practice they recognized the need to look for alternative approaches that can reliably manage deep uncertainty. They analyzed different options and decided to use RDM on one of their projects.

Several limitations were raised regarding RDM: first, in their experience, it is still open to debate which design is the best choice when simulating the system's life cycle properties. RDM does not provide a "simple" answer and the analysis results must be further interpreted in the decision making process. Second, RDM-based assessments of, for instance, 100-year life cycle system properties are based on current data, even if they are analyzed and interpreted differently. Third, there are projects where the use of RDM is not justified, that is, projects involving only the first three levels of uncertainty, where similar engineering solutions exist, where uncertainties of mostly stochastic nature are present, and where the lifetime is fairly short. Clearer guidance is needed about when RDM effectively adds value to LCE decisions, and when it does not.

The case company agrees that there are LCE tasks in projects where higher levels of uncertainty are present and that the traditional approaches currently employed only offer modeling capabilities corresponding to the first three levels of uncertainty. These cases are where the life cycle performances of a one-of-a-kind bridge with a 100-year lifetime have to be assessed, and they are dealing with first-of-a-kind solutions for engineering problems, novel technologies, new locations, more stakeholders and significantly longer lifetimes. Often, as in the water management example, traditional modeling approaches require them to make "precise" predictions based on the limited information available. The approaches introduced in this chapter can significantly support uncertainty management through their provision of more thorough analyses of possible alternative futures.

Reflections

The theoretical sampling was also appropriate in this case. The chosen company is highly specialized in risk management and runs projects that often involve deep uncertainties. More interviews would need to be conducted to entirely comprehend their modeling capacities with RDM, but informal discussions confirmed their remarkable computational skills. In this way, the work presented here constitutes a solid contribution and opens topics up for future research (see Section 10.4).

Despite all these positive aspects, some limitations did emerge. During the PhD project, no opportunity arose to check the company's actual modeling. For instance, the possibility of developing a synthetic case with them, on which more detailed discussions provided key insights, would have been very relevant and would have allowed findings to emerge that could have opened up better possibilities for triangulation.

7.6. Conclusions

There are a number of methods on hand to deal with uncertainty, so it is important to select the method best suited to the particular uncertainty in question. It would be desirable to have a single method capable of quantifying all types of uncertainty. Traditionally, one candidate for this task is probability theory.

As previously mentioned, engineering systems design risk and uncertainty management practice have so far relied heavily on probability-based methods when treating uncertainty. I acknowledge the great merit of probability-based methods, but I also point out limitations that lead to the need for frameworks beyond probability. This has triggered the development of alternative approaches in other fields. The methods introduced in this chapter rely on the idea that imprecision and adaptivity correspond better to the weak information available in LCE, as one approach to ensuring the desirable life cycle properties (i.e., "-ilities") of engineering systems.

The contribution of this chapter is in the *"monitor and adapt"* paradigm, which is suggested for application in LCE to improve risk and uncertainty management practices. I raise the importance of distinguishing deep uncertainty from uncertainty due to variance, and point out the complexities that it brings to decision making. Given the evident need to go beyond probabilities when dealing with deep uncertainty, I provide insights into the contributions offered by novel approaches and where they have been used. These approaches need further adaptation to the conditions of LCE.

I further introduce RDM as a specific method for coping with deep uncertainty in LCE. Nevertheless, in order to demonstrate the full benefit of RDM for LCE, real case studies are needed, as well as illustrative examples/synthetic cases. Future research in that direction would not only allow better treatment of deep uncertainty, but it would also broaden our understanding of decision making support in such situations. In my view, it is essential for the field to consider these relatively recently developed methods. Of particular significance is their application potential when looking for more appropriate solutions to analyzing and quantifying uncertainty in LCE and, correspondingly, overall engineering systems design.

8. Tailoring risk management: Risk and uncertainty quantification as part of the overall risk management process

"It's always the small pieces that make the big picture"

-Unknown-

Acknowledgement: The content of this chapter is based on the previously published article (Tegeltija et al., 2018b). The chapter quotes the article with only minor edits to harmonize the language with the remainder of the thesis. As the first author, I have initiated the research question, collaboration with a scholar from another University, collaboration with industry and conceptual development of the proposed approach. I have carried out the empirical work and drafted the first version of the manuscript. Lastly, I have consolidated the comments from other authors and polished the final version of the manuscript.

While risk quantification research has grown over the last few decades, a limited number of studies have addressed the overall process integration of these approaches in engineering systems design risk management, that is, tailoring risk management methods to the specific requirements and conditions of a design project. This chapter argues that the choice of risk quantification method has strong implications for several aspects of the risk management process, as well as the integration of risk management results into decision making processes. I investigate current risk management maturity models and suggest an expansion to accommodate the requirements originating from the choice of quantification method, as well as informing the choice of quantification method based on other process parameters. This is validated through three case companies. Additionally, three more companies were approached to provide their feedback on the developed approach.

In the context of this thesis, this step is important. Up to this point, the thesis has investigated advanced risk and uncertainty quantification methods (introduced under the non-probabilistic framework). This chapter concerns itself with the broader question of when and how to integrate these methods into an overall risk management process (answering research question 4).

The chapter is structured as follows: Section 8.1 provides a short introduction to the motivation and need for tailoring risk management in engineering systems design through theoretical and empirical considerations related to the current practice. Section 8.2 further describes the specifics of risk management in the field and reviews risk management maturity models. Section 8.3 describes the conceptual development of the risk management tailoring approach, depending on the risk management maturity. In Section 8.4 the approach is illustrated through case companies from different sectors and the empirical work is described. In Section 8.5 I discuss different risk management tailoring approaches, and lastly, in Section 8.6, I provide concluding remarks about the presented research and highlight the importance of the proper integration of risk quantification in engineering systems design to enhance its full potential.

8.1. Introduction

The positive impact of risk management activities on design and product development outcomes has been confirmed multiple times by different scholars (Wieland & Wallenburg, 2012), but the need for risk management differs between different organizations (Oehmen *et al.*, 2014). While some organizations have identified the requirement for rigorous and very strict organization-wide risk management processes in all aspects of their businesses, others simply require some basic understanding of risk management practice. Different project types and the associated risks have to be managed according to the context − one size does not fit all − and the enduring need to tailor the wide range of activities and approaches in the field is confirmed, for example by recent reviews (Kaplan & Mikes, 2012; Škec *et al.*, 2014; Herrmann *et al.*, 2018).

One part of the overall risk management process that requires good integration is risk and uncertainty quantification – such as the methods developed and discussed in the previous chapters of this thesis. Organizations wishing to implement a formal quantification approach, or to improve their practices, need a benchmark against which to review their processes. In this regard, although a number of risk management maturity frameworks are available in the literature, they lack a focus on quantification methods and their impact on and implications for the overall design risk management process. This chapter seeks to address this gap through a proposed tailoring approach, based on maturity grids, that allows a two-fold tailoring: firstly, tailoring the design risk management process to a chosen risk and uncertainty quantification approach, and secondly, tailoring risk and uncertainty quantification options to the capabilities of the overall design risk management process. This chapter will introduce the reader to the importance of maturity grids in benchmarking and as a strategy for improvement, suggesting five categories that will help practitioners choose their risk quantification method: 1) understanding of the needs, 2) method sophistication for risk quantification, 3) quality of data, 4) awareness regarding risk in organizational culture, and 5) impact of risk assessments in decision making. Within these categories, improvements are made to the overall risk management processes, which will ultimately assist companies in systematically planning their desired advancement in practice.

8.2. Risk and uncertainty in design

8.2.1. Risk management in design

Engineering systems design is vulnerable to various risks, which can emerge during the design process. Some even argue that the design process can be perceived as a process of uncertainty and risk management (Gericke, 2011), and suggest that a key attribute of a designer is the ability to manage uncertainty (Cross, 2011). Standardized and structured design processes, accompanied by the use of appropriate methods and tools (such as lean, six sigma and total quality management) may reduce uncertainty and risk in general, but nevertheless a considerable amount of residual uncertainty remains, which needs to be addressed and treated in design processes. Management of risk in these processes has received attention from researchers in engineering design (Lough, Stone, & Tumer, 2009), and related studies have been carried out in project management (Raz & Michael, 2001), and safety-related risk management (Paté-Cornell, 1996).

Despite the wide study of risk management in engineering systems design, only a few authors have tackled the issue of systematization and classification of risk management methods, and especially the need for formulation of recommendations with respect to method application and the associated tailoring of the overall risk management process. The application of risk management requires familiarity with methods, appropriately trained employees and an understanding of context, and if any of the above-mentioned aspects is not implemented correctly, the value that risk management brings to design can decrease significantly. For these reasons, maturity models have been introduced as one approach to guiding organizations in their risk management implementation and benchmarking themselves against best practice (Maier, Moultrie, & Clarkson, 2012).

8.2.2. Risk management maturity models

Maturity-based assessments, for example, in the form of maturity grids or models, are a structured approach to exploring how well the behaviors and practices of an organization are adapted to delivering required outcomes, usually expressed as a series of structured levels presented in matrix form. For a review of existing models see, for example, Maier, Moultrie, and Clarkson (2012). The underlying idea behind maturity-based assessments is that they provide a framework that seeks to capture "good practice" in order to guide and structure both assessment and improvement in capability. The authors of these models begin with the underlying assumption that there is a link between the higher levels of maturity and improved performance in the (relevant) organizational capabilities. Organizations advance through a series of stages or levels of maturity, with levels often represented as ranging from initial, to repeatable, defined, managed, and optimized. While the underlying rationale for the levels may differ (Maier, Moultrie, & Clarkson, 2012), the levels often describe an evolutionary path ranging, for example, from ad hoc, chaotic processes or capabilities to mature, disciplined processes and, in this case, defining the degree to which a process is institutionalized and effective. Stepping through the levels can be seen as representing progress towards an optimum capability. A prominent example of such a maturity model is the Software Engineering Institute's Capability Maturity Model Integration (CMMI) (Humphrey, 1988). The approach has been tailored, modified and further developed for various applications in different domains, including the organizational project management maturity model (OPM3) program of the Project Management Institute (Pennypacker & Grant, 2003), knowledge management (Paulzen *et al.*, 2002) and innovation (Chiesa, Coughlan, & Web, 1996). But while maturity models may share a common structure, their content differs, and for this reason maturity models are very often developed anew. A review of existing models and guidance for the development of new models is given by Maier, Moultrie, and Clarkson (2012).

In terms of risk management, a maturity model was first introduced by Hillson (1997). This was followed by the PMI's RISKSIG extension of the model with new criteria and a further model, with a slight variation, was developed for complex product systems projects (Ren & Yeo, 2004). Table 10 shows the PMI RISKSIG's maturity levels.

Although a good basis for evaluation, current risk management maturity models have some limitations. The underlying assumption of many maturity models is "the higher the better." However, different organizations have different risk management needs, and achieving higher levels of risk management maturity does not necessarily imply a better "fit" of risk management to the organization's requirements. In this thesis, the extension of the model based on the proposed tailoring approach allows companies to engage in a discussion around the maturity model to find and agree on the most adequate risk quantification approach in their case. Furthermore, previous risk management maturity models do not have a strong method focus. Finally, a range of models is available, but all of them neglect the impact of a chosen method on the overall process; we need to be more explicit about the selection and application of the methods.

Table 10 PMI RISKSIG risk management maturity levels (2002)

Attribute	Level 1 (Ad hoc)	Level 2 (Initial)	Level 3 (Repeatable)	Level 4 (Managed)
Definition	Unaware of the need for management of uncertainties	Experimenting with RM through a small number of individuals	Management of uncertainty built into all organizational processes	Risk-aware culture with proactive approach to risk management
Culture	No risk awareness	RM used only on selected projects	Accepted policy for RM	Top-down commitment to RM, leadership by example
Process	No formal process	No generic formal process	Generic processes applied to most projects	Risk-based organizational processes
Experience	No understanding of risk principles of practice	Limited to individuals with little or no formal training	In-house core of expertise	All staff risk aware and able to use basic risk skills
Application	No structured application	Inconsistent application of resources	Routine and consistent application to all projects	Risk ideas applied to all activities

8.3. Conceptual development of the risk management tailoring approach depending on the risk management maturity level

The wide diversity in engineering systems designs and the uncertainty that arises during a design process has led to the development of a number of risk management approaches. To support key phases of risk assessment, including risk identification, analysis and evaluation, different methods and tools have emerged. Some are qualitative, as they mostly serve for risk identification and for when the information is not very easily quantifiable, such as brainstorming, checklists or the Delphi method. Other approaches are semi-quantitative, such as interviewing, risk mapping or the NUSAP tool (Brocéliande team, 2015), and provide quantitative results accompanied by qualitative, descriptive information. Monte Carlo simulations, sensitivity analysis, Bayesian networks and other probability-based approaches provide quantitative uncertainty modeling (Cagliano, Grimaldi, & Rafele, 2014).

I refer to all of the above approaches as "quantification approaches" since organizations with lower levels of risk management maturity only need to identify risks and prioritize them as the first steps towards reaching higher levels. When feasible, companies with higher levels of risk management maturity aim to employ purely quantitative approaches that can vary in their level of sophistication − in terms of mathematical complexity and data requirements (Paté-Cornell, 1996).

The literature is rich in methods, tools and conceptual frameworks for risk quantification. However, scholars have reported limitations and pitfalls in terms of both their methodological foundation and their application. For instance, the probability-based approaches to risk and uncertainty analysis, as those most commonly applied, can be challenged under the frequently found conditions of limited or poor knowledge, in which case the information available does not provide a strong basis for a specific probability assignment (Walley, 1991; Flage *et al.*, 2014). In such cases, precision in probabilistic results may lead to a false degree of certainty (Beer, Ferson, & Kreinovich, 2013). Furthermore, some of the limitations of the methods relate to the fact that correlations among risks are often not modeled and may lead to serious consequences, if not taken into account (Kujawski & Angelis, 2009). Subjectivity in risk assessments is also an issue (Hubbard, 2009). The quality of data used in the analyses has strong implications for the reliability of the results, and this is not reflected in the current approaches. Risk analyses often involve a number of assumptions that, if not presented to decision makers, may lead to false directions (Aven *et al.*, 2014).

As evidence of the low application of quantitative risk methods, Crossland, Williams and McMahon (2003) documented the fact that relatively few engineering systems design companies make use of such models in their risk management practices. They demonstrate the wide applicability of such approaches to engineering systems design, describing three different quantitative modeling approaches and illustrating both the simplicity of the approaches and the benefits of their usage.

The limitations of the current approaches and the gap between practice and research has led recent research to be focused on two research themes. The first is research into novel (more advanced) approaches that will bridge the existing limitations (Walker, Hassnoot, & Kwakkel, 2013; Flage *et al.*, 2014). For instance, some propose uncertainty modeling (i.e., imprecise probabilities) that can be used to explicitly express the precision with which something is known (Aughenbaugh & Paredis, 2005).

The second theme is overviews of existing models, and clarification of both the advantages and disadvantages of their usage is increasingly attracting attention. Classifications of risk management techniques are available in Cagliano, Grimaldi, and Rafele (2014); Raz and Hillson (2005); Dikmen, Birgonul, and Arikan (2004); and Marle and Gidel (2012). To support advancements in practice, it is important to clarify and be transparent about these limitations and disadvantages, and propose to the practitioners ways to overcome these challenges, both when choosing a method and when looking for ways to improve it.

To overcome some of these limitations and enable companies to knowledgably and systematically choose and plan their risk quantification, I propose to extend current risk management maturity models with quantification criteria, building on the work of Crossland, Williams and McMahon (2003); Grubisic, Gidel and Ogliari (2011); and Škec *et al.* (2014). I derived the five categories from the literature review and our empirical work, and iteratively developed this tailoring framework with three engineering companies. The purpose is to benchmark risk management quantification processes in the companies and adapt them to their needs. The framework itself also serves to codify boundary objects for organizational learning about risk management, thereby allowing organizations to understand where specifically to improve.

In particular, I propose a risk management tailoring approach that includes the five categories shown in Figure 17 and described below. The five categories were developed to support all the steps of the entire risk management process. Starting with the method sophistication and quality of data arising from the above-mentioned literature (see also Aven & Zio, 2011), I included three more categories (understanding of the needs, awareness regarding risk in organizational culture, impact of risk assessments in decision making) to incorporate the case companies' registered necessities and challenges in practice and experience.

Category 1: Understanding of the needs

To professionally approach risk and uncertainty in engineering systems design, an organization should be able to understand its needs, and those of its stakeholders, and the necessary approach to this will depend on the organizational structure, the applicable processes and the types and sizes of projects. The understanding of the concepts of risk and uncertainty is important for the ability to manage risk. The nature and type of uncertainty determine in part what kinds of methods are applicable, and thus a heightened level of understanding of uncertainty enables more mature risk management.

Category 2: Method sophistication for risk quantification

Higher accuracy of estimates enables better decision making support. Given their design challenges, some organizations may only need approaches that allow the identification of risks. Others may face challenges that require in-depth analysis. The level of sophistication of analysis will depend substantially on the method chosen for the analysis. Any limitations of the approach should be reported and communicated to decision makers. To improve their quantification, besides choosing a more sophisticated method, practitioners also need to synchronize advancements with other categories to ensure the greatest benefits of their risk management.

Category 3: Quality of data

The quality and availability of data will impact the results, as well as the number of assumptions supporting the analysis. In some cases, it is feasible to spend resources on acquiring high-quality data. In others, we need to proceed with the engineering systems design (often due to time pressures) and be aware of the arbitrariness in the quality of data we use and the number of assumptions we make prior to the analysis of choice. In the absence of that kind of transparency (achieved, for example, through visualization tools), central pitfalls may be overlooked. The quality of data should correspond to the method, as using a more sophisticated method on a low quality of data arguably does not add desired value.

Category 4: Awareness regarding risk in organizational culture

It is of great importance to build awareness of risk management processes, activities, value creation and impact for all employees across the different levels of an organization's hierarchy. To properly support decision making, decision makers need to be aware of its value and other employees need to be informed about why it is important that they provide certain information, attend associated meetings, and why the whole process deserves attention. Communication and (professional) language can vary, even within organizations. While some employees may have an educational background that corresponds to risk management needs, the way they inform and interact with others in the company needs to be adapted to correspond to their knowledge base.

Category 5: Impact of risk assessments in decision making

Employees may not appreciate the analysis and may have too little trust in the results to base decisions on them. Some of the complex mathematical calculations may be challenging for managers to comprehend properly, which may lead to them being neglected. Furthermore, the way the responses are planned and handled needs to be synchronized with the overall engineering systems design.

These categories, the associated maturity levels and a mapping of the categories to the ISO 31000 process, are shown in Figure 17. The proposed approach consists of the iteration of the following steps, inspired by the process outlined in Figure 17: 1) identifying and articulating the needs, 2) analyzing the current state of the risk management in the organization and identifying existing levels of maturity, 3) re-evaluating the needs to match the desired levels of maturity, and finally, 4) developing individual recommendations in order to achieve the desired practice according to specific cases.

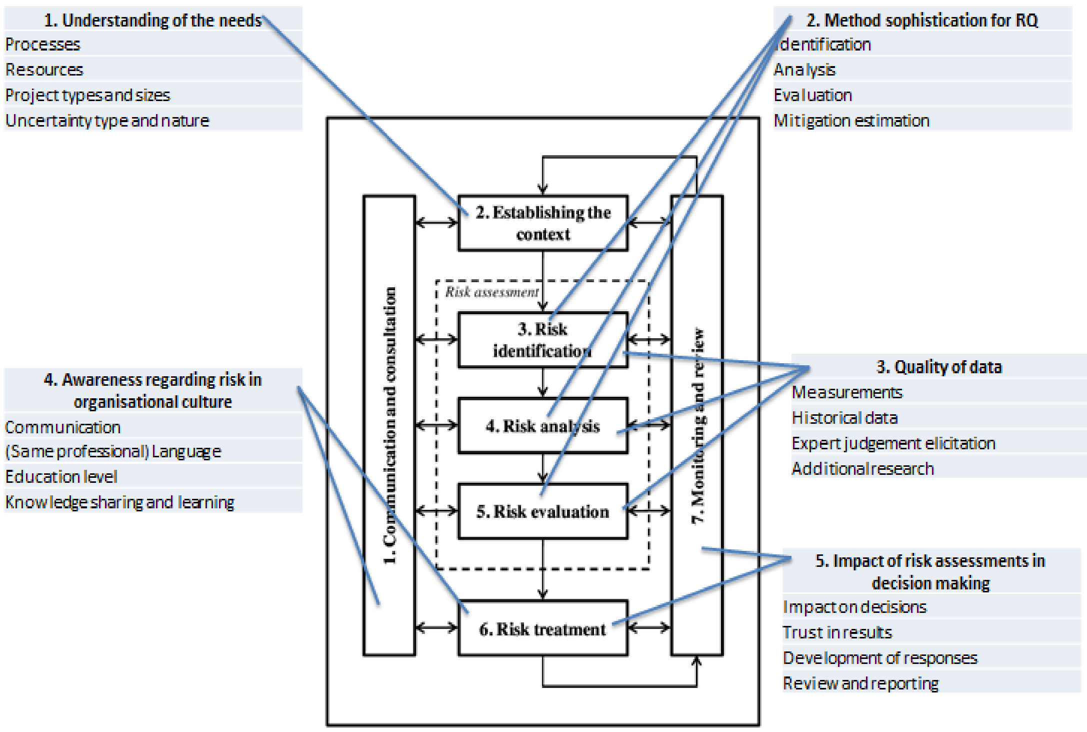


Figure 17 Relationship of maturity categories and the ISO 31000 risk management process (based on Tegeltija *et al.*, 2018b).

8.4. Case companies: Validation of proposed tailoring framework for risk management in design

8.4.1. Research method

In order to examine the applicability of the proposed approach for tailoring risk management in engineering systems design based on the extension of the risk management maturity model, I approached six companies to obtain feedback. As the proposed approach aims to provide support for a broader range of different practices, I selected companies with completely different application domains (areas of design work) with different risk management procedures.

I conducted semi-structured interviews with employees (as described in detail in the following sub-chapters) in order to understand the organizations' contexts. This also allowed me to modify, extend and enrich the initial set of questions and, when needed, to organize follow-up interviews. By doing so, I deepened our understanding of the risk management practices that were encountered. The guide to interview questions and related grouping of codes is available in Appendix 2.

The first set of interviews included discussions with the interviewees on: 1) their area of work and the design challenges they face, in order to understand their specific engineering systems design risk and uncertainty profile; 2) their risk management process, how it relates to their overall organizational structure, how it is designed and compares with risk management standards and maturity models, and 3) the different quantification techniques they use and their relationship to the five categories introduced previously. I then analyzed and coded the collected data in ATLAS.ti, according to the proposed customization approach, as pilot applications, and developed recommendations for process adaptations. This included follow-up phone calls where clarification was necessary. The results of these pilot applications of the proposed customization approach were presented in a second set of interviews, and the interviewees were invited to comment on their possible application, usability and contribution, as well as any limitations and challenges they might foresee. I used Support Evaluation (Blessing & Chakrabarti, 2009) as part of the continuous testing of the design support.

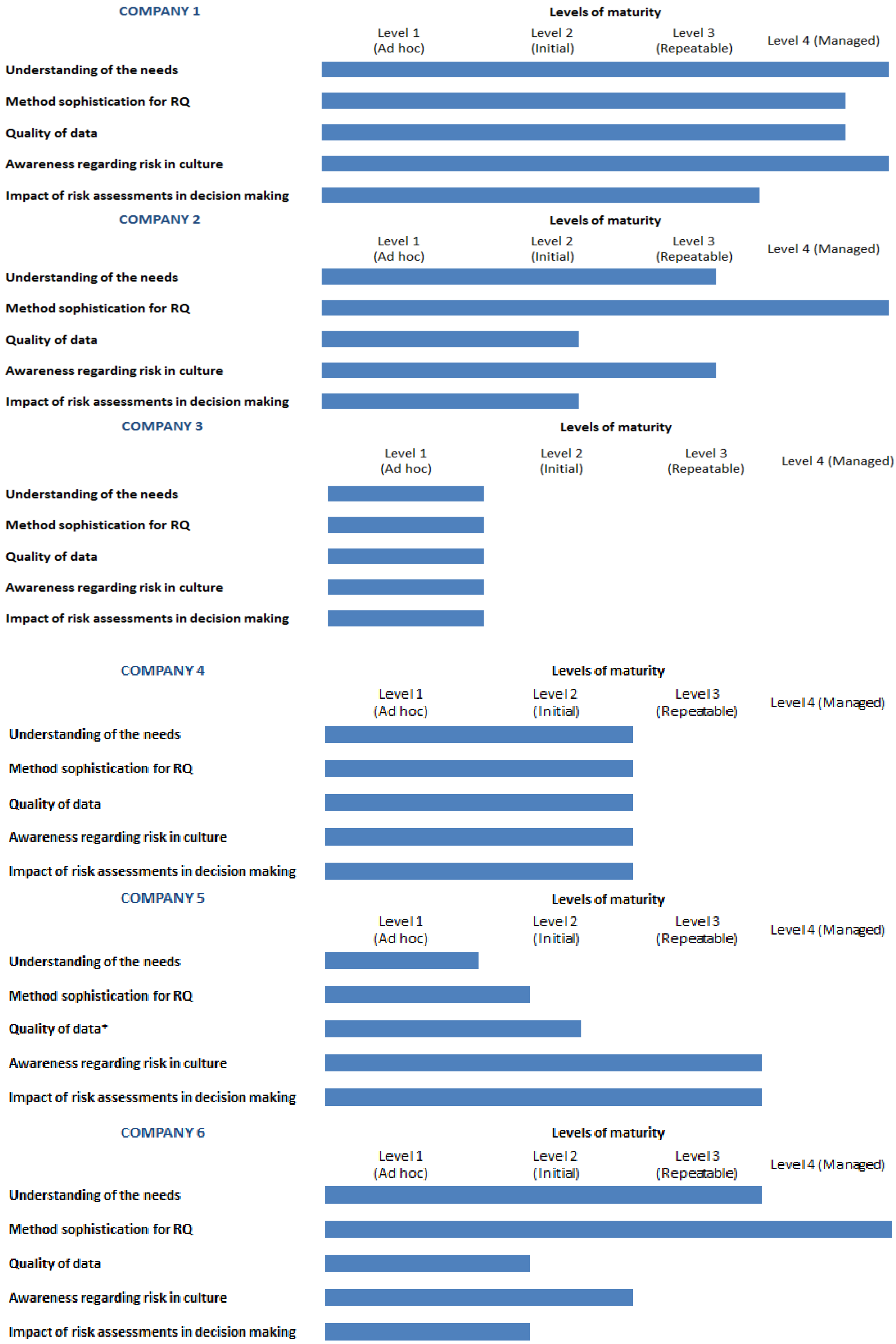


Figure 18 Overview of case companies' levels of maturity (1-6) (extended from Tegeltija *et al.*, 2018b).

8.4.2. Companies involved and their risk management context

A summary of the risk management maturity of the case companies is shown in Figure 18. Given the information collected from the practitioners, and based on coding with respect to each of the five categories, I evaluated companies on the introduced Level 1 - Level 4 maturity scale. There is one company with a very ad hoc profile (Company 3) and one that has established some initial risk management practice (Company 4). Another company has an almost completely managed profile (highly structured approach in Company 1). As a public entity, Company 5 has a completely different profile than all the other cases. Finally, Company 2 and Company 6 have profiles that explicitly illustrate the need for the tailoring being addressed by this thesis, as their current quantifications need improvements in terms of the other four categories.

8.4.2.1. Company 1: Design of large-scale engineering systems

Area of work and design challenges: The first case relates to a large Danish company with extensive experience in designing and managing large-scale engineering projects, such as long life cycle infrastructure systems. They have projects of different sizes (from megaprojects to small design solutions), for instance, they are designing for first-of-a-kind engineering projects in which they face severe uncertainties, but they also help small practices achieve their goals. Their risk management approach needs to provide support for their whole spectrum of design activities and ensure proper and timely responses and monitoring.

Risk management process and link to other organizational processes: The company is a large, highly structured organization comprising many departments. I interviewed the head of the risk management department twice. The department was established to oversee risk management for the company's projects as well as to provide consultancy services to other companies. The department consists of highly specialized risk (and safety) experts, working on different aspects of the risk management process during the design and construction project phases. They all have an appropriate educational background, are familiar with the applicable standards (ISO, 2009) and practitioners' guidelines (PMI, 2008) as well as following the relevant advances and courses in the field. Furthermore, their project, program and portfolio managers are familiar with and rely on the department's results, and other employees are aware that such practice exists in the organization.

Given the broad scope of design activities and the number of projects carried out, there was a need to engage all sorts of methods (from qualitative, through semi-quantitative to quantitative), but also sometimes to employ some of the most sophisticated quantification approaches. For each of the projects they would need to select the most appropriate method, and after the analyses had been carried out, present and communicate the results to the managers. When necessary, special interest and focus would be placed on gathering data. Depending on the specific design and associated uncertainties, they use risk registers and historical data, and they organize workshops and/or hire experts/consultants for particular issues (for instance, when estimating the number of railway passengers in the next 50 years).

Proposed design risk customization and evaluation: Even though the company has already reached a high level of risk management maturity, they still seek frameworks for further improvement and carry out constant re-evaluation. They show a high level of understanding of the impact and the importance of the choice of risk method and its proper usage, which is why my recommendation was to examine the proposed approach in terms of the consultancy services they provide. The feedback to the proposed recommendation was that, from their perspective, the tailoring approach allows them to systematically, and in a structured way, explain and argue why they propose specific risk actions (and even more specifically why they use a certain risk method). The risk management maturity of their clients varies a lot and they would face challenges in adjusting the recommendations and communication to the clients' level. The selected maturity model and proposed extension are seen as a clear, easily understandable and manageable approach for different clients. While until now the clients have relied on the manager's experience to understand their needs and also to convince them of his choices, the presented approach, in contrast, would support and clarify the manager's recommendations in those situations. Documenting their practice in such a structured way (through discussions and associated decisions for all five categories) would also help to ensure a learning and knowledge sharing environment. In this way, other managers, as well as new employees, would get a chance to develop their expertise more rapidly.

8.4.2.2. Company 2: Oil and gas exploration, designing new systems

Area of work and design challenges: Exploration and commercial production of oil and gas are the main business of the second case company. A significant risk in the design and early execution of a new production project is the placement of exploration drill wells. Their objective is to find new oil or gas fields, based on a sound analysis of the prospect's risks and of the potential hydrocarbon volumes: what is the chance that a drilling (well) will locate hydrocarbons, and what quantity could there be? The design challenges are to understand what the best process and infrastructure design are to discover and explore these fields. They explore different locations and prospects, and their performance depends directly on the success rate of drilling, determined in the early design phase of the project. Test drillings are very expensive and represent a significant investment. To increase the success rate with regard to identifying prospective oil deposits, the opinions of multiple experts are solicited as part of the early engineering systems design risk management. Given that the subsequent detailed design of the whole production system is based on these analyses, attaining higher accuracy in their estimates is of great engineering and financial importance.

Risk management process and link to other organizational processes: I interviewed the head of enterprise risk management twice as part of this case. The interviewee is in charge of facilitating risk quantification workshops. The company reaches high levels of maturity in terms of quantification and also has employees with outstanding risk quantification training. Furthermore, the overall managerial skills of these employees, in terms of running the whole risk management process, are at a high level. Yet there is some space left for improvement, mostly regarding the awareness of their work in the organization and communication to decision makers on the quality of data aspects. Moreover, the impact of the method choice has drawn their attention in recent years, due to the need for greater accuracy in the estimates. As they sometimes face severe uncertainties, they have not, until now, had a framework that would better support their argument for the additional research needs.

Proposed design risk customization and evaluation: Their challenge is to ensure that all parties (not just highly trained people) follow the rationale for any change introduced during and through the risk quantification, and that they are able to illustrate its impact on the different levels of the project and organization. The recommendation in this case was to use the tailoring approach to facilitate the conversation with less risk-aware managers, but also with those managers without an extensive mathematical background. Since they run very complex calculations, it is essential that the managers understand the critical issues in terms of data quality. This can be achieved through the discussion of the five tailoring categories. The feedback I got mostly relates to the fact that the proposed approach would allow a structured conversation among different departments and at different hierarchical levels. Yet the desired changes in their risk management routine (introducing new quantification and visualization methods with a greater focus on data quality) would also impact other processes, such as financial analyses, which can be challenging to achieve, as the organization relies on the current estimates at portfolio and strategic levels.

8.4.2.3. Company three: SME, design in construction

Area of work and design challenges: The third example organization is an engineering and consulting SME that provides design services for construction projects. They experienced several risks in the design phase and have heavy delays on their currently most challenging project.

Risk management process and link to other organizational processes: Since there is no established culture regarding risk in their practice, there is no awareness of the need for it, of ways to incorporate it, or any appreciation of its role in value creation. To understand their practice and established ways of working, I talked to key stakeholders of the project. I conducted separate interviews with a risk management employee, a fire and safety engineer, the project manager, a structural engineer, the HVAC design manager, an electrical engineer-designer, an architect and design manager, and the project owner. The employees have no educational background in risk management, are not aware of any bodies of knowledge, or any sort of risk management procedures. However, one of the engineers was actively exploring professional risk management online training material and professional conferences.

Only after one of their larger projects (the design of a hotel complex) faced heavy delays, did the organization consider investigating methods to help them manage uncertainties. The understanding of such a need comes from managers, whereas some of the engineers see absolutely no reason even to jointly discuss possible risks. In their view, they are the experts for a particular design matter and they take full responsibility for that aspect, expecting the others to do the same in their own domains, without appreciating the potential challenges that are present at the intersections of the domains.

Proposed design risk customization and evaluation: The challenge they are facing is to establish initial risk management practice. The recommendation in this case was to use the proposed tailoring approach to facilitate the conversation about their needs and the importance of establishing risk management practice, from the beginning informing them that the method (tool, technique) they choose should be based on an informed and knowledgeable choice, not just something copied from another company. Furthermore, I highlighted the requirement to ensure a discussion took place about what kind of data they need to support the chosen analysis and how they can ensure its proper quality. The feedback was that the communication through the proposed approach was clear enough for the employees to understand and follow the concept, and some initial understanding and awareness of risk management was achieved.

8.4.2.4. Company 4: Consultancy for the Design phase

Area of work and design challenges: This international, multidisciplinary engineering consultancy company is an example of an organization that provides design services for construction projects. They provide consultancy services for projects such as the design of airports, transportation systems, hospitals and similar projects. They also have some of their own projects in construction.

Risk management process and link to other organizational processes: To understand their practice and established ways of working, I talked to their business and risk management consultant, who works for many different clients in construction. The risk consultant had previously worked across different sectors: the pharmaceutical sector, the utility sector, the beer and beverage sector, and others. They have two types of projects: those where he works as a risk manager as a specialist and those where his work is incorporated into other services provided to the client as part of a bigger service for larger projects, or megaprojects. As part of working on the first type of project, they are helping clients to increase their maturity in terms of risk management. At the moment, he is the one who has the lead on risk management and in terms of resources he collaborates with other departments working on the same projects (for example, the financial team, or a planner). They are aware of the ISO 31000 standard and other practitioners' guidelines.

Proposed design risk customization and evaluation: There was recognition of the great need to develop this type of tailoring approach. Furthermore, the practitioners from this company had been looking into available risk management maturity models before we established the collaboration. They appreciated having an overview of the maturity models described in Section 8.2.2, as it summarizes the basic information they need.

However, the proposed approach is not seen as straightforwardly implementable. The approach is understandable for a risk manager, risk analyst or those very informed on risk management, but in order to present it to the others on the team it would need further adjustments. However, the approach is seen as a valuable supplement to established practice. For instance, it could be used to support risk managers in their work and the communication of the need to improve the quantification. The interviewee also highlighted the fact that this and similar approaches should not be too academic, so as to be more easily applicable in practice.

8.4.2.5. Company 5: Public Organization

Area of work and design challenges: An international organization providing design services for a number of different projects and systems. They provide services to other NGOs, governments and lastly the private sector and private foundations. They are mainly focused on procurement services, project management, and infrastructure, plus they offer some financial management services (such as managing grants) and human resource services (some organizations outsource their recruitment processes). What is very interesting is that they are currently designing their risk and quality framework. The biggest challenge is to design a framework for the whole spectrum of their practice (applicable and manageable for everyone, from those working in the field in war zones to those working in senior positions in offices).

Risk management process and link to other organizational processes: I conducted the interview with their risk and quality group. Previously, risk management was a part of project management. However, the decision was made to establish a separate team and dedicate more attention to the risk management process. Their risk management is part of the wider governance risk and compliance framework.

They are aware of the standards and the practitioners' guidelines, and have also got in touch with some of the practitioners with other companies that already have established risk management. Every project they start has to go through a qualitative risk assessment with 27 questions divided into four risk categories with predefined answers (the level of criticality from 0 to 4 that is tailored to each of the 27 questions). This assessment is followed up quarterly and updated. Operational risk management is daily risk management. They want to find a common basis for everyone to be able to complete their risk management task/report. For more systematic and quantitative information, they could use additional forms/questionnaires.

Proposed design risk customization and evaluation: The proposed approach, in their view, requires expertise, and that could be the main challenge for the implementation of the approach in their organization. As a public entity, they have quite a wide range of projects. Some of the projects involve employees who are aware of risk and project management, but some employees need to be introduced to even the basics. They find the discussion on to what extent they need to go into analyses very interesting and very relevant. Finding the right balance is seen as the key to effective risk management.

8.4.2.6. Company 6: Large-scale high-tech infrastructure design in energy sector

Area of work and design challenges: The sixth case company is involved in designing and deploying large-scale high-tech infrastructure in the energy sector. Designing and operationalizing both onshore and offshore systems is part of their expertise.

Risk management process and link to other organizational processes: Their risk management is recognized as one of the best practices, due to their advanced way of dealing with risk and uncertainties throughout the process, tools and decision making they have adopted and further developed. We previously conducted nine interviews with their senior project risk manager on their risk management and challenges in practice, as well as analyzing the implementation of a complex, quantitative engineering design and deployment project risk model in Primavera. Additionally, we conducted a follow-up interview on the proposed approach.

Proposed design risk customization and evaluation: The company has already reached a high level of risk management maturity and, like Company 1, they constantly seek frameworks for further improvement and carry out re-evaluation. More concretely, they look into ways to improve their processes through advanced risk quantification techniques. Undoubtedly, they show a high level of understanding of the impact and the importance of the choice of risk method and its proper usage. However, the importance, resources (in terms of cost, time, and employees) needed, and the quality of data are some of their main concerns. The feedback on the possibilities of using and implementing this approach is therefore mostly related to the opportunities it provides in terms of better communication to the managers on these two core aspects: why they need better quality of data and why they need to improve their quantification. The downside of the approach is seen in the sense that, in order to properly use it (go in detail through all the criteria), a lot of time would be required. Often, the managers they report to allocate insufficient attention to the long process of reporting (presenting and questioning).

8.5. Discussion

The case companies' interviews supported the view that I had identified in the literature about the diversity of engineering systems designs and, consequently, also of risk management practice. The first company showed a high level of understanding of the impact and the importance of the choice of risk method and its proper usage, and the maturity framework that this thesis proposed would help them to identify detailed improvements, especially in their consulting activities.

For the second company, the proposed customizations would support their communication to less risk-aware managers, as well as to those without an extensive mathematical background. It would also help them identify where they have a greater need for better accuracy in their estimates, the approaches they might take to achieve these, and the necessary argumentation for additional research. The proposed approach could contribute to the sustainability and effectiveness of their risk management process.

For the third organization, I consider that the proposed approach would be useful to help establish a practice to facilitate a conversation about their needs, from the beginning informing them that the method (tool, technique) they choose should be based on an informed and knowledgeable choice, not just "copied and pasted" from another company.

The practitioners from the fourth company highly appreciated this research direction. While the proposed approach is seen as overly "academic" to be directly used in their practice, it is also seen as a valuable way to help risk managers to communicate to others in the company the need to improve the quantification.

The fifth case company has a rather specific risk management profile, due to the nature of their organization. The concrete contribution that the approach introduced in this chapter could bring to their practice was hard to determine, as they are going through the restructuring of risk management process. However, the approach could support their workshops as well as enable support from the academic point of view in the argument for the need to approach the whole risk management process more structurally.

And finally, the sixth company showed a high level of understanding of the impact of the choice of risk method and especially of its proper usage. The proposed approach would help them to start the conversation with their managers about the need to actively improve the quality of data on which they base their assessments, as this represents one of the main concerns in their current practice.

The interviews also supported the view, which I had identified from the research literature, of the need for any engineering systems design risk management process to be adapted to the particular circumstances of the project and of the organization undertaking it. As Loch (2000) demonstrates, based on 90 product development projects, there is no "best practice" product development process; rather, a tailoring approach is needed to help companies achieve their strategic innovation needs. They propose a strategy deployment procedure for product development, which can help an organization to manage its innovation efforts proactively. This and our research are aligned with other related work. For example, Mulqueen, Maples, and Fabisinski (2012) describe tailoring of systems engineering processes with a specific focus on the conceptual design environment. Cabannes *et al.* (2014) propose an approach for taking into account the maturity of information in risk assessments and providing meta-information on the risk estimations, given that there is uncertainty related to information during the design process (particularly in the early design stages). Fontoura and Price (2008) propose a systematic approach to managing risks in software development projects through process tailoring, with the aim of elaborating a defined process to a project suitable to the project's context, taking advantage of agile methods, planned or hybrid, while preventing identified risks for the project. All these approaches are aligned with the approach we propose.

However, tailoring is not an easy task; it requires experience and knowledge in related processes, and concrete recommendations that go beyond the statement that *"you should tailor your risk management process"* are scarce. Furthermore, changes in large organizations can take time and are difficult to implement. Starting from the number of approvals on different hierarchical levels needed for proceeding with a change, to training employees for the new process, and ensuring proper integration with other processes, implementing change represents a challenging task. Therefore, organizations need to treat the implementation itself as a strategic change project. This requires articulating clear objectives as well as success criteria, proper planning and resources, and effective monitoring and control.

The approach taken in this chapter, based on existing risk management process maturity frameworks, with the addition of specific components that enable a concrete tailoring of risk management processes (e.g. decision making) to specific quantification approaches, makes contributions in both these respects.

The proposed tailoring can also be seen from a fit-for-purpose point of view. We believe this also makes the contribution of ensuring that risk management is fit-for-purpose as the dimensions we discuss (understanding of the risk management needs; method sophistication for risk quantification; quality of data; awareness regarding risk in organizational culture and impact of risk assessments in decision making) have a significant impact on it. This promises potential to develop the proposed customization framework into a tool incorporating significant detail on the process level, thus also enabling organizations with less design risk management context knowledge to significantly improve their overall process quality. However, the proposed tailoring approach requires further detailing and application in industry. This would allow reporting of the potential impact of the approach in an organization and its learning and knowledge sharing capacities.

In summary, the key insights obtained through the case companies' validation were:

- Success in using the same tailoring framework at three different companies facing three very different risk profiles and design tasks; the three companies approached to provide their feedback on the developed approach raise the importance of the approach and suggest further improvements;
- the framework yielded practical suggestions to adapt the design risk management process model that were seen both as fitting and relevant by the interview partners;
- while the current application of the framework still requires significant risk management context knowledge (one of the challenges of the current state), the prototypical adaptation has already enabled us to collect concrete examples of alternative modes of executing risk management when using different quantification techniques.

8.6. Conclusions

During engineering systems design, companies deal with uncertainty. The types and degrees of uncertainty vary significantly as the design process progresses, and the choice of methods to deal with risk and uncertainty play a crucial role in achieving the desired results. Therefore, in this chapter I present the research on developing a framework to tailor risk management to the specific company's needs. I accomplished this objective by linking risk management maturity concepts to previous research on product development, project management, and risk management methods, deriving five categories to guide practitioners in the choice of the appropriate method. The proposed framework advances the state of the art by taking into account the quality of the available data, the corporate culture and awareness of risk, and the way responses are planned. I preliminarily tested the validity of our approach in six different companies, showing its value in tailoring risk management to the specific needs and challenges of each of the companies.

Risk management awareness usually occurs after companies have already digested other management practices. These companies have usually already adopted strategic management cultures and methods such as, for instance, product/project portfolio management. The proposed approach enables further improvements of management practices by informing different hierarchical levels on the need for a more adequate process/method, accompanied awareness and its value.

As discussed in more detail in Chapter 9, to fully take advantage of the potential that advanced risk and uncertainty quantification approaches offer, a company (or the analysts/managers responsible) must understand and articulate their overall risk management needs and resources available. In order to improve the accuracy of quantities produced, it is not enough to simply apply a more advanced method. Often, more resources need to be used, such as: hiring trained people or sending already employed experts to training courses, buying new software, getting adequate IT (more powerful computers, data storage, etc.), investing in data collection when needed and enabling data storage for learning purposes.

For that reason, in my view, the greatest potential of these methods is by far in large-scale engineering systems and large-scale projects. It makes sense to invest a bit more in the quality of the analyses on which the important decisions are made, given the impact they later have, as the difference in terms of the outcome can be significant.

However, projects and design solutions that are not necessarily large-scale, but that deal with deep uncertainties, can also benefit significantly from applying advanced approaches. For instance, an IT solution can be designed for a particular online banking service or a tax system, or a transportation system that improves the experience for the end user. It impacts a large number of users, it becomes part of an extremely large system and better understanding is needed of its effect on the overall system. Yet, since it is a first-of-a-kind solution, it cannot be tested elsewhere and therefore it should be recognized that there is a need to cope better with deep uncertainties.

9. Discussion: The non-probabilistic framework and its connection to the current state-of-the-art

"All models are wrong, but some are useful"

- George E.P. Box –

To allow a more thorough reflection on the contributions of this PhD thesis and the general limitations potentially impacting the results, this chapter intends to wrap up the findings of this thesis and elaborate on the replicability and reproducibility of the research.

In the context of this thesis, this is the final step: up to this point, the thesis has investigated advanced risk and uncertainty quantification methods (introduced under the non-probabilistic framework) and the methods' integration into the overall risk management process. This chapter concerns itself with the broader question of the integration of the non-probabilistic framework into the existing state of the art in engineering systems design risk management, potential recommendations for its usage and specific method selection. This is followed by a discussion on methodological reflections and limitations.

The chapter is structured as follows: Section 9.1 provides a rationale for considering the non-probabilistic framework as an additional step – an extension of the current probabilistic (view on) risk management. Section 9.2 further describes this extension through the integration of risk assessment tools and techniques, and the comparison of probabilistic and non-probabilistic methods. Additionally, recommendations depending on types of situations in engineering systems design are provided. Section 9.3 specifically addresses the limitations.

9.1. The extension of the probabilistic view on quantification in risk management

The methods investigated in this thesis are presented as a complement to probabilistic methods to quantify epistemic uncertainty. The basic rationale behind this is that the non-probabilistic methods introduced in this thesis should complement the probabilistic processes in specific situations and scenarios that are dominated by epistemic uncertainty. They are not intended to substitute already established (entire) processes and analyses, but rather to support them in situations in which it has been demonstrated that conventional approaches face challenges (examples are provided in Sections 3.3 and 3.4). Specific situations are documented, and their integration and recommendations are described in Section 9.2.

Moreover, the literature on uncertainty quantification explains that probabilistic methods are basically a special case of the non-probabilistic ones. The main points for making such claims are following:

1. The first group of methods: Mathematically imprecise probabilities are seen as a natural extension of probabilities, because different studies show that, with sufficient evidence (information), the intervals converge to precise estimates (Walley, 1999; Weichselberger, 2000).

   An exception to this line of thinking is Dempster-Shafer theory, which is created on completely different pillars: degree of belief and no condition to sum probability intervals to unity (Beynon, Curry, & Morgan, 2000).
2. The second group of methods: Semi quantitative methods are technically a combination of probabilistic approaches and qualitative descriptions and/or visualizations that support the evidence behind the conducted analyses (Aven, 2008; Boone *et al.*, 2010).
3. The third group of methods: Exploratory modeling enables the sampling of a large number of scenarios for different plausible futures, on a large number of different inputs, for which conventional scenario analysis becomes only a subset of the new scope (Bankes, Walker, & Kwakkel, 2013).

   These methods are aimed at enabling modeling for system resilience and providing business continuity support.

As an example, consider cyber security risks. Over the last five to ten years this type of risk has risen significantly in terms of size, number of attacks, their mutation capacities and effect. Major organizations can almost certainly expect to be the potential subject of an attack speaking in 3-5 year horizon. While they cannot be sure of the exact probability of this happening, they can prepare recovery systems, that are as good as and as fast as possible, to minimize the potential harm to their systems. In this context, the non-probabilistic methods assessment supports such aims.

9.2. Overview of the existing approaches and the comparison with non-probabilistic methods

To allow a discussion on the integration of the non-probabilistic methods into the broader, widely available and used set of methods (a broadly accepted collection), both the academic and practitioners' communities need to consider various angles in which such integration contributes to the field. In that view, this thesis adds to the current state of the art by: 1) uplifting the importance of specifically addressing epistemic uncertainty, 2) providing the means to do so, 3) integrating developed methods into the broadly accepted collection of methods and 4) articulating recommendations to both research and practice.

In the previous chapters, this thesis has explained the first two aspects. After introducing, describing and analyzing different concepts, the explanations for their integration into the overall risk management process are discussed through the proposed tailoring approach. There is a need to more thoroughly discuss the concrete choice of a specific method. The third aspect aims to address this gap. As explained at the beginning of this thesis, the methods used in analyzing risks can be qualitative, semi-quantitative or quantitative. The degree of detail required will depend upon the particular application, the availability of reliable data and the decision-making needs of the organization. Some methods and the degree of detail of the analysis may be prescribed by legislation in a field. Other types of design solutions/projects may have certain domain tendencies ("common practice"), some can be contractual requirements, or there may be no usual approach at all.

The work presented here builds on the existing literature and what is considered to be the best practice. More concretely, the foundation is found in the ISO 31000 risk management process description and ISO 31010 list of approaches (Institute, 2011). In light of this, the extension of Applicability of Tools used for Risk Assessment in ISO 31010 is presented in Appendix 3. Such an extended list allows the creation of awareness about the existence and availability of these approaches, and their applicability possibilities.

Furthermore, a more detailed integration is developed and presented in Appendix 4, where methods are briefly described and further characterized. The significant extension here lies in two groups of methods: 1) software assessments and 2) non-probabilistic methods.

Over the last decade, software (computationally-based and web-based) solutions have emerged. As they provide a valuable contribution to the current practices, this thesis suggests expanding the current most-widely accepted list of Tools and Techniques with this group of methods. Namely, these solutions allow the usage of multiple tools and techniques from the same list in a structured, transparent and traceable way. This proposed extension is provided in Appendix 4 under "Software Assessment."

The included examples are: Oracle's Primavera Risk Analysis, RamRisk, RAMAS, Resilinc, and Risk Calc (for references, please see Appendix 4). These types of methods brought a number of benefits, from which some were documented through the empirical work presented in Section 3.4. For instance, they allow the integration of both predeveloped risk registers and newly developed risk registers, they can identify common scheduling pitfalls, they can report confidence levels with regard to finishing dates, costs, internal rates of return, and net present values. Important features include more employees having the access to the same data and/or analysis at the same time, the possibility of allocating tasks and responsibilities, round-the-clock access from multiple devices, confidentiality (possibilities for restrictions), customized reporting, mail notifications, data storage, benchmarking against peers, etc.

Therefore, these cloud-based platforms provide a powerful array of tools for risk management, model development, benchmarking and business continuity. These methods are user friendly, and project members can contribute directly and collaborate efficiently when performing analyses. They allow integration of multiple analyses, maturity assessments, and the use of various visualization options to customize solutions, access and reporting. For these reasons, they represent a valuable contribution to the field.

The methods introduced in this thesis represent the second important extension of the overall collection (see Appendix 4). Non-probabilistic methods are in this way included in the list, and their merits have previously been discussed.

Other than integrating developed methods to the broadly accepted collection of methods, this thesis also provides a comparison of the methods with some of the most widely used ones (Table 11). As illustrated, it is important to understand what the required inputs are for each of the analyses and what the expected outcomes are. The first three methods require more specific input, with predefined system variables. The three non-probabilistic methods instead open up various possibilities for situations when such predefined information is not available: by elicitation of different formats, by providing qualitative explanations regarding the assumptions and by sampling a large number of scenarios.

The comparison of the outputs reveals that, in the case of the first three methods, we can expect a precise outcome and a clear recommendation. In contrast, the non-probabilistic methods provide a broader description of a wider range of different outcomes according to the initial input. Strengths and weaknesses discussed in this thesis are summarized and presented in Table 11.

Table 11 Comparison of the probabilistic and non-probabilistic methods

Method	Input	Output	Strengths	Weaknesses
Monte Carlo Analysis *Largely taken from the ISO 31010 standard	- Good model of the system and information on the types of inputs, the sources of uncertainty that are to be represented and the required output - Uniform, triangular, normal and log normal distributions are often used for this purpose	- The output could be a single value - It could be a result expressed as the probability or frequency distribution - It could be the identification of the main Standards' functions within the model that has the greatest impact on the output - In general, a Monte Carlo simulation will be used to assess either the entire distribution of outcomes or key measures from the distributions	- Models are relatively simple to develop and can be extended as the need arises - Sensitivity analysis can be applied to identify strong and weak influences - Software is readily available and relatively inexpensive	- The accuracy of the solutions depends upon the number of simulations which can be performed (this limitation is becoming less important with increased computer speeds) - It relies on being able to represent uncertainties in parameters by a valid distribution - Large and complex models may be challenging to the modeler and make it difficult for stakeholders to engage with the process - The technique may not adequately weigh high consequence/low probability events and therefore not allow an organization's risk appetite to be reflected in the analysis

Consequence Probability Matrix *Largely taken from the ISO 31010 standard	- Customized scales for consequence and probability - A matrix which combines the two	- The output is a rating for each risk or a ranked list of risk with significance levels defined	- Relatively easy to use - Provides a rapid ranking of risks into different significance levels	- It is difficult to define the scales unambiguously - Use is very subjective and there tends to be significant variation between raters - Risks cannot be aggregated (i.e. one cannot define that a particular number of low risks or a low risk identified a particular number of times is equivalent to a medium risk) - It is difficult to combine or compare the level of risk for different categories of consequences
Bayesian analysis *Largely taken from the ISO 31010 standard	- Define system variables - Define causal links between variables - Specify conditional and prior probabilities - Add evidence to net - Perform belief updating - Extract posterior beliefs	- The graphical output provides an easily understood model and the data can be readily modified to consider correlations and sensitivity of parameters	- All that is needed is knowledge of the priors - Inferential statements are easy to understand - Bayes' rule is all that is required - It provides a mechanism for computing subjective beliefs in a problem	- Defining all interactions in Bayes nets for complex systems is problematic - Bayesian approach needs the knowledge of a multitude of conditional probabilities, which are generally provided by experts - Software tools can only provide answers based on these assumptions
Imprecise Probabilities (Coherent upper and lower probabilities, (Walley, 1991))	- Expert judgments elicitation in different formats	- Depending on the data format, the corresponding aggregation – single point, distribution, envelope, c-box	- Various options for aggregating expert opinions - Imprecision explicitly manifests the degree of knowledge or ignorance - The greater the interval the greater our ignorance is	- Harder to communicate the results - Harder to compute - Immaturity of the field - Can be a large imprecision that makes conclusions impractical

Semi-quantitative methods (The NUSAP tool, (Funtowicz & Ravetz, 1990))	- Quantitative assessment followed by qualitative information on all the steps, assumptions and potential limitations	- Precise quantitative output supplemented by descriptions of quality of data and visualizations representing uncertainty surrounding the results	- Important information about the assumptions included in the analysis are provided to decision makers - Facilitate conversation about uncertainties, convenient for discussions with lay public	- Not widely used, lack of awareness of its abilities/capacities - Can be time consuming to present all the details
Exploratory modeling (RDM, (Walker, Haasnoot, & Kwakkel, 2013))	- Development of agents - Sampling of a large number of scenarios - Use trusted simulation models to consider a wide spectrum of plausible futures, each with different input parameters	- A robust solution for a system/design that works satisfactorily over a broad range of possible tested futures	- Advanced and thorough approach - Highly relevant for complex systems	- Computational complexity - Seeks robustness rather than optimal solution - Demands a strong set of skills for its usage

Regarding the fourth aspect in which this thesis adds to the current state of the art, it develops recommendations for the specific situations, previously identified as critical. This recommendation is framed as a combination of findings from the comparison of the methods, from literature and from the empirical studies (Table 12).

It is important to clarify the *Resource dependency in the first recommendation. As the methods introduced in this thesis can be demanding in terms of resources, a number of other factors need to be considered when deciding on the method (apart from risk and uncertainty quantification considerations). The scope, size, budget, lifetime, and current level of risk management of projects play an important role. For instance, the greatest potential of these methods is in large-scale systems, also because of the fact that their budget allows such results to bring value to the organization. In the case of smaller systems, it can simply be too costly to be appreciated, even though it may lead to better results. Employing these methods involves having adequately educated employees, specialized in the topic, which is more reasonably to be expected in larger organizations. In the case of time pressures, there is a tendency to neglect all the benefits from a more thorough investigation of system components and use tools that allow a rough estimate.

Table 12 Recommendations for method selection for the key situations

Situation	Probabilistic	Non-probabilistic
"Low likelihood, high impact" Resource dependent*	Advantage: Well-known and well established in practice. Disadvantage: Have been challenged in the accuracy and reliability of their results, making them arbitrary to use. Do not really cope with black swans.	Advantage: Bring more relevant information to decision makers. Disadvantage: Do not provide a single, simple, straight forward answer, which is harder for decision makers to comprehend. New methods, representing a challenge for practice to integrate with other processes.
"No information", or hardly any knowledge or prior experience available. Typically, a first-of-a-kind situation	Advantage: Relying on managers' experience. Disadvantage: Arbitrary results. Possibilities for misrepresentation.	Advantage: **Recommended,** by providing the means to do so. Disadvantage: Computational requirements and an adequate educational level of employees that is needed.
"50-50 %" Possibility of distinguishing between the actual 0.5 outcome and default assigned probability	Advantage: Disadvantage: Assigning 0.5 by default making it hard to further explain to managers.	Advantage: **Clear.** Disadvantage: A new method that the whole team needs to switch to. It also needs to be integrated into other processes.

"Historical data sets available", several experienced experts involved, experience in providing similar solutions-systems	Advantage: **Clear.** Disadvantage: Computational complexity and costs of conducting the full analysis can be high. Time-consuming when there is a pressure to make fast decisions.	Advantage: Disadvantage: Add additional complexity and costs (in terms of time, money and other resources) that can be hard to justify.

It is possible to anticipate that in the future, with an advance in methods and computational developments, and a wider availability of risk management educational programs (MSc studies and PhD projects), that this dependency may no longer be as high. However, at the moment, it represents one possibility for future research that would make methods more hands-on and their results easier for communication.

9. 3. Methodological reflections and limitations of the study

This PhD study has been carried out to investigate practical challenges taking place in the real world, which is not always controllable. As such, this research differs from studies taking place in controlled labs, where replicability can be achieved and demonstrated in an easier manner. Thus, the choice of this type of research may have consequences for validity, generalizability and repeatability.

Validity

Validity refers to the extent to which the findings of a research study accurately reflect the studied phenomenon (Collis & Hussey, 2009). There is an ongoing discussion around the criticism of case studies in different scientific communities. The most vociferous opponents indicate that such studies tend to be subjective and are often biased in the researchers' interpretations. However, others, for example, Flyvbjerg (2006a), argue that case studies are neither more subjective nor more biased than other methods of inquiry, if additional wits are considered.

To do so, and by following Lincoln and Guba (1986), *triangulation*, *peer debriefing* and *member check* were applied. I employed method triangulation (Guba, 1981) by combining case studies work with interviewing technique. On the other hand, a noteworthy limitation of this study lies in the inability to employ data triangulation. Due to the confidentiality issues, it was not feasible to arrange access to actual companies' documentation and records or to perform observations. Such insights could have significantly enriched and strengthened the findings.

Peer debriefing was organized in both the scientific and practitioners' communities (Guba, 1981). As part of the scientific peer debriefing, I shared my research challenges with the research group of which I was part, and received critical feedback and challenging questions that enabled me to consider various angles and alternatives to my interpretations. During my first external research stay at Delft University of Technology, researchers from the Multi-Actor Systems Department challenged the extensiveness of the non-probabilistic framework. This resulted in the expansion of the framework with a whole new group of methods (the third group). During my second external research stay at Applied Biomathematics, USA, the researchers working at the Institute challenged the development of the case studies and the generalizability of the data. This resulted in me refining my argumentation and improving the data collection and coding. Furthermore, I presented subanalyses and received feedback from the broader researcher communities (design, project management and risk analysis).

Peer debriefing in the practitioners' communities was carried out through participation and presentation at professional conferences, seminars and forums, where I received feedback from risk managers, project managers, portfolio managers and policy makers, as well as business modelers and analysts. Their experiences have put this research into a very practical application perspective, encouraging me to develop solutions that seek to be applicable in and relevant for practice.

The third step in improving the validity and reliability of this study was member check, which was based on the presentation of initial constructs and results to the interviewees and participants of the case studies. This enabled me to validate interpretations and helped me clarify that the data generating code was conducting representative samples.

Generalizability and transferability

In research, generalizability refers to the extent to which the results of one study can be extended to a wider sample (Collis & Hussey, 2009). More specifically, in his work, Yin (2013) points out the limitations of (single) case study generalizability that applies to this study as well. However, by introducing the term "transferability," Flyvbjerg (2006a) argues that achieving learning from examples is a valuable outcome of a case study. Transferability refers to the fact that principles from one or more case studies can be transferred to similar contexts. While case studies cannot generalize beyond their sample (Kathleen M. Eisenhardt, 1989) and create/provide conclusions that will be applicable in all situations, they can result in knowledge relevant for similar contexts. Lincoln and Guba (1986) propose two ways to enable transferability—*purposive sampling* and *descriptive data*.

As purposive sampling is not intended to be representative of a wider sample/population (Guba, 1981), it should allow comparison between cases based on specific criteria. The case companies (where interviews were conducted, presented in Chapter 8) were selected on the maximum variation criterion, which is a criterion for purposive sampling (Flyvbjerg, 2006a). Identification of commonalities across cases that vary to the maximum degree are considered highly transferable. However, the focus of sampling was on companies within engineering systems design. As such, this sector has characteristics that differ from other sectors. To have substantiated arguments for claiming transferability within risk analysis, but in other domains (such as ecology, or the pharmaceutical or financial industries), would require certain adaption or additional research. However, fields such as project, portfolio and megaproject management have significant overlap with design, making the results potentially transferable to that domain.

Descriptive data, on the other hand, provide extensive information about a specific case (Lincoln & Guba, 1986). As transferability depends on matching characteristics between contexts (Lincoln & Guba, 1986), this thesis includes thorough case descriptions and case company descriptions. By providing those data, this thesis enables readers to compare the cases with other possible contexts and future applications.

Repeatability

As mentioned above, due to the high context dependency, the reliability of case studies has been questioned in research communities. This emerges due to the fact that the replication of a case study within a different yet related context will not necessarily result in consistent findings (Thomas, 2015). Reliability refers to the absence of differences in results if the research is repeated (Collis & Hussey, 2009). However, others suggest that the reliability of case studies can be increased through *audits* (Lincoln & Guba, 1986).

Audits involve allowing other researchers or companies to check the conducted analysis (Lincoln & Guba, 1986). This study involves audit initiatives, both from researchers and practitioners. Furthermore, for both case studies, the setup was discussed in another large company to ensure its repeatability, consistency and accuracy.

For all these reasons, I argue that the main principles of the findings and the non-probabilistic framework followed by the tailoring approach can (with adaptions) be transferred to risk management in other sectors. In this way, the findings, the non-probabilistic framework and the tailoring approach have potential in the broader context of risk management.

10. Conclusions: Learnings from applying advanced risk quantification methods to engineering systems design

"We are what we repeatedly do. Excellence, then, is not an act, but a habit."

- Aristotle-

Risk assessment methods have become widespread in supporting decision making processes. These methods have to create the necessary level of confidence in their results to satisfy the decision makers. To create this confidence in risk management, the key is to have a transparent and systematic analysis and representation of uncertainty.

As outlined in Chapter 3, engineering systems design risk management practice has so far widely relied on probability-based methods when treating uncertainty. However, both theoretical and practical challenges have emerged, within engineering systems design as well as in risk management in other domains. On the other hand, the non-probabilistic methods emerged in other fields as alternatives to the challenging, and still not fully understood, epistemic uncertainty quantification. Only a few studies have investigated these non-probabilistic methods, and little or no attention has been devoted to their application and integration into engineering systems design. Yet these methods are essential when exploring different means to represent uncertainty as part of risk management in engineering systems design. Therefore, it is crucial – in terms of both research and practice – to extend our knowledge base on risk management (quantification) in engineering systems design.

As part of a comprehensive investigation and rethinking of uncertainty quantification in engineering systems, industry involvement was initiated through conducting case studies and interviews with practitioners to extract their professional knowledge (Chapters 3, 5, 6, 7 and 8). These practical insights provided a unique opportunity to investigate the non-probabilistic methods in the engineering systems design context and extend our knowledge base. The study was guided by four research questions (see Chapter 2) that supported the development of the non-probabilistic framework to support: 1) researchers in analyzing and representing (epistemic) uncertainty and 2) practitioners in planning, facilitating and performing their risk management.

The following conclusions summarize the core findings, and relate them to the existing research and practice.

10.1. Core contributions of this thesis

This thesis makes the following contributions:

1) The thesis clarifies different concepts in uncertainty quantification and their limitations, and synthesizes challenges with the currently most widely used methods

In Chapter 3, this thesis confronts research question 1. The literature review and conducted empirical studies provide the following finding: current risk management practice has challenges in terms of uncertainty quantification. The specific challenges in terms of modeling, quality of background knowledge and use and integration of results are documented and discussed. Furthermore, an overview of the main concepts in uncertainty quantification is provided, followed by the literature findings on their limitations, misuse or misrepresentation. The subsequent finding is that there is a need to investigate alternative approaches to adequately represent epistemic uncertainty.

2) The thesis develops the non-probabilistic framework

In Chapter 4, this thesis addresses research question 2. The main findings are literature-based: other fields have dealt with similar issues, with the result that those fields have developed advanced methods for coping with uncertainty. The conceptual development here argues for the introduction of the methods to the engineering systems design field and exploring their application potential in the field. The contribution to our knowledge base in engineering systems design is collecting a broad range of methods, providing their systematization and categorization, and conceptually discussing their application in the engineering systems design context. Furthermore, the literature-based findings about the limitations of the non-probabilistic methods are acknowledged – their complexity has been recognized as the main reason for their not being widely used in terms of the analyses and resources needed.

3) The thesis creates a usable, practical toolkit for practitioners

In Chapters 5, 6 and 7, this thesis addresses research question 3. It transfers more general non-probabilistic methods into usable tools: a method for eliciting expert judgment in different data formats, a method for aggregating different data formats in expert judgment, a qualitative Pedigree scoring for quality of data, a visualization method for uncertainty around data, and a bias correction method. This is done through examples of case study applications in the oil and gas industry (Chapters 5 and 6), followed by their comparison with several traditional probability approaches in representative situations (Chapter 9). The main findings here are based on empirical research work: for the engineering design situations and scenarios tested in this thesis, the non-probabilistic methods provided a more reliable representation of uncertainty.

4) The thesis develops a tailoring approach to tie the quantification needs to the overall risk management process capabilities

In Chapter 8, this thesis addresses research question 4. After introducing, describing and analyzing different concepts, the explanations for their integration into the overall risk management process are discussed through the proposed tailoring approach. The main finding highlights the fact that the success of choosing a specific quantification method from the available options is context dependent, and a broader risk management process view needs to be carefully considered when tailoring risk management to specific design situations, rather than the simple picking of a specific advanced quantification method.

In the end, there is a need to discuss more thoroughly the actual choice of a specific method. Appendices 3 and 4 provide insights into the rich set of available methods (including the non-probabilistic extension) to fit and address different problems. This thesis proposes that the choice of the methods is dependent on context, resources and capabilities (the discussion is opened in Chapter 9). This means that once practitioners have understood their needs (through the proposed tailoring approach), a conscious choice needs to be made in terms of the allocation of resources, the adequate usage of a method, and the acknowledgement of each method's limitations (instead of hiding such obstacles). Furthermore, the importance of an adequate level of employees' education (in terms of risk management and risk quantification) is raised, as methods involving higher mathematical sophistication require certain skills in both conducting analyses and presenting, communicating and explaining the results to various stakeholders involved in decision making.

10.2. Implications for research in engineering systems design

This thesis has the following implications for research:

- The thesis offers a clarification of different concepts of risk and uncertainty quantification, documenting the challenges and limitations of the current practice of uncertainty quantification. Researchers can use this information to guide the development of approaches to overcome these limitations. Furthermore, the clarification provided in this thesis could be a good starting point for young researchers seeking to explore their research directions.

- The field of engineering systems design is enriched by a collection of advanced risk and uncertainty representation approaches, introduced through the non-probabilistic framework. The non-probabilistic methods extend the umbrella of available approaches, which can also be researched in other case applications.

- The field of risk management is enhanced by concrete examples and case studies for the particular needs of one insufficiently researched domain: engineering systems design. This contributes to the overall verification of risk management methods and tools and their applicability, usability and generality.

- It is critical to clarify the fact that we should not expect unrealistic answers from science. The intention of quantification methods is not to eliminate uncertainty, but rather to provide its effective management. It is necessary to raise awareness about this central distinction among other researchers in the community, as well as the lay public.

10.3. Implications for practice in risk management

This thesis has the following implications for practice:

- Managerial implications include support for decision making under uncertainty in engineering systems design. This can help practitioners to be aware of the pitfalls of current practices and reflect on the opportunities for improving their risk management process.

- Throughout the project, one of the main goals and contributions to the practice has been ensuring a higher level of understanding of uncertainty, its nature and types, and the need to cope with it knowledgably in the field by presenting and participating in various events.

Another goal was ensuring practitioners' greater awareness of the need to more thoroughly revise the way they decide on a risk quantification method and, in particular, how they deal with epistemic uncertainty, through active industry engagement, different industry involvement and presentation across various levels of managerial hierarchy.

- A tailoring risk management approach is developed, based on a risk management maturity model, which can be seen as an extension of or additional feature of the ISO 31000 Standard and the risk management process. The extended process view provided through this tailoring approach enables a better understanding of the need for and applicability of the approaches introduced in this thesis.

- The contributions are highlighted through the lens of the current trends in design, such as industrial product-service systems and the integration of various systems (particularly those that face epistemic uncertainty due to their innovative nature and first-of-a-kind solutions). Globalization and rapid technological changes demand proactive monitoring and timely reactions and decisions, while keeping options open for future possibilities. Risk management should therefore keep pace with the evolving and dynamic nature of engineering systems.

10.4. Directions for future research

The paragraphs that follow contain suggestions for relevant future research in relation to the four core findings and the non-probabilistic framework. Future studies can continue the development of "beyond probabilistic" thinking in risk and uncertainty quantification in design as this thesis provides the basis for 1) collecting/adding/developing more (and new) methods, 2) addressing particular challenges through the application of selected approaches in specific design situations, and 3) further fostering the impact of such applications on decision making processes and the overall quality of design solutions.

Core contribution 1 and related findings open up the need for future research of other types of risk management challenges in engineering systems design. For instance, purely behavioral aspects are discussed only to a certain degree in this thesis. Different perceptions of different stakeholders when communicating non-probabilistic results are still to be investigated. Additionally, methods that specifically address or integrate the treatment of ambiguity are an interesting direction.

Core contribution 2 and related findings suggest further investigation of other alternative approaches and their integration into the non-probabilistic framework. For instance, specific literature searches in fields such as mechanical engineering or IT have been outside of the scope of this thesis, but could potentially provide more resources.

Core contribution 3 and related findings reveal opportunities for various types of research. First, additional cases from different industries and different contexts would help in identifying and documenting both opportunities for, and the limitations of, the application of the approaches introduced in this thesis. As the generalizability of the conducted studies is limited, these additional studies would allow a clearer understanding of specific boundaries. Second, applications in real case studies, or even past projects (but on actual data) would provide new insights. Such applications would enable further adjustments of the methods for particular cases and design challenges. Those adjustments would provide practitioners with more hands-on tools, and provide researchers with the ability to streamline some of the more mathematically complex approaches and to seek other application domains.

Third, analyzing actual data is of great importance as it could demonstrate the ability of the introduced approaches to cope with “the noise.” Often, due to time pressures, mistakes or intentional approximations of some form accompany the documentation recorded in organizations. Many of the approaches can first be tested on the historical data (which is the biggest strength for the companies that store this information). Such valuable material can allow the comparison of the approaches and their results, and inform on the level of professionals’ ability to read and comprehend the results on different levels.

Finally, the diversity of design projects and their solution space have a broad range. Core finding 4 and related findings allow a proper understanding of the particular needs in terms of risk management, highlighting specific situations in which advanced methods are essential. The positive feedback on the tailoring approach opened the door for additional discussions, implementation examples and the commercialization potential of the approach. In this way, it could be the basis for the systematic improvement of risk management and its integration into the broader managerial processes. Future studies could apply the approach in different companies, making the findings more generalizable and opening up the possibility of a bigger scientific contribution. The proposed methods, and the related tailoring approach and gained knowledge, can also be introduced and used in other domains, such as project and program management and supply change management, where many of the challenges are based on design issues or are emerging in design.

An acceptable and largely affordable way for companies to further explore non-probabilistic methods could be through encouraging and supporting its employees to study further and engage in MSc projects and other research activities. In this way, a company could potentially benefit from the results of the application of the methods, educating its employees and possibly identifying areas for improvement in its risk management process.

My recommendation for future studies is, if feasible, to design a study over a longer period of time. This would enable the capturing of a more nuanced research process and allow sufficient time to provide a significant amount of detail in each of the analyses conducted. I would seek to develop case studies in various companies involved in engineering systems design, evaluating their processes and methods applications. It is essential to establish a fruitful industrial collaboration built on mutual trust and respect, and to be open to understanding practitioners' needs (both short- and long-term ones). This is a key element in adjusting risk and uncertainty quantification methods for their direct, practical usage, and in opening supplementary research questions.

This thesis concludes with the following insight: there is no single best method for quantifying every type of uncertainty. Context, resources and application skills play a major role, as well as a proper understanding of different schools of thought when applying methods. The thesis looks into and identifies situations when the non-probabilistic methods are more adequate to use, describes these situations and provides the means (tools) to apply them in the identified contexts. Nevertheless, this high potential of the non-probabilistic methods in engineering systems design is dependent on their integration into the overall risk management and associated processes. These must be carefully and knowledgably planned and carried out in order to harness this potential and to achieve an actual design impact in practice.

References

Ainsworth, S., & Loizou, A. T. (2003). The effects of self-explaining when learning with text or diagrams. *Cognitive Science*, *27*(6), 669–681. http://doi.org/10.1016/j.cogsci.2003.08.001

Alting, L., & Legarth, J. B. (1995). Life Cycle Engineering and Design. *CIRP Annals*, *44*(2), 569–580.

André, J. C. S., & Lopes, D. R. (2012). On the use of possibility theory in uncertainty analysis of life cycle inventory, 350–361. http://doi.org/10.1007/s11367-011-0364-9

Apostolakis, G. E. (2004). How Useful Is Quantitative Risk Assessment? *Risk Analysis*, *24*(3), 515–520. http://doi.org/10.1111/j.0272-4332.2004.00455.x

Argyris, C., & Schon, D. A. (1989). Participatory Action Research and Action Science Compared. *American Behavioral Scientist*, *32*, 612–623.

Atie, S. A. (2008). Wicked Problems and Applied Economics. *American Journal of Agricultural Economics*, *90*(5), 1176–1191. http://doi.org/10.1111/j.1467-8276.2008.01202.x

Aughenbaugh, J. M., & Paredis, C. J. J. (2005). The value of using imprecise probabilities in Engineering Desing. In *ASME 2005 Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (pp. 1–13).

Aven, T. (2008). A semi-quantitative approach to risk analysis, as an alternative to QRAs. *Reliability Engineering & System Safety*, *93*(6), 790–797. http://doi.org/10.1016/j.ress.2007.03.025

Aven, T. (2011). The risk concept — historical and recent development trends. *Reliability Engineering and System Safety*, *99*(951), 33–44. http://doi.org/10.1016/j.ress.2011.11.006

Aven, T. (2013a). On the meaning of a black swan in a risk context. *Safety Science*, *57*, 44–51. http://doi.org/10.1016/j.ssci.2013.01.016

Aven, T. (2013b). Practical implications of the new risk perspectives. *Reliability Engineering and System Safety*, *115*, 136–145. http://doi.org/10.1016/j.ress.2013.02.020

Aven, T. (2014). *Risk, surprises and black swans: fundamental ideas and concepts in risk assessment and risk management*. Routledge.

Aven, T., Baraldi, P., Flage, R., & Zio, E. (2014). *Uncertainty in Risk Assessment: The Representation and Treatment of Uncertainties by Probabilistic and Non-Probabilistic Methods*. John Wiley and Sons Ltd.

Aven, T., Ben-Haim, Y., Andersen, H. B., Cox, T., Droguett, E. L., Greenberg, M., Zio, E. (2015). SRA glossary.

Aven, T., & Renn, O. (2010). *Risk Management.* In: Risk Management and Governance. Risk, Governance and Society, vol 16. Springer, Berlin, Heidelberg.

Aven, T., & Zio, E. (2011). Some considerations on the treatment of uncertainties in risk assessment for practical decision making. *Reliability Engineering & System Safety*, *96*(1), 64–74. http://doi.org/10.1016/j.ress.2010.06.001

Aven, T., & Anthony, C. (2015). Introduction to Special Virtual Issue: Foundations of Risk Analysis. *Risk Analysis*.

Bankes, S. (2002). Tools and techniques for developing policies for complex and uncertain systems. In *The National Academy of Sciences* (Vol. 99).

Bankes, S., Walker, W. E., & Kwakkel, J. H. (2013). Exploratory modeling and analysis. In *Encyclopedia of operations research and management science* (pp. 532–537). Springer US. https://doi.org/10.1007/978-1-4419-1153-7_314

Barrett, J., & Lampard, D. (1955). An expansion for some second-order probability distributions and its application to noise problems. *IEEE Transactions on Information Theory*, *1*(1), 10–15. http://doi.org/10.1109/TIT.1955.1055122

Bauer, M. (1997). Approximation algorithms and decision making in the Dempster-Shafer theory of evidence—An empirical study. *International Journal of Approximate Reasoning*, *17*(2), 217–237. http://doi.org/http://dx.doi.org/10.1016/S0888-613X(97)00013-3

Beard, A. N. (2004). Risk assessment assumptions. *Civil Engineering and Environmental Systems*, *21*(1), 19–31. http://doi.org/10.1080/10286600310001605489

Beer, M., Ferson, S., & Kreinovich, V. (2013). Imprecise probabilities in engineering analyses. *Mechanical Systems and Signal Processing*, *37*(1–2), 4–29. http://doi.org/10.1016/j.ymssp.2013.01.024

Ben Abdallah, N., Mouhous-Voyneau, N., & Denœux, T. (2013). Combining statistical and expert evidence using belief functions: Application to centennial sea level estimation taking into account climate change. *International Journal of Approximate Reasoning*, *55*(1 PART 3), 341–354. http://doi.org/10.1016/j.ijar.2013.03.008

Bernardo, J. M., & Smith, A. F. M. (2009). *Bayesian theory*. Wiley.

Berner, C., & Flage, R. (2015). Strengthening quantitative risk assessments by systematic treatment of uncertain assumptions. *Reliability Engineering and System Safety*, *151*, 46–59. http://doi.org/10.1016/j.ress.2015.10.009

Berner, C. L. (2016). *Contributions to Improved Risk Assessments –To Better Reflect the Strength of Background Knowledge*. PhD Thesis.

Berner, C. L., & Flage, R. (2016). Comparing and integrating the NUSAP notational scheme with an uncertainty based risk perspective. *Reliability Engineering and System Safety, 156*, 185–194. http://doi.org/10.1016/j.ress.2016.08.001

Bernstein, P. (1996). *Against the gods: the remarkable story of risk.* New York: Wiley.

Beuren, F. H., Gitirana, M., Ferreira, G., & Miguel, P. A. C. (2013). Product-service systems : a literature review on integrated products and services. *Journal of Cleaner Production*, *47*, 222–231. http://doi.org/10.1016/j.jclepro.2012.12.028

Beynon, M., Curry, B., & Morgan, P. (2000). The Dempster-Shafer theory of evidence: An alternative approach to multicriteria decision modelling. *Omega*, *28*, 37–50. http://doi.org/10.1016/S0305-0483(99)00033-X

Blessing, L., & Chakrabarti, A. (2009). *DRM, a Design Research Methodology*. Springer. http://doi.org/10.1007/978-1-84882-587-1

Bolger, F., & Rowe, G. (2015). The Aggregation of Expert Judgment: Do Good Things Come to Those Who Weight? *Risk Analysis*, *35*(1), 5–11. http://doi.org/10.1111/risa.12272

Boone, I., Van Der Stede, Y., Dewulf, J., Messens, W., Aerts, M., Daube, G., & Mintiens, K. (2010). NUSAP: A method to evaluate the quality of assumptions in quantitative microbial risk assessment. *Journal of Risk Research*, *13*(3), 337–352. http://doi.org/10.1080/13669870903564574

Bourgeois, B. S., Elmore, P. A., Avera, W. E., & Zambo, S. J. (2016). Achieving comparable uncertainty estimates with Kalman filters or linear smoothers for bathymetry data. *Geochemistry, Geophysics, Geosystems*, *17*(7), 2576–2590. http://doi.org/10.1002/2015GC006239

Brocéliande team. (2015). NUSAP: Numeral Unit Spread Assessment Pedigree. Retrieved January 1, 2015, from https://proxy.eplanete.net/galleries/broceliande7/nusap-numeral-unit-spread-assessment-pedigree

Bryant, B. P., & Lempert, R. J. (2010). Thinking inside the box : A participatory , computer-assisted approach to scenario discovery. *Technological Forecasting & Social Change*, *77*(1), 34–49. http://doi.org/10.1016/j.techfore.2009.08.002

Buede, D. M., & Girardi, P. (1997). A target identification comparison of Bayesian and Dempster-Shafer multisensor fusion. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*. http://doi.org/10.1109/3468.618256

Burgman, M., Fidler, F., Mcbride, M., Walshe, T., & Wintle, B. (2006). *Eliciting Expert Judgments: Literature Review*.

Cabannes, G., Mey, Y., Nadège, G., Gidel, T., & Mcmahon, C. (2014). Taking account of information maturity in assessing product risk. *International Journal on Interactive Design and Manufacturing*, 243–253. http://doi.org/10.1007/s12008-014-0228-1

Cagliano, A. C., Grimaldi, S., & Rafele, C. (2014). Choosing project risk management techniques . A theoretical framework framework. *Journal of Risk Research*, *9877*(Mar), 1–17. http://doi.org/10.1080/13669877.2014.896398

Cantarelli, C. C., Flyvbjerg, B., Molin, E. J. E., & van Wee, B. (2010). Cost overruns in large-scale transportation infrastructure projects: Explanations and their theoretical embeddedness. *European Journal of Transport and Infrastructure Research*, *10*(1), 5–18.

Caron, F., & Ruggeri, F. (2016). Project Management in the Oil & Gas Industry – A Bayesian Approach. In *Wiley StatsRef: Statistics Reference Online, 1*. http://doi.org/doi.org/10.1002/9781118445112.stat07893

Carruccio, E., & Quigly, I. (2006). The Logic of Aristotle. In *Mathematics and Logic in*

History and in Contemporary Thought. New York: Routledge.

Cash, P. J. (2018). Developing theory-driven design research. *Design Studies*, *56*, 84–119. http://doi.org/10.1016/j.destud.2018.03.002

Chalupnik, M. J., Wynn, D. C., & Clarkson, P. J. (2013). Comparison of ilities for protection against uncertainty in system design. *Journal of Engineering Design*. http://doi.org//doi.org/10.1080/09544828.2013.851783

Chang, D., Lee, C. K. M., & Chen, C. (2014). Review of life cycle assessment towards sustainable product development. *Journal of Cleaner Production*, *83*, 48–60. http://doi.org/10.1016/j.jclepro.2014.07.050

Cheng, P. C. H., Lowe, R. K., & Scaife, M. (2001). Cognitive science approaches to diagrammatic representations. *Artificial Intelligence Review*, *79–94*(15), 79–94. http://doi.org/10.1023/A:1006641024593

Chiesa, V., Coughlan, P., & Web, C. (1996). Development of a Technical Innovation Audit. *Journal of Product Innovation Management*, *13*(2), 105–136.

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, *5*, 559–583.

Clemen, R. T., & Winkler, R. L. (1999). Combining Probability Distributions From Experts in Risk Analysis. *Risk Analysis*, *19*(2).

Collis, J., & Hussey, R. (2009). Business research: a practical guide for undergraduate and postgraduate students. *Palgrave Macmillan UK*, *3rd*, 420. http://doi.org/10.1038/142410a0

Colyvan, M. (2008). Is Probability the Only Coherent Approach to Uncertainty ? *Risk Analysis*, *28*(3). http://doi.org/10.1111/j.1539-6924.2008.01058.x

Cooke, R. M. (1991). *Experts in Uncertainty: Opinion and Subjective Probability in Science*. *booksgooglecom*. http://doi.org/10.1016/0040-1625(93)90030-B

Cooke, R. M., & Goossens, L. H. J. (2000). Procedures Guide for Structural Expert Judgement in Accident Consequence Modelling. *Radiation Protection Dosimetry*, *90*(3), 303–309. http://doi.org/10.1093/oxfordjournals.rpd.a033152

Cooke, R. M., & Goossens, L. H. J. (2004). Expert judgement elicitation for risk assessments of critical infrastructures. *Journal of Risk Research*, *7*(6), 643–656. http://doi.org/10.1080/1366987042000192237

Coolen-Schrijner, P., Coolen, F. P. A., Troffaes, M. C. M., Coolen, F. P. A., & Troffaes, M. C. M. (2011). Imprecision in Statistical Theory and Practice. *Journal of Statistical Theory and Practice*, *3*(1). http://doi.org/10.1080/15598608.2009.10411907

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, *13*(1), 3–21. http://doi.org/10.1007/BF00988593

Cross, N. (2011). *Design thinking: Understanding how designers think and work*.

Crossland, R., Williams, J. S., & McMahon, C. (2003). The practical application of design risk assessment models. *Journal of Engineering Manufacture*, *217*, 227–234.

Dani, B., & Joan, G. (2004). Statistical Literacy, Reasoning, and Thinking: Goals, Definitions, and Challenges. In *he Challenge of Developing Statistical Literacy, Reasoning and Thinking*. Dordrecht: Springer.

de Weck, O. L. ., Ross, A. M., & Rhodes, D. H. (2010). Investigating Relationships and Semantic Sets amongst System Lifecycle Properties (Ilities). In *Third International Engineering Systems Symposium, CESUN*.

de Weck, O., Roos, D., & Magee, C. (2011). *Engineering Systems: Meeting Human Needs in a Complex Technological World*. MIT Press.

Dempster, A. P. (1967). Upper and Lower Probabilities Induced by a Multivalued Mapping. *The Annals of Mathematical Statistics*. http://doi.org/10.1214/aoms/1177733256

Denoeux, T. (1995). A k-nearest neighbor classification rule based on Dempster-Shafer theory. *IEEE Transactions on Systems, Man, and Cybernetics*, *25*(5), 804–813. http://doi.org/10.1109/21.376493

Denoeux, T. (2016). 40 years of Dempster–Shafer theory. *International Journal of Approximate Reasoning*, *79*, 1–6. http://doi.org/10.1016/j.ijar.2016.07.010

Dewar, J., Builder, C., Hix, W., & Levin, M. (1993). *Assumption-Based Planning: A Planning Tool for Very Uncertain Times*.

Dey, P. (2001). Decision support system for risk management: a case study. *Management Decision*, 634–649. http://doi.org/10.1108/00251740110399558

Dezfuli, H., Stamatelatos, M., Maggio, G., Everett, C., & Youngblood, R. (2010). Risk Informed Decision Making (NASA/SP-2010-576). Office of Safety and Mission Assurance, NASA Headquarters.

Dikmen, I., Birgonul, M. T., & Arikan, A. E. (2004). A critical review of risk management support tools. In *20th Annual ARCOM Conference*. 20th Annual ARCOM Conference, Heriot-Watt University, Edinburgh, UK.

DoD. (2006). Risk Management Guide for DoD Acquisition, 6th ed. United States Department of Defense, Office of the Secretary of Defense, Washington, D.C.

Dubois, D. (2010). Representation, propagation, and decision issues in risk analysis under incomplete probabilistic information. *Risk Analysis*, *30*(3), 361–368. http://doi.org/10.1111/j.1539-6924.2010.01359.x

Dubois, D., & Prade, H. (2009). Formal representations of uncertainty. In *Decision-Making Process: Concepts and Methods* (pp. 85–156).

Dubois, D., & Prade, H. (2012). Possibility Theory. *Computational Complexity*. http://doi.org/10.1007/978-1-4614-1800-9

Dunn, J. M., & Epstein, G. (1977). *Modern Uses of Multiple-Valued Logic*. Springer.

Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *Academy of*

Management Review, *14*(4), 532–550. http://doi.org/10.5465/amr.1989.4308385

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory Building from Cases: Oppurtinities and Challanges. *The Academy of Management Journal*, *50*(1), 25–32. http://doi.org/10.2307/20159839

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*.

Eweje, J., Turner, R., & Müller, R. (2012). Maximizing strategic value from megaprojects: The influence of information-feed on decision-making by the project manager. *International Journal of Project Management*, *30*(6), 639–651. http://doi.org/10.1016/j.ijproman.2012.01.004

Ferson, S. (1996). What Monte Carlo Methods Cannot Do. *Human and Ecological Risk Assessment*, *2*(4), 990–1007.

Ferson, S., & Ginzburg, L. R. (1996). Different methods are needed to propagate ignorance and variability. *Reliability Engineering and System Safety*, *54*(2–3), 133–144. http://doi.org/10.1016/S0951-8320(96)00071-3

Ferson, S., Ginzburg, L., & Akcakaya, R. (1996). Whereof one cannot speak: when input distributions are unknown. *Risk Analysis*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.7668&rep=rep1&type=pdf

Ferson, S., Kreinovich, V., Ginzburg, L., Myers, D. S., & Sentz, K. (2003). Constructing Probability Boxes and Dempster-Shafer Structures. *Sandia National Laboratories*, (January), 143.

Fischhoff, B. (1995). Risk Perception and Commuication Unpluged: Twenty Years of Process. *Risk Analysis*, *15*(2), 137–145. http://doi.org/10.1111/j.1539-6924.1995.tb00308.x

Flage, R., & Terje, A. (2009). Expressing and communicating uncertainty and bias in relation to Quantitative Risk Analysis. *Safety and Reliability: Methodology and Applications*, *2*(13), 1691–1699. http://doi.org/10.1201/b17399-230

Flage, R., Aven, T., Zio, E., & Baraldi, P. (2014). Concerns, Challenges, and Directions of Development for the Issue of Representing Uncertainty in Risk Assessment. *Risk Analysis*, *34*(7), 1196–1207. http://doi.org/10.1111/risa.12247

Flyvbjerg, B. (2006a). Five misunderstandings about case-study research. *Qualitative Inquiry*, 420–434. http://doi.org/10.1177/1077800405284363.CITATIONS

Flyvbjerg, B. (2006b). From Nobel Prize To Project Management: Getting Risks Right. *Project Management Journal*, *37*(3), 5–15. http://doi.org/10.1002/smj.476

Flyvbjerg, B. (2007). Truth and Lies about Megaprojects. *Delft University of Technology, September*, *2002*, 1–27. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Truth+and+Lies+About+Megaprojects+Inaugural+speech#1

Flyvbjerg, B., Holm, M. S., & Buhl, S. (2007). Underestimating Costs in Public Works

Projects: Error or Lie? *Journal of the American Planning Association*.

Flyvbjerg, B., Morris, P. W. G., Pinto, J. K., & Söderlund, J. (2013). Over budget, over time, over and over again: Managing Major Projects, *The Oxford Handbook of Project Management*, Oxford: Oxford University Press, pp. 321–344. http://doi.org/10.1093/oxfordhb/9780199563142.003.0014

Fontoura, L., & Price, R. T. (2008). Systematic Approach to Risk Management in Software Projects through Process Tailoring. In *Combining SOA and BPM Technologies for Cross-System Process Automation*.

Fortin, M. J., & Gagnon, C. (2006). Interpreting major industrial landscapes: Social follow-up on meanings, the case of two aluminium smelters, Alcan (Alma, Canada) and Pechiney (Dunkirk, France). *Environmental Impact Assessment Review*, *26*(8), 725–745. http://doi.org/10.1016/j.eiar.2006.06.002

Funtowicz, S., & Ravetz, J. (1990). *Uncertainty and quality in science for policy research*. Kluwer Academic Publishers.

Gallagher, B. P. (1999). Software Acquisition Risk Management Key Process Are (KPA) - A Guidebook. Software Engineering Institute, Pittsburgh, PA.

Geraldi, J. G., Lee-Kelley, L., & Kutsch, E. (2010). The Titanic sunk, so what? Project manager response to unexpected events. *International Journal of Project Management*, *28*(6), 547–558. http://doi.org/10.1016/j.ijproman.2009.10.008

Gericke, K. (2011). *Enhancing Project Robustness: A Risk Management Perspective*. Berlin, Germany.

Gibbert, M., Ruigrok, W., & Wicki, B. (2008). What passes as a rigorous case study? *Strategic Management Journal*, *29*(13), 1465–1474. http://doi.org/10.1002/smj.722

Gidel, T., Gautier, R., & Duchamp, R. (2005). Decision-making framework methodology: an original approach to project risk management in new product design. *Journal of Engineering Design*, *16*(1), 1–23. http://doi.org/10.1080/09544820512331325238

Glendon, A. I., Clarke, S., & McKenna, E. (2016). *Human Safety and Risk Management* (Second edi). CRC Press.

Goerlandt, F., & Reniers, G. (2015). On the assessment of uncertainty in risk diagrams. *Safety Science*, *84*, 67–77. http://doi.org/10.1016/j.ssci.2015.12.001

Goh, Y. M., McMahon, C. A., & Booker, J. D. (2007). Development and characterisation of error functions in design. *Research in Engineering Design*, *18*(3), 129–148. http://doi.org/10.1007/s00163-007-0034-x

Goodman, I. R., & Nguyen, H. T. (1999). Probability updating using second order probabilities and conditional event algebra. *Information Sciences*, *121*(3–4), 295–347. http://doi.org/10.1016/S0020-0255(99)00089-4

Groves, D. G., & Lempert, R. J. (2007). A new analytic method for finding policy-relevant Scenarios. *Global Environ. Change*, *17*(73–85).

Grubisic, V. V. F., Gidel, T., & Ogliari, A. (2011). Recommendations for risk identification method selection according to product design and project management maturity, product innovation degree and project team. In *International Conference On Engineering Design, ICED11*.

Guba, E. G. (1981). Criteria for Assessing the Trustworthiness of Naturalistic Inquiries. *ERIC/ECTJ Annual Review PaperJ Annual Review Paper*, *29*(4), 75–91. http://doi.org/10.1126/science.146.3642.347

Guyonnet, D., Bourgine, B., Dubois, D., Fargier, H., Côme, B., & Chilès, J. P. (2003). Hybrid approach for addressing uncertainty in risk assessments. *Journal of Environmental Engineering*, *129*(1), 68.

Haasnoot, M., Kwakkel, J. H., & Walker, W. E. (2013). Dynamic adaptive policy pathways : A method for crafting robust decisions for a deeply uncertain world. *Global Environmental Change*, *23*(2), 485–498. http://doi.org/10.1016/j.gloenvcha.2012.12.006

Haasnoot, M., Middelkoop, H., Offermans, A., Beek, E. Van, & Deursen, W. P. A. Van. (2012). Exploring pathways for sustainable water management in river deltas in a changing environment. *Climatic Change*, 795–819. http://doi.org/10.1007/s10584-012-0444-2

Hamarat, C., Kwakkel, J. H., & Pruyt, E. (2013). Technological Forecasting & Social Change Adaptive Robust Design under deep uncertainty. *Technological Forecasting & Social Change*, *80*(3), 408–418. http://doi.org/10.1016/j.techfore.2012.10.004

Hammer, J. (2015). How Berlin Airport Became a $6 Billion Embarrassment. *Bloomberg Business*. Retrieved from https://www.bloomberg.com/news/features/2015-07-23/how-berlin-s-futuristic-airport-became-a-6-billion-embarrassment

Harbaugh, J. W., Davis, J. C., & Wendebourg, J. (1995). *Computing Risk for Oil Prospects: Principles and Programs*. Pergamon.

Hargreaves, R. (1971) *Mr. Tickle.* Egmont Publishing.

Hellweg, S., & Canals, L. M. I. (2014). Emerging approaches, challenges and opportunities in life cycle assessment. *Science, 344*(6188), 1109–1114.

Helton, J. C. (1993). Risk, uncertainty, and the EPA release limits for radioactive-waste disposal. *Nuclear Technology*, *101*(1), 18–39.

Helton, J. C., & Burmaster, D. E. (1996). Guest editorial: treatment of aleatory and epistemic uncertainty in performance assessments for complex systems. *Reliability Engineering & System Safety*, *54*(2–3), 91–94. http://doi.org/10.1016/S0951-8320(96)00066-X

Helton, J. C., & Oberkampf, W. L. (2004). Alternative representations of epistemic uncertainty. *Reliability Engineering and System Safety*, *85*(1–3), 1–10. http://doi.org/10.1016/j.ress.2004.03.001

Herman, J., Harrison, Z., Reed, P., & Characklis, G. (2014). Beyond optimality: Multistakeholder robustness tradeoffs for regional water portfolio planning under deep uncertainty. *Water Resources Research*, 7692–7713. http://doi.org/10.1002/2014WR015338

Herrmann, W. J., Guenther, A., Oehmen, J., Schulte, J., Sörensen, C. A., & Willumsen, P. L. (2018). Risk Management in Product Development: Risk Identification, Assessment, and Mitigation. *Tbp.*

Hillson, D. (1997). Towards a risk maturity model. *The International Journal of Project and Business Risk Management*.

Hogg, R. V., McKean, J., & Craig, A. T. (2005). *Introduction to Mathematical Statistics (6th Edition).* All Books and Monographs by WMU Authors. Retrieved from https://scholarworks.wmich.edu/books/119

Holton, G. A. (2004). Defining Risk. *Financial Analysts Journal*, *60*(6).

Hubbard, D. W. (2009). *The Failure of Risk Management: Why It's Broken and How to Fix It*. Hoboken, NJ: John Wiley and Sons, pp. 281.

Humphrey, W. S. (1988). Characterizing the software process: a maturity framework. *IEEE Software*, *5*(2). http://doi.org/10.1109/52.2014

INCOSE. (2007). Systems Engineering Handbook v3.1. Internation Council on Systems Engineering, Seattle.

Institute, A. N. S. (2011). American National Standard. *ANSI/ASSE/IEC/ISO 31010 (Z690.3-2011) Risk Assessment Techniques*, *31010*, 545.

Ip, H. H. S., & Ng, J. M. C. (1994). Human face recognition using Dempster-Shafer theory. In *IEEE International Conference on Image Processing* (Vol. 1, pp. 292–295).

ISO. (2009). ISO 31000:2009(E) - Risk Management - Principles and Guidelines. International Organization for Standardization, Geneva.

Jaynes, E. T. (2003). *Probability theory: the logic of science*. Cambridge University Press.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, *33*(7), 14–26. http://doi.org/10.3102/0013189X033007014

Kaplan, R. S., & Mikes, A. (2012). Managing Risks: A New Framework. *Harvard Business Review*, *90*(6).

Kaplan, R. S., & Garrick, B. J. (1981). On the quantitative definition of risk. *Risk Analysis*.

Klinke, A., & Renn, O. (2002). A new approach to risk evaluation and management: Risk-based, precaution-based, and discourse-based strategies. *Risk Analysis*, *22*(6), 1071–1094. http://doi.org/10.1111/1539-6924.00274

Knight, F. H. (1921). *Risk, uncertainty and profit.* New York: Houghton Mifflin Company (republished in 2006 by Dover Publications, Mineola, NY).

Kozin, I. (2017). Lecture notes: Risk and decision-making, DTU 42172.

Kozin, I., & Petersen, K. E. (1996). Obtaining Reliability and Safety Assessments on the Basis of Non-Probabilistic Methods. In *Intelligent Systems and Soft Computing for Nuclear Science and Industry. 2 International FLINS Workshop*.

Kozine, I., & Utkin, L. (2002). Processing unreliable judgements with an imprecise hierarchical model. *Risk Decision and Policy*, *7*, 325–339.

Kozine, I., & Utkin, L. (2005). Constructing imprecise probability distributions. *International Journal of General Systems*, *34*(4), 401–408.

Kreye, M. E. (2011). *Uncertainty Analysis in Competitive Bidding for Service Contracts*. University of Bath.

Kujawski, E., & Angelis, D. (2009). Monitoring risk response actions for effective project risk management. *Systems Engineering*, n/a-n/a. http://doi.org/10.1002/sys.20154

Kulikova, N. (2016). *Managing risks in large engineering projects: tools, standards and practices*. Technical University of Denmark.

Kvale, S. (1996). The Interview Situation. *Interviews: An Introduction to Qualitative Research Interviewing*. http://doi.org/10.1037/0022-3514.90.4.644

Kvale, S. (2008). *Doing interviews*. SAGE Publications.

Kwakkel, J. H., & Jaxa-rozen, M. (2016). Improving scenario discovery for handling heterogeneous uncertainties and multinomial classified outcomes. *Environmental Modelling and Software*, *79*, 311–321. http://doi.org/10.1016/j.envsoft.2015.11.020

Kwakkel, J. H., Walker, W. E., & Marchau, V. A. W. J. (2010). Classifying and communicating uncertainties in model-based policy analysis, *10*(4).

Kwiatkowska, M., Norman, G., & Parker, D. (2011). Verification of Embedded Real-time Systems. In *23rd International Conference on Computer Aided Verification (CAV'11)* (pp. 1–25). http://doi.org/10.1007/978-3-658-09994-7_1

Kyburg, H. E. (1987). Bayesian and non-bayesian evidential updating. *Artificial Intelligence*, *31*(3), 271–293. http://doi.org/10.1016/0004-3702(87)90068-3

Lakoff, G. (1975). Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts. In *Contemporary Research in Philosophical Logic and Linguistic Semantics* (pp. 221–271). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-94-010-1756-5_9

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*(1), 65–100.

Lavasani, M. R., Wang, J., Yang, Z., & Finlay, J. (2011). Application of Fuzzy Fault Tree Analysis on Oil and Gas Offshore Pipelines. *International Journal of Marine Science and Engineering*, *1*(1), 29–42.

Lejewski, C., & Łukasiewicz, J. (1967). Encyclopedia of Philosophy. In *MacMillan* (pp. 104–107).

Lempert, R. J. (2002). *A new decision sciences for complex systems* (Vol. 99). Santa Monica.

Lempert, R. J., Groves, D. G., Popper, S. W., & Bankes, S. C. (2006). A General, Analytic Method for Generating Robust Strategies and Narrative Scenarios. *Management Science*, *52*(4), 514–528.

Lempert, R. J., Popper, S., & Bankes, S. (2003). *Shaping the Next One Hundred Years: New Methods for Quantitative, Long Term Policy Analysis*.

Lempert, R. J., Schlesinger, M. E., & Bankes, S. C. (1996). When we don't know the cost or the benefits: Adaptive strategies for abating climate change. *Climatic Change*, *33*, 235–274.

Lempert, R. J., Sriver, R. L., & Keller, K. (2012). *Characterizing Uncertain Sea Level Rise Projections to Support Investment Decisions*. Sacramento.

Levi, I. (1990). *Hard choices. Decision Making under Unresolved Conflict*. Cambridge University Press.

Lincoln, Y. S., & Guba, E. G. (1986). But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. *New Directions for Program Evaluation*, *1986*(30), 73–84. http://doi.org/10.1002/ev.1427

Locatelli, G. (2018). Why are Megaprojects, Including Nuclear Power Plants, Delivered Overbudget and Late? Reasons and Remedies. *Center for Advanced Nuclear Energy Systems (CANES), Massachusetts Institute of Technology*, *2018*.

Locatelli, G., & Mancini, M. (2010). Risk management in a mega-project: the Universal EXPO 2015 case, *2*(3), 236–253. http://doi.org/10.1504/IJPOM.2010.035342

Loch, C. (2000). Tailoring Product Development to Strategy : Case of a European Technology Manufacturer. *European Management Journal*, *18*(3), 246–258.

Lough, K. G., Stone, R., & Tumer, I. Y. (2009). The risk in early design method. *Journal of Engineering Design*, *20*(2), 155–173. http://doi.org/10.1080/09544820701684271

Lovallo, D., & Kahneman, D. (2003). Delusions of Success. *Harvard Business Review*, *5*, 5–10.

Love, P. E. D., Edwards, D. J., & Irani, Z. (2012). Moving beyond optimism bias and strategic misrepresentation: An explanation for social infrastructure project cost overruns. *IEEE Transactions on Engineering Management*, *59*(4), 560–571. http://doi.org/10.1109/TEM.2011.2163628

Maier, A. M., Moultrie, J., & Clarkson, P. J. (2012). Assessing Organizational Capabilities : Reviewing and Guiding the Development of Maturity Grids. *IEEE Transactions on Engineering Management*, *59*(1), 138–159.

Malcolm, G. (2005). *Blink: The Power of Thinking Without Thinking*. Little Brown & Company 13.

Marle, F., & Gidel, T. (2012). Assisting Project Risk Management Method Selection. *International Journal of Project Organisation and Management*.

Matrosov, E. S., Woods, A. M., & Harou, J. J. (2013). Robust Decision Making and Info-Gap Decision Theory for water resource system planning. *Journal of Hydrology*, *494*, 43–58. http://doi.org/10.1016/j.jhydrol.2013.03.006

Merrow, E. W. (2013). *Industrial Megaprojects: Concepts, Strategies, and Practices for Success*. Wiley.

Morris, S. (1990). Cost and Time Overruns in Public Sector Projects. *Economic and Political Weekly*, *15*, 154–168.

Mosleh, A., Bier, V. M., & Apostolakis, G. (1988). A Critique of Current Practice for the Use of Expert Opinions in Probabilistic Risk Assessment. *Reliability Engineering & System Safety*, *20*, 63–85.

Mulqueen, J., Maples, C. D., & Fabisinski, L. (2012). Tailoring Systems Engineering Processes in a Conceptual Design Environment : a case study at NASA Marshall Spaceflight Center 's ACO. In *INCOSE*.

Nilsen, T., & Aven, T. (2002). Models and model uncertainty in the context of risk analysis. *Reliability Engineering & System Safety*, *79*(3), 309–317. http://doi.org/10.1016/S0951-8320(02)00239-9

Nisbett, R. (1993). *Rules for Reasoning*. New York: Psychology Press.

Oberkampf, W. L., Helton, J. C., Joslyn, C. A., Wojtkiewicz, S. F., & Ferson, S. (2004). Challenge problems: Uncertainty in system response given uncertain parameters. *Reliability Engineering and System Safety*, *85*(1–3), 11–19. http://doi.org/10.1016/j.ress.2004.03.002

Oehmen, J., Olechowski, A., Robert Kenley, C., & Ben-Daya, M. (2014). Analysis of the effect of risk management practices on the performance of new product development programs. *Technovation*, *34*(8), 441–453. http://doi.org/10.1016/j.technovation.2013.12.005

Otway, H., & Winterfeldt, D. von. (1992). Expert Judgment in Risk Analysis and Management: Process, Context, and Pitfalls. *Risk Analysis*. http://doi.org/doi.org/10.1111/j.1539-6924.1992.tb01310.x

Ouchi, F. (2004). *A Literature Review on the Use of Expert Opinion in Probabilistic Risk Analysis* (No. 3201). Washington, DC. Retrieved from https://openknowledge.worldbank.org/handle/10986/15623

Paté-Cornell, M. E. (1996). Uncertainties in risk analysis: Six levels of treatment. *Reliability Engineering & System Safety*, *54*(2–3), 95–111. http://doi.org/10.1016/S0951-8320(96)00067-1

Paulzen, O., Doumi, M., Perc, P., & Cereijo-Roibas, A. (2002). A Maturity Model for Quality Improvement in Knowledge Management. In *ACIS*.

Pawson, R., Wong, G., & Owen, L. (2011). Known Knowns, Known Unknowns, Unknown Unknowns. *American Journal of Evaluation*, *32*(4), 518–546. http://doi.org/10.1177/1098214011403831

Pender, S. (2001). Managing incomplete knowledge: Why risk management is not sufficient. *International Journal of Project Management*, *19*(2), 79–87. http://doi.org/10.1016/S0263-7863(99)00052-6

Pennypacker, J. S., & Grant, K. (2003). Project Management Maturity: An Industry Benchmark. *Project Managemet Journal*, *34*.

PMI. (2008). A Guide to the Project Management Body of Knowledge (PMBOK GUIDE), 4th ed. Project Management Institute, Drexel Hill, PA.

PMI. (2016). *PMI's 2016 Pulse of the Profession: The High Cost of Low Performance*.

Poirier, D. J. (2014). Frequentist and Subjectivist Perspectives on the Problems of Model Building in Economics. *Journal of Economic Perspectives*, *20*(1), 25–46. http://doi.org/10.1257/jel.48.2.281

Prelec, D., & Loewenstein, G. (1991). Decision Making Over Time and Under Uncertainty: A Common Approach. *Management Science*, *37*(7), 770–786. http://doi.org/10.1287/mnsc.37.7.770

Ramas. (2016). Risk Imaging. Retrieved from http://www.ramas.com/riskimaging

Ramsey, F. P. (2009). On a Problem of Formal Logic. In *Classic Papers in Combinatorics* (pp. 1–24). Boston, MA: Birkhäuser Boston. http://doi.org/10.1007/978-0-8176-4842-8_1

RAND Corporation. (2009). *Data Collection Methods. Semi-Structured Interviews and Focus Groups*.

Raz, T., & Hillson, D. (2005). A Comparative Review of Risk Management Standards. *Risk Management*, *7*(4), 53–66.

Raz, T., & Michael, E. (2001). Use and benefits of tools for project risk management. *International Journal of Project Management*, *19*(1), 9–17. http://doi.org/10.1016/S0263-7863(99)00036-8

Ren, Y. T., & Yeo, K. T. (2004). Risk Management Capability Maturity Model for Complex Product Systems (CoPS) Projects. In *International Engineering Management Conference* (pp. 807–811).

Renn, O. (1998). Three decades of risk research: accomplishments and new challenges. *Journal of Risk Research*, (September 2015). http://doi.org/10.1080/136698798377321

Renn, O., Klinke, A., & Van Asselt, M. (2011). Coping with complexity, uncertainty and ambiguity in risk governance: A synthesis. *Ambio*, *40*(2), 231–246. http://doi.org/10.1007/s13280-010-0134-0

Riabacke, A. (2006). Managerial Decision Making Under Risk and Uncertainty. *International Journal of Computer Science*, *32*(November), 453–459. Retrieved from http://www.iaeng.org/IJCS/issues_v32/issue_4/IJCS_32_4_12.pdf

RISKSIG Risk Management Special Interest Group, PMI (2002), Risk Management Maturity Model Version 1. [online] Available at: http://www.risksig.com/articles/rm%20report%20final%20version%2012.doc

Ricci, N., Fitzgerald, M. E., Ross, A. M., & Rhodes, D. H. (2014). Architecting Systems of Systems with Ilities : an Overview of the SAI Method. *Procedia - Procedia Computer Science*, *28*(Cser), 322–331. http://doi.org/10.1016/j.procs.2014.03.040

Rowe, F. (2014). What literature review is not: diversity, boundaries and recommendations. *European Journal of Information Systems*, *23*(3), 241–255.

http://doi.org/10.1057/ejis.2014.7

Rowe, G., & Wright, G. (1999). The Delphi technique as a forecasting tool: issues and analysis. *International Journal of Forecasting*.

Rowe, D. (1975). *An "Anatomy" of Risk*. Environmental Protection Agency Washington, D. C.

Sahlin, P. N. E. (2012). Unreliable Probabilities, Paradoxes, and Epistemic Risks. In *Handbook of Risk Theory* (pp. 477–498). http://doi.org/10.1007/978-94-007-1433-5

Schoemaker, P. (1995). Scenario Planning : A Tool for Strategic Thinking. *Sloan Management Review*, (January 1995).

Sentz, K., & Ferson, S. (2002). Combination of Evidence in Dempster- Shafer Theory. *Contract*, (April), 96. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.122.7929&rep=rep1&type=pdf

Seong, S., Popper, S., & Zheng, K. (2005). *Strategic Choices in Science and Technology: Korea in the Era of a Rsing China;* Santa Monica.

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. *The Journal of Finance*, *19*(3), 425–442. http://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Sims, C. A. (2001). Pitfalls of a minimax approach to model uncertainty. *American Economic Review*, *91*(2), 51–54. http://doi.org/10.1257/aer.91.2.51

Singh, S. P., & Singh, V. K. (2011). Three-level AHP-based heuristic approach for a multi-objective facility layout problem. *International Journal of Production Research*, *49*(4), 1105–1125. http://doi.org/10.1080/00207540903536148

Škec, S., Štorga, M., Rohde, D., & Marjanović, D. (2014). Tailoring risk management approach for the product development environment. In *International Design Conference - Design 2014* (pp. 385–396).

Smithson, M. (1989). *Ignorance and Uncertainty: Emerging Paradigms*. Springer-Verlag.

SNL. (2016). Sandia National Laboratories. Livermore, California. Retrieved from https://www.sandia.gov/

Speirs-Bridge, A., Fidler, F., McBride, M., Flander, L., Cumming, G., & Burgman, M. (2010). Reducing overconfidence in the interval judgments of experts. *Risk Analysis*, *30*(3), 512–523. http://doi.org/10.1111/j.1539-6924.2009.01337.x

Swart, R., Bernstein, L., Ha-Duong, M., & Petersen, A. (2009). Agreeing to disagree: Uncertainty management in assessing climate change, impacts and responses by the IPCC. *Climatic Change*, *92*(1–2), 1–29. http://doi.org/10.1007/s10584-008-9444-7

Swiler, L., Paez, T., & Mayes, R. (2009). Epistemic uncertainty quantification tutorial. *Proceedings of the 27th International Modal Analysis Conference*.

Takata, S., & Umeda, Y. (2007). Advances in Lyfe Cycle Engineering for Sustainable

Manufacturing Businesses. In *14th CIRP Conference on Life Cycle Engineering*.

Taroun, A. (2014). Towards a better modelling and assessment of construction risk: Insights from a literature review. *International Journal of Project Management*, *32*(1), 101–115. http://doi.org/10.1016/j.ijproman.2013.03.004

Taroun, A, & Yang, J. B. (2013). A DST-based approach for construction project risk analysis. *Journal of the Operational Research Society*, *64*(8), 1221–1230. http://doi.org/10.1057/jors.2013.38

Tegeltija, M., Oehmen, J., Kozin, I., & Kwakkel, J. H. (2018a). Exploring Deep Uncertainty Approaches for Application in Life Cycle Engineering. In *Procedia CIRP* (Vol. 69). http://doi.org/10.1016/j.procir.2017.12.006

Tegeltija, M., Oehmen, J., McMahon, C. A., Maier, A., Kozin, I., & Škec, S. (2018b). Tailoring Risk Management In Design. In *Proceedings of the DESIGN 2018 15th International Design Conference* (pp. 667–678). http://doi.org/10.21278/idc.2018.0385

Thissen, W., Kwakkel, J. H., & Mens, M. (2017). Dealing with Uncertainties in Fresh Water Supply : Experiences in the Netherlands. *Water Resources Management*, 703–725. http://doi.org/10.1007/s11269-015-1198-1

Thomas, G. (2015). *How to do your case study*. London: SAGE Publications.

Tsoukiàs, A. (2008). From decision theory to decision aiding methodology. *European Journal of Operational Research*.

Unger, D., & Eppinger, S. (2011). Improving product development process design: a method for managing information flows , risks, and iterations. *Journal of Engineering Design*, *4828*. http://doi.org/10.1080/09544828.2010.524886

van der Sluijs, J. P., Craye, M., Funtowicz, S., Kloprogge, P., Ravetz, J., & Risbey, J. (2005). Combining quantitative and qualitative measures of uncertainty in model-based environmental assessment: the NUSAP system. *Risk Analysis : An Official Publication of the Society for Risk Analysis*, *25*(2), 481–492. http://doi.org/10.1111/j.1539-6924.2005.00604.x

van Asselt, M. B. A., & Rotmans, J. (2002). Uncertainty in Integrated Assessment Modelling. *Climatic Change*. http://doi.org/10.1023/A:1015783803445

Walker, W. E., Harremoes, P., Rotmans, J., Sluijs, J. P. van der, Asselt, M. B. A. van, Janssen, P., & M.P. Krayer von Krauss. (2003). Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment*.

Walker, W. E., Haasnoot, M., & Kwakkel, J. H. (2013). Adapt or Perish: A Review of Planning Approaches for Adaptation under Deep Uncertainty. *Sustainability*, 955–979. http://doi.org/10.3390/su5030955

Walker, W. E., Lempert, R. J., & Kwakkel, J. H. (2013). Deep uncertainty. In *Encyclopedia of Oper. Res. and Management Science*. http://doi.org/10.1007/978-1-4419-1153-7

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman & Hall.

Walley, P. (1996). Measures of uncertainty in expert systems. *Artificial Intelligence*, *83*(1), 1–58. http://doi.org/10.1016/0004-3702(95)00009-7

Walley, P. (1999). Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, *24*(2–3), 125–148. http://doi.org/10.1016/S0888-613X(00)00031-1

Wardekker, J. A., van der Sluijs, J. P., Janssen, P. H. M., Kloprogge, P., & Petersen, A. C. (2008). Uncertainty communication in environmental assessments: views from the Dutch science-policy interface. *Environmental Science and Policy*, *11*(7), 627–641. http://doi.org/10.1016/j.envsci.2008.05.005

Weaver, C. P., Lempert, R. J., Brown, C., Hall, J. A., Revell, D., & Sarewitz, D. (2013). Improving the contribution of climate model information to decision making: the value and demands of robust decision frameworks. *Wiley Interdisciplinary Reviews: Climate Change*, *4*(1), 39–60. http://doi.org/10.1002/wcc.202

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future : Writing a Literature Review, *26*(2).

Weichselberger, K. (2000). The theory of interval-probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning*, *24*(2–3), 149–170. http://doi.org/10.1016/S0888-613X(00)00032-3

Werner, C., Bedford, T., Cooke, R. M., Hanea, A. M., & 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. http://doi.org/10.1016/j.ejor.2016.10.018

Wieland, A., & Wallenburg, C. M. (2012). Dealing with supply chain risks: Linking risk management practices and strategies to performance. *International Journal of Physical Distribution & Logistics Management*, *42*(10), 887–905. http://doi.org/10.1108/09600031211281411

Wilby, R. L., & Dessai, S. (2010). Robust adaptation to climate change. *Weather*, *65*(7), 180–185. http://doi.org/10.1002/wea.504

Xiao, M., Gao, L., Xiong, H., & Luo, Z. (2015). An efficient method for reliability analysis under epistemic uncertainty based on evidence theory and support vector regression. *Journal of Engineering Design*, *26*(10–12), 340–364. http://doi.org/10.1080/09544828.2015.1057557

Yen, J. (1989). GERTIS: a Dempster-Shafer approach to diagnosing hierarchical hypotheses. *Communications of the ACM*.

Yin, R. K. (2013). Case study research: Design and methods. *Applied Social Research Methods Series*, *18*(2), 54–56. http://doi.org/10.1097/00001610-199503000-00004

Zadeh, L. A. (1978). Fuzzy Sets As a Basis for a Theory of Possibility. *Fuzzy Sets and Systems*, *1*, 3–28.

Zio, E. (2007). From complexity science to reliability efficiency: A new way of looking at complex network systems and critical infrastructures. *International Journal of Critical*

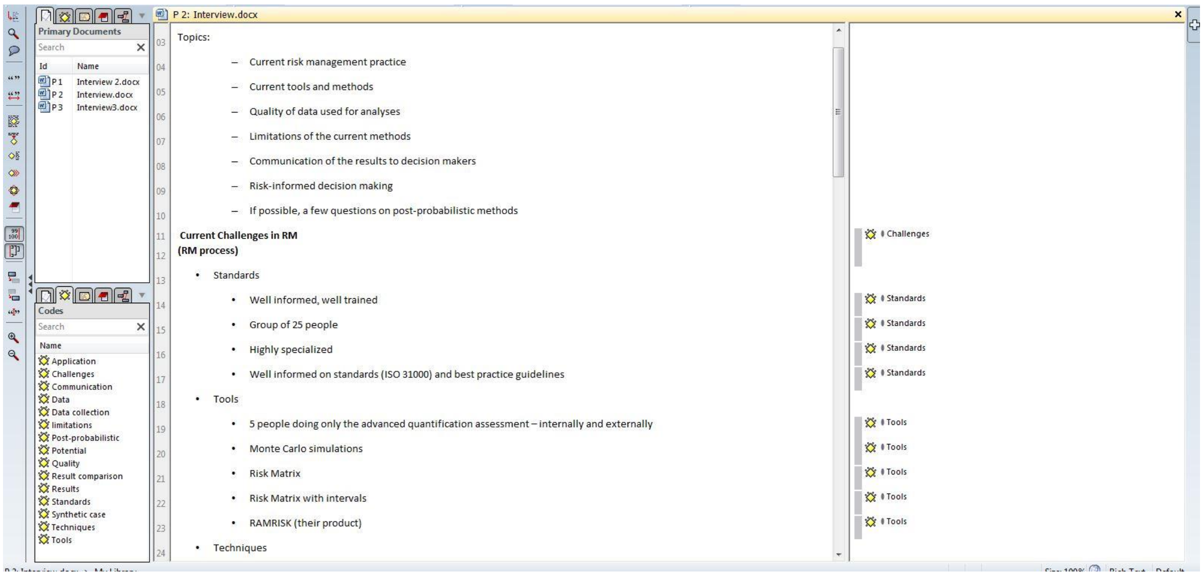
Infrastructures, *3*(3–4), 488–508.

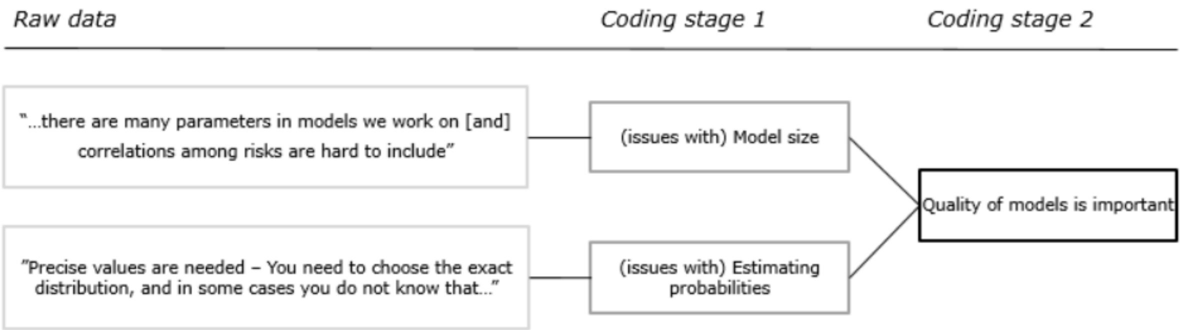
Zio, E. (2009). Reliability engineering: Old problems and new challenges. *Reliability Engineering and System Safety*, *94*(2), 125–141. http://doi.org/10.1016/j.ress.2008.06.002

Zwikael, O., & Ahn, M. (2011). The Effectiveness of Risk Management: An Analysis of Project Risk Planning Across Industries and Countries. *Risk Analysis*, *31*(1), 25–37. http://doi.org/10.1111/j.1539-6924.2010.01470.x

Appendix 1: Coding preview

Interviews' coding preview





Appendix 2: Interview Guides

Interview guide for documenting challenges in risk management practice	
Introduction	I introduce myself and my background. I briefly present the PhD project. This includes an explanation of the expectations of the interview and plans for its use.
Background of the interviewee	Can you give me a short introduction to your current position (including exact title, hierarchical level, etc.)? Also about your professional background, including examples of previous projects, tasks, duties? What are your current roles and responsibilities in the organization?
Previous experience with risk management	Are you aware of risk management standards and procedures? Have you had any prior experience with risk management? What was it? Educational level—ask for details if relevant. Have you had any training on the topic?
Design (solution) in question	What kind of design work does your organization do? Who are the users of your solutions? (Affecting whom?) What is the scope/budget and lifetime of your projects? (Portfolio of design activities.) Do you cover an entire system's life-cycle or only the design? The project you are currently working on is in which design phase?
Risk management in the organization	Can you briefly present how risk management is performed in your practice? Do you rely on the current best-practice recommendations (ISO 31000, PMBok guide, etc.)? How applicable are those recommendations to your practice? How would you describe the position of risk management (risk managers) in your organization? (The 1st, 2nd or 3rd line of defense?) Who is in the team? Who do they report to? On what? How often? In which form (written reports, presentations, oral-dialog)?
Risk quantification (Risk and uncertainty modelling)	In your practice do you rely on any quantification methods? If yes, which ones? If not, what do you use instead? What are the tools you currently use? Do you use one or more methods? How many of you work on it? Can you share more about your experiences with employing the method? Do you focus on representing epistemic uncertainty?
Data collection	Data availability: What kind of data do you use in your assessments? How are they collected? Who has the access to it? How sensitive is it?
	Data quality: Do you consider how reliable your data are?

	How do you challenge the accuracy of the experts' estimates (surveys, measurements or other ways in which you collect your data)?
	Lack of data: In case of missing data, what do you do? Do you sometimes experience time pressures? If so, how does that affect your modeling process and/or data quality? Do you have means (tools, protocols, visualizations) to report and separately discuss the quality of data on which you performed the analyses (within the risk management team and with decision makers)?
Limitations of the chosen approach	What are the challenges with the method you are using? Please elaborate. What additional features of the method would you like/need to have? How much time do you need to perform your assessment? What resources you need in order to do so? Do you have managerial support for collecting all the information you need?
Communication of the results	How do you prepare for the communication of the results? What is your main concern when doing so, what do you try to highlight? What is the main message you send? Only the top risks or the overall findings? Or more? Is a discussion on data included? Is the reliability of results considered? What is your biggest challenge when communicating the result?
Risk-informed decision making	How are the results of your assessment taken into account? What challenges do you experience? Do you experience any follow-ups, meaning requests for additional research/assessment/analyses? What other aspects can impact decision making?
Other (Some of the often recognized, additional, questions)	Do you store your data and knowledge as some form of historical data? Is there a possibility for knowledge sharing from project to project? What additional challenges have you noticed in today's risk management? If you had the opportunity to choose, would you change the tool you are using? What happens if you are dealing with a first-of-a-kind solution, for which you do not have prior experience? Please describe how you approach modeling (assessment) in such a case. Did you try using other methods, for instance Monte Carlo simulation? What are your experiences? Do you consider having insurance to cover some of the risks materializing? Who decides on when and which ones? Did you try using other methods, for instance Monte Carlo simulation? What are your experiences? Do you consider having insurance to cover some of the risks materializing? Which ones, who and when decides on that?

Interview guide for the risk management tailoring approach	
Introduction (shortened if the interviewee was previously contacted)	I introduce myself and my background. I briefly present the PhD project. This includes an explanation of the expectations of the interview and plans for its use.
Background of the interviewee (shortened if the interviewee was previously contacted)	Can you give me a short introduction to your current position (including exact title, hierarchical level, etc.)? Also about your professional background, including examples of previous projects, tasks, duties? What are your current roles and responsibilities in the organization?
Previous experience with risk management (shortened if the interviewee was previously contacted)	Are you aware of risk management standards and procedures? Have you had prior experience with risk management? What was it? Educational level – ask for details if relevant. Have you had any training on the topic?
Area of work and design (solution) in question	What kind of design work does your organization do? Who are the users of your solutions? (Affecting whom?) What is the scope/budget and lifetime of your projects? (Portfolio of design activities). Do you cover an entire system's life-cycle or only the design? The project you are currently working on is in which design phase?
Challenges with the type of design	What are the challenges with the solutions you are developing (in your organization)? How much novelty or new technology do they include? Do you have first-of-a-kind solutions? How dynamic is the market? How competitive is the field? What is the time horizon of the solutions that you are developing? How do you integrate and prioritize your activities? How early in the process do you take risk and uncertainty into account? What are the milestones?
Risk management in the organization	Can you briefly present how risk management is performed in your practice? Do you rely on the current best-practice recommendations (ISO 31000, PMBok guide, etc.)? How applicable are those recommendations to your practice? How would you describe the position of risk management (risk managers) in your organization? (The 1st, 2nd or 3rd line of defense?) Who is in the team? Who do they report to? On what? How often? In which form (written reports, presentations, oral-dialog)? Do you benchmark your practice to others? Do you use maturity models? How? How is risk management integrated into the overall organizational structure?

Details regarding the approach	An explanation from my side. Figure 17 Relationship of maturity categories and the ISO 31000 risk management process, from this thesis.
Understanding of the needs	Do you consider the broader impact and integration of risk management in your organization? How are resources distributed and how much is allocated for this type of analysis (in terms of employees, software, access to different information, etc.)? Do you discuss project types and sizes? Do they differ in terms of risk management needs? Do you distinguish between aleatory and epistemic uncertainty? Do you focus on representing epistemic uncertainty? What are your needs in terms of quantification?
Method sophistication for risk quantification (Risk and uncertainty modelling)	In your practice do you rely on any quantification methods? If yes, which ones? If not, what do you use instead? What are the current tools you use? Do you use one or more methods? How many of you work on it? Can you share more about your experiences with employing the method? How did you choose that approach? What did you use before? Did you try any other methods? How much time do you need to perform your assessment? How computationally demanding (mathematically sophisticated) is it?
Quality of data	What kind of data do you use in your assessments? How are they collected? Who has the access to it? How sensitive is it? What kinds of measurements (and data collection methods) do you use? How do you challenge the accuracy of the experts' estimates (surveys, measurements or other ways in which you collect your data)? In case of missing data, what do you do? Do you sometimes experience time pressures? If so, how does that affect your modeling process and/or data quality? Do you have means (tools, protocols, visualizations) to report and separately discuss the quality of data on which you performed the analyses (within the risk management team and with decision makers)? How do you get the budget for additional research—data collection (in terms of time, costs and trained employees)? Do you have managerial support for collecting all the information you need?
Awareness regarding risk in organizational culture	How do you prepare for the communication of the results? What is your main concern when doing so, what do you try to highlight? What is the main massage you send? Only top risks or the overall findings? Or more? What is your biggest challenge when communicating results?

	That is when communicating to your colleagues from the team or your superiors. But what about other employees? How do others perceive your work (risk management related tasks)? How would you describe the risk management culture in your organization? Do they understand the value it creates? Do you and how do you adjust your communication and vocabulary depending on the position of the employee you are interacting with (some may have had no experience with risk management and therefore need some basic explanations)? Do you get the opportunity to attend professional conferences or other types of training? Is there an opportunity for knowledge sharing from project to project?
Impact of risk assessments in decision making	How are the results of your assessment taken into account? Generally, do you perceive there is trust in the results? Please elaborate. Is there a discussion on data included? Is the reliability of results considered? Does it comply with the development of the responses? What challenges do you experience? Do you experience any follow-ups, meaning requests for additional research/assessment/analyses? What other aspects can impact decision making?
Feedback	In your opinion, which criteria from the tailoring approach need to be specifically addressed within your organization? What are the challenges with the proposed approach? Where do you see it having the highest potential? How would it support your current practice?
Other (Some of the often recognized, additional, questions)	What additional challenges have you noticed in today's risk management? If you had the opportunity to choose, would you change the tool you are using? What happens if you are dealing with first-of-a-kind solutions, of which you do not have prior experience? Please describe how you approach modeling (assessment) in such a case. If you consider the approach too "academic," can you please elaborate why you think so? Which parts are too demanding?

Appendix 3: Tools and Techniques

*Table adapted and extended from the ISO31010 Standard.

Table A.3 Applicability of tools used for Risk Assessment

Tools and Techniques	**Risk Assessment Process**				
	Risk Identification	**Risk Analysis**			**Risk Evaluation**
		Consequence	Probability	Level of risk	
Brainstorming	SA[1]	NA[2]	NA	NA	NA
Structured or Semi-Structured Interviews	SA	NA	NA	NA	NA
Delphi	SA	NA	NA	NA	NA
Checklists	SA	NA	NA	NA	NA
Primary Hazard Analysis	SA	NA	NA	NA	NA
Hazard and Operability Studies (HAZOP)	SA	SA	A[3]	A	A
Hazard Analysis and Critical Control Points (HACCP)	SA	SA	NA	NA	SA
Environmental Risk Assessment	SA	NA	NA	NA	NA
Structure <<What if?>> (SWIFT)	SA	NA	NA	NA	NA
Scenario Analysis	SA	SA	A	A	A
Business Impact Analysis	A	SA	A	A	A
Root Cause Analysis	NA	SA	SA	SA	SA
Failure Mode Effect Analysis	SA	SA	SA	SA	SA
Fault Tree Analysis	A	NA	SA	A	A
Event Tree Analysis	A	SA	A	A	NA
Cause and Consequence Analysis	A	SA	SA	A	A

Cause-and-Effect Analysis	SA	SA	NA	NA	NA
Layer Protection Analysis (LOPA)	A	SA	A	A	NA
Decision Tree	NA	SA	SA	A	A
Human Reliability Analysis	SA	SA	SA	SA	A
Bow Tie Analysis	NA	A	SA	SA	A
Reliability Centered Maintenance	SA	SA	SA	SA	SA
Sneak Circuit Analysis	A	NA	NA	NA	NA
Markov Analysis	A	SA	NA	NA	NA
Monte Carlo Simulation	NA	NA	NA	NA	SA
Bayesian Statistics and Bayes Nets	NA	SA	NA	NA	SA
FN Curves	A	SA	SA	A	SA
Risk Indices	A	SA	SA	A	SA
Consequence/Probability Matrix	SA	SA	SA	SA	A
Cost/Benefit Analysis	A	SA	A	A	A
Multi-Criteria Decision Analysis (MCDA)	A	SA	A	SA	A
Expert judgement elicitation process with IPs	SA	SA	SA	SA	SA
The NUSAP tool	A	A	A	SA	SA
Robust Decision Making	SA	SA	SA	SA	A

[1] Strongly Applicable.
[2] Not Applicable.
[3] Applicable.

Appendix 4: Types of Risk Assessment Tools

*Table adapted and extended from the ISO31010 Standard.

Table A.4 Attributes of a Selection of Risk Assessment Tool

Type of Risk Assessment Technique	Description	Relevance of Influencing Factors			Can Provide Quantitative Output
		Resources and Capability	Nature and degree of Uncertainty	Complexity	
LOOK-UP METHODS					
Checklists	A simple form of risk identification. A technique which provides a listing of typical uncertainties which need to be considered. Users refer to a previously developed list, codes or standards.	Low	Low	Low	No
Primary Hazard Analysis	A simple inductive method of analysis whose objective is to identify the hazards and hazardous situations and events that can cause harm for a given activity, facility or system.	Low	High	Medium	No
SUPPORTING METHODS					
Structured Interview and Brainstorming	A means of collecting a broad set of ideas and evaluation, ranking them by a team. Brainstorming may be stimulated by prompts or by one-on-one and one-on-many interview techniques.	Low	Low	Low	No
Delphi Technique	A means of combining expert opinions that may support the source and influence identification, probability and consequence estimation and risk evaluation. It is a collaborative technique for building consensus among experts. Involving independent analysis and voting by experts.	Medium	Medium	Medium	No
SWIFT Structured <<What-If>>	A system for prompting a team to identify risks. Normally used within a facilitated workshop. Normally linked to a risk analysis and evaluation technique.	Medium	Medium	Any	No
Human Reliability Analysis (HRA)	Human reliability assessment (HRA) deals with the impact of humans on system performance and can be used to evaluate human error influences on the system.	Medium	Medium	Medium	Yes
SCENARIO ANALYSIS					
Root Cause Analysis (Single Loss Analysis)	A single loss that has occurred is analyzed in order to understand contributory causes and how the system or process can be improved to avoid such future losses. The analysis shall consider what controls were in place at the time the loss occurred and how controls might be improved.	Medium	Low	Medium	No

Scenario Analysis	Possible future scenarios are identified through imagination or extrapolation from the present and different risks considered assuming each of these scenarios might occur. This can be done formally or informally qualitatively or quantitatively.	Medium	High	Medium	No
Toxicological Risk Assessment	Hazards are identified and analyzed and possible pathways by which a specified target might be exposed to the hazard are identified. Information on the level of exposure and the nature of harm caused by a given level of exposure are combined to give a measure of the probability that the specified harm will occur.	High	High	Medium	Yes
Business Impact Analysis	Provides an analysis of how key disruption risks could affect an organization's operations and identifies and quantifies the capabilities that would be required to manage it.	Medium	Medium	Medium	No
Fault Tree Analysis	A technique which starts with the undesired event (top event) and determines all the ways in which it could occur. These are displayed graphically in a logical tree diagram. Once the fault tree has been developed, consideration should be given to ways of reducing or eliminating potential causes/sources.	High	High	Medium	Yes
Event Tree Analysis	Using inductive reasoning to translate probabilities of different initiating events into possible outcomes.	Medium	Medium	Medium	Yes
Cause/ Consequence Analysis	A combination of fault and event tree analysis that allows inclusion of time delays. Both causes and consequences of an initiating event are considered.	High	Medium	High	Yes
Cause-and-Effect Analysis	An effect can have a number of contributory factors which may be grouped into different categories. Contributory factors are identified often through brainstorming and displayed in a tree structure or fishbone diagram.	Low	Low	Medium	No
FUNCTION ANALYSIS					
FMEA and FMECA	FMEA (Failure Mode and Effect Analysis) is a technique which identifies failure modes and mechanisms, and their effects. There are several types of FMEA: Design (or product) FMEA which is used for components and products. System FMEA which is used for systems. Process FMEA which is used for manufacturing and assembly processes. Service FMEA and Software FMEA. FMEA may be followed by a criticality analysis which defines the significance of each failure mode, qualitatively, semi-qualitatively, or quantitatively (FMECA). The criticality analysis may be based on the probability that the failure mode will result in	Medium	Medium	Medium	Yes

	system failure, or the level of risk associated with the failure mode, or a risk priority number.				
Reliability-Centered Maintenance	A method to identify the policies that should be implemented to manage failures so as to efficiently and effectively achieve the required safety, availability and economy of operation for all types of equipment.	Medium	Medium	Medium	Yes
Sneak Analysis (Sneak Circuit Analysis)	A methodology for identifying design errors. A sneak condition is a latent hardware, software, or integrated condition that may cause an unwanted event to occur or may inhibit a desired event and is not caused by component failure. These conditions are characterized by their random nature and ability to escape detection during the most rigorous of standardized system tests. Sneak conditions can cause improper operation, loss of system availability, program delays, or even death or injury to personnel.	Medium	Medium	Medium	No
HAZOP Hazard and Operability Studies	A general process of risk identification to define possible deviations from the expected or intended performance. It uses a guideword based system. The criticalities of the deviations are assessed.	Medium	High	High	No
HACCP Hazard Analysis and Critical Control Points	A systematic, proactive, and preventive system for assuring product quality, reliability and safety of processes by measuring and monitoring specific characteristics which are required to be within defined limits.	Medium	Medium	Medium	No
CONTROLS ASSESSMENT					
LOPA (Layers of Protection Analysis)	(May also be called barrier analysis). It allows controls and their effectiveness to be evaluated.	Medium	Medium	Medium	Yes
Bow Tie Analysis	A simple diagrammatic way of describing and analyzing the pathways of a risk from hazards to outcomes and reviewing controls. It can be considered to be a combination of the logic of a fault tree analyzing the cause of an event (represented by the knot of a bow tie) and an event tree analyzing the consequences.	Medium	High	Medium	Yes
STATISTICAL METHODS					
Markov Analysis	Markov analysis, sometimes called *State-space* analysis, is commonly used in the analysis of repairable complex systems that can exist in multiple states, including various degraded states.	High	Low	High	Yes

Monte Carlo Analysis	Monte Carlo simulation is used to establish the aggregate variation in a system resulting from variations in the system, for a number of inputs, where each input has a defined distribution and the inputs are related to the output via defined relationships. The analysis can be used for a specific model where the interactions of the various inputs can be mathematically defined. The inputs can be based upon a variety of distribution types according to the nature of the uncertainty they are intended to represent. For risk assessment, triangular distributions or beta distributions are commonly used.	High	Low	High	Yes
Bayesian Analysis	A statistical procedure which utilizes prior distribution data to assess the probability of the result. Bayesian analysis depends upon the accuracy of the prior distribution to deduce an accurate result. Bayesian belief networks model cause-and-effect in a variety of domains by capturing probabilistic relationships of variable inputs to derive a result.	High	Low	High	Yes
SOWFTWARE ASSESSMENT (Examples)					
Oracle's Primavera Risk Analysis	Primavera Risk Analysis integrates directly with project schedules and cost estimates to provide quick and easy techniques to model risks and analyze the cost and schedule impacts of mitigating them. Use distribution to determine confidence levels for project pans and schedule and cost contingencies.* Available at: https://www.oracle.com/applications/primavera/products/risk-analysis.html	High	Medium	High	Yes
RamRisk	RamRisk is a web-based risk register specifically developed to support the optimal handling of risks and opportunities. RamRisk complies fully with ISO 31000, 'Risk management – Principles and guidelines'. User-friendly and flexible user interface that makes it easy to setup new projects with templates and make personal user defined views* Available at: https://ramrisk.com/	High	Medium	High	Yes
RAMAS	RAMAS Risk Calc 4.0 computes with scalars, intervals, fuzzy numbers, probability distributions, and interval bounds on probability distributions. * Available at: http://www.ramas.com/	High	High	High	Yes
Resilinc	The Resilinc is the standard for measuring, benchmarking, and tracking companies' supply chain risk and resiliency. It is a comprehensive assessment of a company's supply chain resiliency. The metric is based	High	Medium	High	No

	on extensive data collected by Resilinc working with tens of thousands of organizations as part of our global supply chain visibility initiative. It builds on key metrics: Transparency, Network Resiliency, Continuity Robustness, Performance and Supply Chain Risk Program Maturity.* Available at: https://www.resilinc.com/				
RiskCalc	The cloud-based platform provides a powerful array of tools for risk management, model development, benchmarking, impairment analysis, capital allocation and strategic business decision making.* Available at: https://rafa.moodysanalytics.com/riskcalc	High	Medium	High	Yes
NON-PROBABILISTIC METHODS					
Imprecise Probabilities	Expand the possibilities of established probabilistic risk quantification to reason more reliably with limited information on actual probability distributions. The approach allows decision makers to review and discuss coherent and plausible ranges of probabilities.	High	High	High	Yes
The NUSAP Tool	The NUSAP Tool adds qualitative information to the uncertainty and risk analysis in a structured manner, informing the modelling, analysis and decision making process by making issues such as data origin, quality and key assumptions transparent.	High	Medium	High	Yes
Robust Decision Making	The main principles of Robust Decision Making are: to explore a wide variety of relevant uncertainties, connect short-term targets to long-term goals, commit to short-term actions while keeping options open and continuously monitor the environment and take actions if necessary.	High	High	High	Yes

Appendix 5: Code in R

Code in R

```
# 10 experts (replicates of each other; could be 20)

# 6 models (c-boxes, beta distributions, point values, intervals, weighted intervals, Burgman
elicitations)

# 4 aggregation methods (average, mixture, enveloping, pooling)

#####################################################################################
#

# number of experts

some = 10          # but note that only a max of 9 are plotted

#####################################################################################
#

# MODELS

# natural frequencies

n = floor(rexp(some,1/15))    # randomly constructed

k = floor(runif(some,0,n))    # randomly constructed


# confidence boxes

CBox <- function(k,n) return(env(beta(k, n-k+1), beta(k+1, n-k)))

c = rep(CBox(0,0),some)

for (i in 1:some) c[[i]] = CBox(k[[i]],n[[i]])


# beta distributions

b = rep(beta(1,1),some)

for (i in 1:some) b[[i]] = beta(0.5+k[[i]],0.5+n[[i]]-k[[i]])


# point values
```

```
x = rep(0,some)

for (i in 1:some) x[[i]] = midpoint(mean(c[[i]]))


# intervals

ci <- function(b,level=0.95,p1=(1-level)/2,p2=1-(1-level)/2) interval(left(cut(b,p1)), right(cut(b,p2)))

v=w=NULL; for (i in 1:some) { CI = ci(c[[i]]); v = c(v,left(CI)); w = c(w,right(CI)) }


# weighted intervals

l = runif(some, 0.5, 1)

y=z=NULL; for (i in 1:some) { CI = ci(c[[i]], level=l[[i]]); y = c(y,left(CI)); z = c(z,right(CI)) }

yzl = list(y=y,z=z,l=l)


# Burgman elicitations

yzlx = list(y=y,z=z,l=l,x=x)



##########################################################################################
#

# Plot the controversy for each model


par(mfrow=c(2,3))


blank <- function() plot(NULL,xlim=c(0,1), ylim=c(0,1), xlab ='', ylab = '')

blank(); title(main = 'C-Box', xlab ='Estimates', ylab = 'Cumulative probability'); for (i in 1:some)
lines(c[[i]], col = i)

blank(); title(main = 'Beta', xlab ='Estimates', ylab = 'Cumulative probability' ); for (i in 1:some)
lines(b[[i]], col = i)

plot(density(x), xlim = c(0, 1), main = 'Estimates Density', xlab ='Estimates')
```

```
plot(NULL, xlim = c(0,1), ylim = c(1,10), main = 'Intervals', xlab ='Estimates', ylab = 'Expert ID'); for (i in 1:some){ lines(c(v[[i]],w[[i]]),rep(i,2), col = i, lwd = 2); points(x[i],i, col = i, pch = 19);}

plot(NULL, xlim = c(0,1), ylim = c(1,10), main = 'Weighted Intervals', xlab ='Estimates', ylab = 'Expert ID'); for (i in 1:some){ lines(c(y[[i]],z[[i]]),rep(i,2), col = i, lwd = 2); points(x[i],i, col = i, pch = 19);}


#####################################################################################
#

# Plot each expert with all models


par(mfrow=c(3,3))

par(mar=c(2.1,4.1,2.1,1.1))

for (i in 1:min(some,9)) {

 plot(c[[i]],col='blue', main = paste('Expert ', i, sep = ''));

 lines(b[[i]],col='red', lwd = 1);

 lines(c(v[[i]],w[[i]]),c(0.3,0.3), lwd = 2)

 lines(c(y[[i]],z[[i]]),c(0.5,0.5), lwd = 2, col = 'green')

 points(x[[i]],0.5, pch = 19, cex=1.5)

 }


par(mfrow=c(1,1))
```

```
#######################################################################################
#

# AGGREGATIONS
#######################################################################################
#
# average

Ac = c[[1]]; for (i in 2:some) Ac = Ac %|+|% c[[i]]; Ac = Ac / some   # convolutive average
Ab = b[[1]]; for (i in 2:some) Ab = Ab %|+|% b[[i]]; Ab = Ab / some   # convolutive average
Ax = mean(x)
Avw = c(mean(v), mean(w))
Ayzl = c(sum(y*l) / sum(l), sum(z*l) / sum(l))
Ayzlx = c(Ayzl, Ax)
otherAyzl = Ayzl
Y = (y - x) * (0.95 / l) + x
Z = (z - x) * (0.95 / l) + x
Y = pmax(0,Y)                    # cannot exceed zero or one
Z = pmin(1,Z)                    # cannot exceed zero or one
Ayzl = c(mean(Y), mean(Z))

Ayzlx = c(Ayzl, Ax)

# debug
# plot(Z, ylim=c(0,1))
# for (i in 1:some) lines(c(i,i), c(y[[i]],z[[i]]), col='blue')
# points(x, col='red')
# lines(l)
```

```
# points(Y)

# points(Z)



################################################################################

# mixture


Mc = mixture.pbox(c)

Mb = mixture.pbox(b)

Mx = histogram(x, mn=0, mx=1,conf=0)

I = rep(pbox(0,0),some)

for (i in 1:some) I[[i]] = pbox(interval(v[[i]],w[[i]]))

Mvw = mixture.pbox(I)

for (i in 1:some) I[[i]] = pbox(interval(y[[i]],z[[i]]))

Myzl = mixture.pbox(I,w=l)

for (i in 1:some) I[[i]] = pbox(interval(Y[[i]],Z[[i]])) # use extrapolations Y and Z

Myzlx = mixture.pbox(I)



##############################################################################

# enveloping


Ec = env(c)

Eb = env(b)

Ex = range(x)

Evw = range(c(v,w))     # i.e., c(min(v), max(w))

Eyzl = range(c(y,z))     # i.e., c(min(y), max(z))   # ignores levels l
```

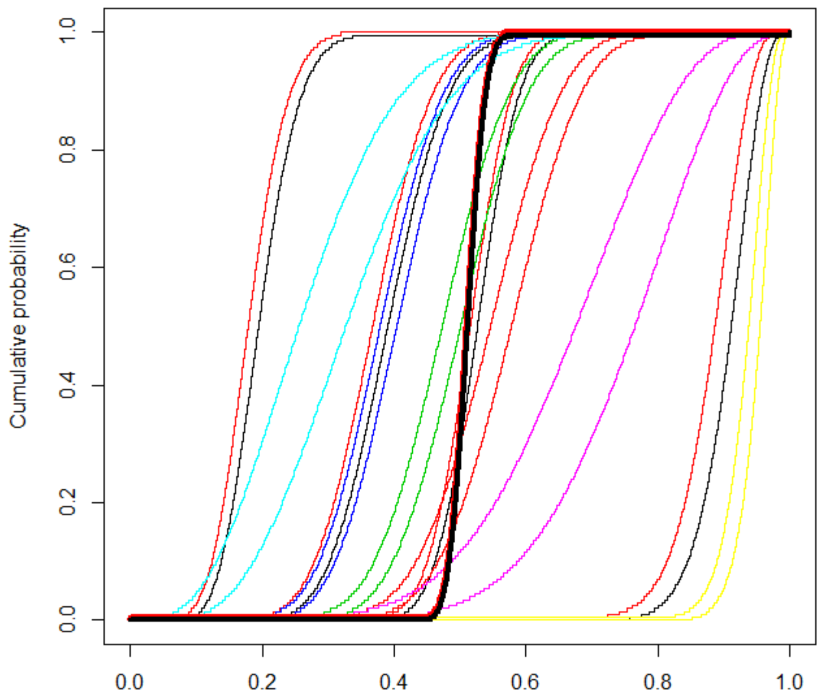
Eyzlx = list(range = Eyzl, mean = Ex)

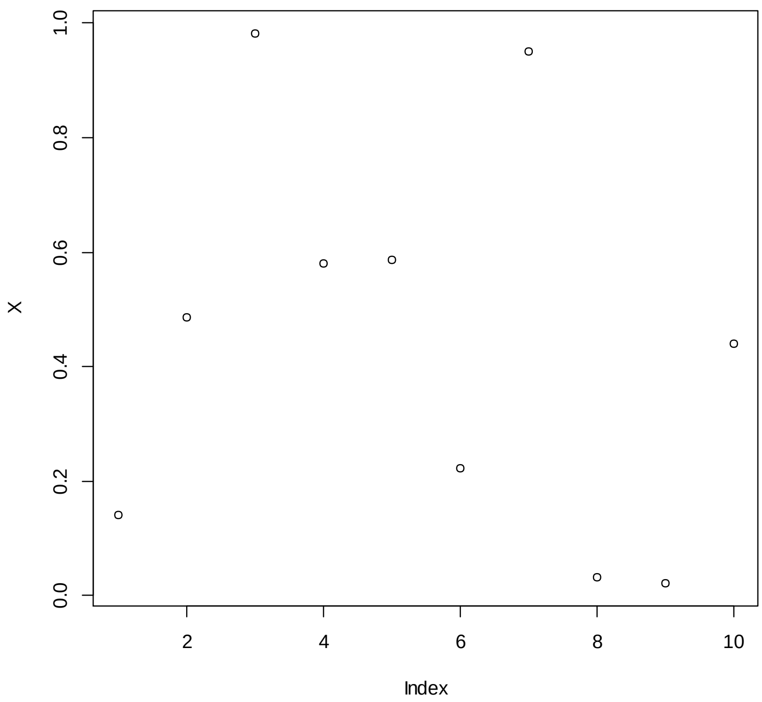
##

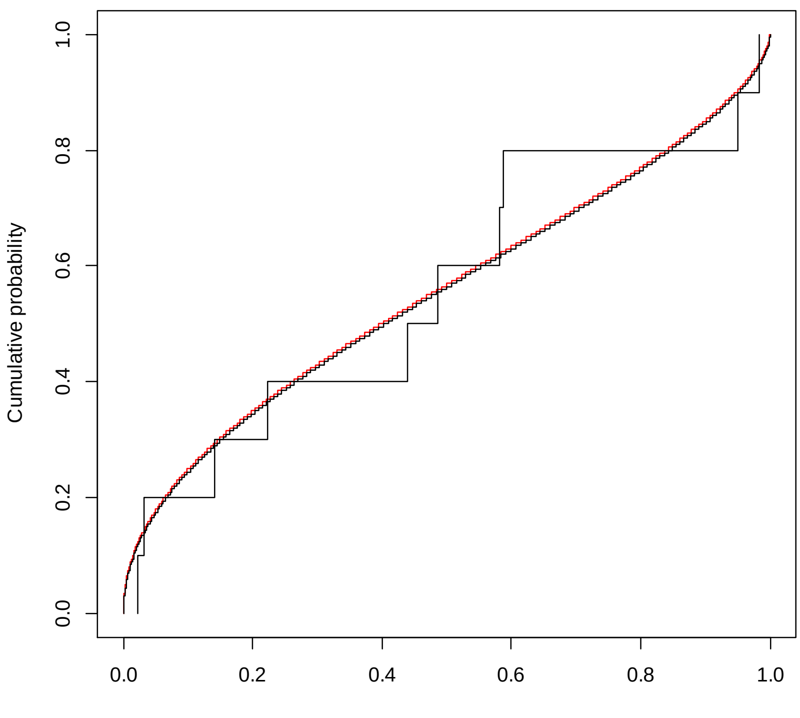
pooling

Pc = CBox(sum(k),sum(n))

Pb = beta(0.5+sum(k),0.5+sum(n)-sum(k))




X
Index
0.0
0.2
0.4
0.6
0.8
1.0
2
4
6
8
10


Cumulative probability
0.0
0.2
0.4
0.6
0.8
1.0

“Everything that can be counted does not necessarily count;
everything that counts cannot necessarily be counted.”

- Albert Einstein -