University of Stavanger

FACULTY OF SCIENCE AND TECHNOLOGY

MASTER'S THESIS

Study programme/specialisation: M.Sc. in Risk Management	Spring semester, 2018 Open
Author: Preben Emil Haugen	…………………………………………. (signature of author)
Programme coordinator: Roger Flage Supervisor(s):	
Title of master's thesis: A study of how to increase precision in the categorization of deep uncertainty, and how to assess risk under a specific level of deep uncertainty	
Credits (ECTS): 30	
Key words: Deep uncertainty Uncertainty taxonomy Risk assessment Probability bounds analysis Imprecise probabilities	Pages: 54 + enclosure: 2 Stavanger, 15/6-2018 Date/year

Intentionally left blank

A study of how to increase precision in the categorization of deep uncertainty, and how to assess risk under a specific level of deep uncertainty

Abstract

Many of the recently published articles that try to resolve challenges related to deep uncertainty have based their understanding of deep uncertainty and the various nuances of uncertainty on Courtney's uncertainty taxonomy. This thesis provides some reflections on some of the foundational pillars this taxonomy is built upon. In doing so, some major challenges and limitations are uncovered. To overcome these challenges and limitations an alternative uncertainty taxonomy is introduced. This taxonomy is built upon the same template as the one used by Courtney, but it contains a higher level of detail, and an additional level of uncertainty. The new level of uncertainty, which covers the transition from moderate to deep uncertainty is added to make sure that every nuance of uncertainty, ranging from low to deep uncertainty is reflected. Later on, a method that can assess risk under the new level of uncertainty is introduced. This method is an adaptation of a regular risk assessment process combined with a probability bounds analysis (PBA) and a qualitative judgement of the assumptions made in the analysis. The PBA form the quantitative basis of the assessment, while the qualitative judgement of the assumptions is used to justify whether the final result of the PBA can be trusted or not. By applying this method to a hypothetical case, it proves itself to be a good tool for assessing risk in cases where the empirical data is limited.

Acknowledgement

First of all, I want to give my utmost gratitude to my supervisor Professor Roger Flage for guiding me through this entire process. You have been sharing your knowledge generously and you have helped me find the way whenever I got stuck. The end result would never have been the same without your guidance and support. Thank you! Secondly, I must express my gratefulness towards my girlfriend, my family, and my closest friends for your everlasting support and constructive criticism. And last but not least, I have to thank my colleagues for taking so many of my shifts so that I could focus on this thesis.

Table of Contents

Abstract iv
Acknowledgement v
List of figures vii
List of tables vii
Abbreviations viii
1. Introduction 1
1.1. Background 1
1.2. Objectives 1
1.3. Scope and limitations of the thesis 2
1.4. Thesis structure 2
2. Theoretical foundations 3
2.1. Risk 3
2.1.1. The risk concept and risk description 3
2.1.2. Risk assessment 6
2.1.3. Risk management 7
2.2. Deep uncertainty 8
2.2.1. Definitions of deep uncertainty 8
2.2.2. Categorizations of deep uncertainty 9
2.2.3. Methods to assess and manage risk under deep uncertainty 12
3. Development of an extended deep uncertainty categorization taxonomy 20
3.1. Challenges with the current categorization taxonomies 20
3.2. An extended deep uncertainty categorization taxonomy 21
4. Adaptation of a method to assess risk characterized by level 3 uncertainty 25
4.1. Method adaptation 25
4.2. Case study 29
4.2.1. Case description 29
4.2.2. A probability bounds analysis approach 30
5. Discussion 40
5.1. The importance of acknowledging the existence of an additional uncertainty level and how the introduction of this level solves challenges that are present in the other uncertainty taxonomies 40
5.2. The newly proposed methods applicability as a tool for assessing risk 41
5.3. Further work 42
6. Conclusion 43
References 44
Appendix A 47

A.1. The calculations behind the bounds on the cumulative probabilities of the fraction of survivors 47

A.2. The calculations behind the best and worst case estimate of the fraction of survivors in the long run 48

List of figures

Figure 1: A risk management process (ISO, 2018)........ 6
Figure 2: Illustration of the p-box from the previous example (Tucker & Ferson, 2003)........ 17
Figure 3: P-box with the bounds of estimated required cut in greenhouse gas emissions to prevent an increase in temperature of more than 2 degrees........ 18
Figure 4: P-box with the bounds of estimated temperature increases with a 35% cut in greenhouse gas emissions........ 18
Figure 5: Illustration of the interval [1, 3] which is guaranteed to contain the sum of σ_1 = [0, 1] and σ_2 = [1, 2]........ 26
Figure 6: Probability distribution of the sum of the random variables σ_1 = uniform(0, 1) and σ_2 = uniform(1, 2) under the assumption that the parameters are independent........ 26
Figure 7: An overview of the situation in the form of an event tree........ 30
Figure 8: Overview of the probability of occurrence of the different outcomes........ 31
Figure 9: Probability distribution of survivors at different wave heights........ 32
Figure 10: Best & worst case cumulative probabilities for the different wave heights........ 33
Figure 11: Bounds on the cumulative probabilities of the fraction of survivors........ 33
Figure 12: Bounds on the cumulative probabilities of the fraction of survivors........ 37
Figure 13: Bounds on the cumulative probabilities of the fraction of survivors........ 39

List of tables

Table 1: The uncertainty taxonomy introduced by Courtney (Walker et al., 2010)........ 10
Table 2: A simplification of the uncertainty classification taxonomy (Aven, 2013)........ 12
Table 3: An alternative uncertainty classification taxonomy (Aven, 2013)........ 12
Table 4: A modified uncertainty classification taxonomy. Compared to Courtney's taxonomy, newly added features are written in italic. The gradual shift in color from green to red represents the increase in uncertainty........ 24
Table 5: Overview of probability of occurrence........ 29
Table 6: Fraction of resident who survive at the different wave height intervals........ 30
Table 7: Probability bounds on the worst and best case scenarios........ 32
Table 8: An overview of the relevant assumptions, their associated justifications and judgement of the strength of these justifications. Part 1 of 2........ 34
Table 9: An overview of the relevant assumptions, their associated justifications and judgement of the strength of these justifications. Part 2 of 2........ 35
Table 10: Numerical overview of the bounds on survival probabilities........ 38

Abbreviations

ARM	Adaptive risk management
CDF	Cumulative distribution function
EU	Expected utility
ICAF	Implied cost of averting a fatality
PBA	Probability bounds analysis
QRA	Quantitative risk assessment
RAC	Risk Acceptance criteria
RDM	Robust decision making
SoK	Strength of Knowledge
SRA	Society for Risk Analysis

1. Introduction

This chapter has three main purposes. The first one is to introduce the reader to the main objectives of the thesis and to makes it clear why it is important to investigate and solve these problems. The second one is to make it clear what the scope and limitation of the thesis are. The final purpose of this chapter is to inform the reader about how the thesis is organized.

1.1. Background

Global warming, the future of the world's economy and the current political instability are some of the greatest challenges the collective community of the world is currently facing, and they are all characterized by deep uncertainty. This is cause for concern because handling a situation that is characterized by deep uncertainty is a main foundational issue risk assessment and risk management (Aven, 2013). To handle a situation like this it is of outmost importance to understand what deep uncertainty really is, and to acknowledge its presence. Numerous definitions of deep uncertainty do exist (Bjerga & Aven, 2015, p. 75; Cox, 2012, p. 1607; Walker et al., 2017, p. 5). They seem to agree that for a situation to be characterized by deep uncertainty the available empirical information must be so limited and the underlying phenomena so poorly understood that it is hard or even impossible to identify possible outcomes and their probability of occurrence. In 2001 Courtney (2001) introduced an uncertainty taxonomy which intended to clarify the different levels of uncertainty, as they progress from low to deep uncertainty. This taxonomy has later formed the foundation for the understanding of deep uncertainty in several articles that try to resolve challenges related to deep uncertainty (Cox, 2012; Walker et al., 2010; Walker et a., 2017). Using this taxonomy as a basis for understanding the nuances of the various levels of uncertainty may not be the best idea, as Aven (2013) pointed out. He argues that "critical questions can be raised regarding its foundations" and later presents an alternative taxonomy (Aven 2013, p.2082). As we shall see later on, this alternative taxonomy is not without limitations. The challenges and limitations found in these taxonomies are the point of departure for this thesis.

1.2. Objectives

The objectives of this thesis have been to:

- Further develop the deep uncertainty taxonomies introduced by Courtney (2001) and Aven (2013).
- Adapt a method to assess risk under a newly introduced level of uncertainty by combining probability bounds analysis with a qualitative judgement of assumptions.

1.3. Scope and limitations of the thesis

The thesis will focus on highlighting and discussing challenges and limitations that are rooted in the uncertainty taxonomies introduced by Courtney (2001) and Aven (2013). The result of this work is then used to develop an alternative uncertainty taxonomy with an additional level of uncertainty, and a higher level of detail. Finally, a risk assessment method that can assess risk under the newly introduced level of uncertainty is adapted to work under the given circumstances and its abilities are illustrated trough a stylized example.

1.4. Thesis structure

This thesis consists of six chapters which cover the development of a new uncertainty taxonomy and the adaptation of method to assess risk under a newly introduced level of uncertainty. The setup of this thesis may differ slightly from the norm, since the discussion is not solely presented in the second to last chapter. This was an active choice taken by the author to give the text a better flow, and to prevent the reader from having to skip back and forth. A short summary of the content of the different chapters can be seen below:

Chapter One: this chapter introduces the reader to the main objectives of the thesis and to make it clear why it is important to investigate and solve these problems. It also presents the scope and limitations of the thesis, as well as an overview of the structure.

Chapter Two: this chapter is a literature review that introduces the reader to relevant theoretical foundations of risk and deep uncertainty that lay the foundation of this study.

Chapter Three: this chapter presents and discusses the challenges and limitations that are imbedded in the existing uncertainty taxonomies. This is done to express the need for a new and alternative taxonomy, which is developed later on in the same chapter.

Chapter Four: this chapter introduces the reader to a method that can be used to assess risk that fall under the previously introduced third level of uncertainty and discusses the background of this method. To illustrate how this method can be used it is applied to a hypothetical case.

Chapter Five: This chapter will introduce two separate discussion topics as well as suggestions for further work. The aim of the first discussion is to discuss why it is important to acknowledge an additional level of uncertainty and to what degree the new alternative uncertainty taxonomy eliminates the challenges that are present in the other uncertainty taxonomies. The aim of the second discussion is to discuss the previously introduced risk assessment method's applicability as a tool for assessing risk.

Chapter Six: This chapter will present the conclusions that can be drawn from the study that has been done here in this thesis.

2. Theoretical foundations

The purpose of this section is to introduce the theoretical foundations that are of importance to the research undertaken in thesis. Rather than being an exhaustive literature review, this section will give an overview of the modern collective thinking of some of the world's leading experts in the areas of risk, risk assessment, risk management and deep uncertainty that are relevant for this thesis. It will also define associated concepts such as probability, uncertainty and knowledge.

2.1. Risk

To give an overview of the modern collective thinking of some of the world's leading experts in the field of risk, the author of this thesis has been looking to the glossary on risk and risk related terms recently published by the Society for Risk Analysis (SRA), as it was developed by a committee of 11 active risk experts from various academic fields (SRA, 2015). Another reason for looking to this glossary is the important premises it is based upon. First, it allows for, and includes different perspectives, meaning that it includes several definitions of risk. Secondly, a clear distinction is made between risk as a concept and the measurement/description of risk. Finally, the included definitions must meet some basic criteria such as being logical, well-defined, understandable, precise, etc.

2.1.1. The risk concept and risk description

To explore the risk concept as described in the SRA glossary it is essential to first define a risk setting. Here we consider some future activity, it could be anything from driving a car to buying a house and define risk in relation to the consequences of this activity with respect to something that is of value to us humans (e.g. economy, health, etc.) (SRA, 2015). There is always at least one consequence that is considered negative or undesired, meaning that risk should not be solely be associated with negative outcomes. With this setting in mind, here are the overall qualitative definitions of risk as given in the SRA glossary (SRA, 2015, p. 3):

a) *Risk is the possibility of an unfortunate occurrence.*
b) *Risk is the potential for realization of unwanted, negative consequences of an event.*
c) *Risk is exposure to a proposition (e.g. the occurrence of a loss) of which one is uncertain.*
d) *Risk is the consequences of the activity and associated uncertainties.*
e) *Risk is uncertainty about and severity of the consequences of an activity with respect to something that humans value.*
f) *Risk is the occurrences of some specified consequences of the activity and associated uncertainties.*
g) *Risk is the deviation from a reference value and associated uncertainties.*

From these definitions we can see that SRA consider risk to be defined through uncertainty. This is in line with the (A, C, U)-perspective of risk presented by Aven (2015), where he says that an event, A, will have some consequences, C, and there is uncertainty, U, about what these consequences will be (Aven, 2015).

As it is understood here, risk is closely related to the concept of uncertainty, which in a risk setting can be interpreted in two ways, either as *aleatory* or *epistemic* uncertainty (Walker, Lempert & Kwakkel, 2017; Walker, Marchau and Swanson, 2010; Aven, 2016; Beer et al., 2013). Aleatory uncertainty is a type of uncertainty which is seen as irreducible, since it *cannot* be reduce further by acquiring more knowledge. It represents the property of a system which is associated with variability or fluctuations, and an example of such a system is the rolling of a die. Epistemic uncertainty is a type of uncertainty that *can* be reduced by acquiring more knowledge, as it results from the analyst not having complete information about the system, in other words a lack of knowledge. This type of uncertainty may also be denoted as ignorance (Ferson & Ginzburg, 1996). An example of such a system could be the long term effects of climate change. It can however be argued that all uncertainties can be seen as epistemic on a fundamental level (Winkler, 1996). E.g. by acquiring knowledge on all the physical variables related to the rolling of a die (spin, weight, height, speed, shape, hardness, etc.) we can reduce the uncertainty related to the outcome.

Since knowledge plays an important role in defining uncertainty it is important to define the meaning of knowledge in this context. The 2015 SRA glossary distinguish between two distinct types of knowledge (SRA, 2015, p. 8): "know-how (skill) and know-that of propositional knowledge (justified beliefs). Knowledge is gained through for example scientific methodology and peer-review, experience and testing." The know-how part of knowledge is a skill that is acquired over time, e.g. driving a car. The know-that part of knowledge or justified beliefs is knowledge gained or strengthened over time by seeing similar results from similar events, e.g. the force of gravity pulls everything towards the ground. Another important aspect of knowledge in a risk context is assumptions. An assumption is "something that you consider likely to be true even though no one has told you directly or even though you have no proof" according to the Macmillan Dictionary (Macmillan Dictionary, 2018). In other words, an assumption is one of your personal beliefs. Even though you assume something to be true it may turn out not to be true, and in a risk context this can have fatal consequences. Flage and Aven have proposed a set of principles as guidelines for assessing the strength your background knowledge in any situation (Flage & Aven, 2009):

The background knowledge is considered as strong if <u>all</u> of the following conditions are met:

- The assumptions made are seen as very reasonable.
- Large amount of reliable and relevant data/information is available.
- There is broad agreement among experts.
- The phenomena involved are well understood; the models used are known to give good predictions.

The background knowledge is considered as poor if one or more of the following conditions are met:

- The assumptions made represent strong simplifications.
- Data/information is non-existent or highly unreliable /irrelevant.
- There is strong disagreement among experts.
- The phenomena involved are poorly understood, models are non-existent or known/believed to give poor predictions.

For cases in between the background knowledge is considered moderate.

If your knowledge or assumptions turn out not to be true, surprises can occur. In a risk context the most extreme surprises are known as black swans (Taleb, 2010; Aven, 2014). Aven has defined three distinct types of black swans (Aven, 2014):

1. Unknown unknown – unknown to everyone
2. Unknown known – known to some, but not to the assessor
3. Events judged not to occur because the probability of occurrence is seen as negligible

To manage risk, we have to be able to measure risk, and this is where risk description comes in. Risk description; abbreviated as (C', Q, K) where C' is the specified consequences, Q is the measurement of uncertainty, often measured by using probability, P, and K is the knowledge on which C' and P is based upon (Aven et. al, 2013). Risk can also be described as (A', C', Q, K), where A' is the specified initiating event (Aven, 2015). The SRA glossary included several other definitions of risk description (SRA, 2015, p. 4):

1. *The combination of probability and magnitude/severity of consequences.*
2. *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.*
3. *Expected consequences (damage, loss), for example computed by:*
   I. *Expected number of fatalities in a specific period of time or the expected number of fatalities per unit of exposure time.*
   II. *The product of the probability of the hazard occurring and the probability that the relevant object is exposed given the hazard, and the expected damage given that the hazard occurs and the object is exposed to it (the last term is a vulnerability metric).*
   III. *Expected disutility.*
4. *A possibility distribution for the damage (for example a triangular possibility distribution).*

Due to its importance in risk description it is essential to define the meaning of probability in this context. According to the SRA glossary, overall, probability is a measure of uncertainty, belief or variation which follows the rules of probability calculus, but various interpretations exist (SRA, 2015). There are however only two probability interpretations that are being frequently used in a risk context, frequentist probabilities and subjective probabilities (Aven, 2013 & Aven & Reniers, 2013). Both are described below. Less frequently used probability interpretations also exist, see (Aven & Reniers, 2013).

(i) *Frequentist probability* (P_f), this is a purely objective probability where P_f represents the relative fraction of times an event occurs if the situation in question were hypothetically "repeated" an infinite number of times. The variation in the outcomes of this repetition which bring about the "true" value of P_f is usually referred to as aleatory or stochastic uncertainty. The "true" underlying frequentist probability P_f can never be known with 100% certainty and has to be estimated.

(ii) *Subjective probability* (P, SoK), often referred to as knowledge based or judgemental probability express the assessor's degree of belief about an occurring event, based on his or her background knowledge. It can be interpreted with reference to a standard event like drawing a specific ball from an urn containing a specific number of balls. If the assignor assigns a probability of 0,4 to an event A, he or she compares his or her uncertainty (degree of belief) of event A occurring to drawing a blue ball from an urn containing ten balls, where four of them are blue. Subjective probability can also be interpreted with reference to betting.

2.1.2. Risk assessment

This subchapter is based on the ISO 31000 standard (ISO, 2018).

The ISO 31000 standard defines a risk assessment as the identification, analysis and evaluation of risk, see figure 1. The identification, being the initial part of risk assessment, aims to find, recognize and describe risk. Through this process one identifies the elements which has the potential to give rise to risk, the initiating events and their potential consequences. This is accomplished by using historical data, theoretical analysis and/or expert opinions. The second step of the risk assessment aims to analyze the risks that are identified, which enables the risk analyst to present an informative risk picture. The final step of the risk assessment is the risk evaluation. Here the risk analyst will compare the results of the concluded risk analysis against given risk criteria. The objective of this step is to determine the significance of risk, and whether the risk is acceptable or not.

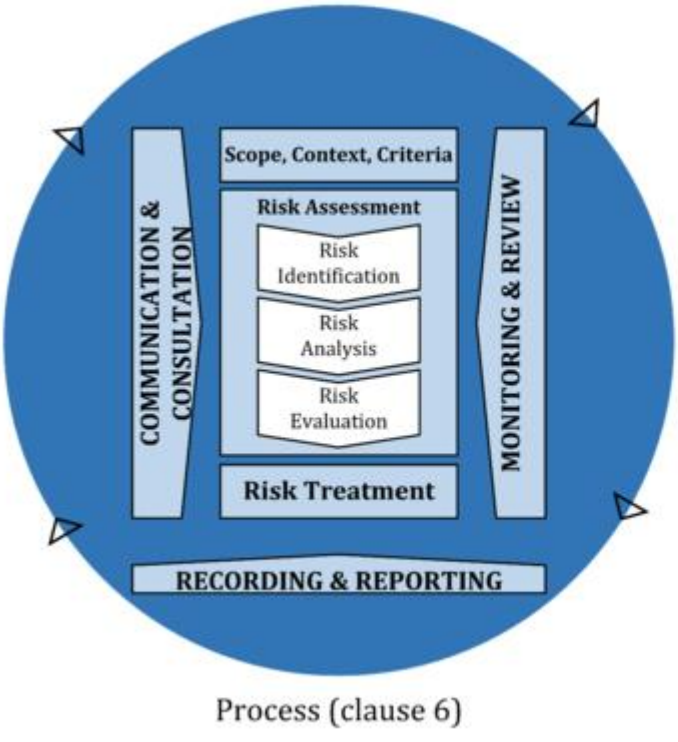


Figure 1: A risk management process (ISO, 2018).

2.1.3. Risk management

This subchapter is based on the ISO 31000 standard (ISO, 2018), unless stated otherwise. The ISO 31000 standard describes risk management as the coordinated activities to direct and control an organization with regards to risk. It has also been defined as all the measures and activities that are carried out to manage risk (Aven, 2015). Risk management plays a vital role in balancing the pursuit of possibly gainful opportunities, and avoiding the losses, accidents and disasters that may follow (Aven, 2015). Figure 1 presents an overview of the risk management process, which includes establishing the context, risk assessment, risk treatment, risk communication and monitoring and review. Establishment of context is the initial step of this iterative process. This step includes describing the internal and external parameters that are to be taken into consideration when managing risk. Furthermore, the process includes setting the scope and defining the criteria (a reference which the significance of risk is evaluated against) for the risk management policy (a statement of the intentions and direction of an organization that are related to risk management). The second step is risk assessment, see chapter 2.1.2. for a detailed description. The third step consists of risk treatment, in which the objective is to mitigate the risk. This could mean refraining from taking certain actions to avoid risk, exploring an opportunity in despite of increased risk, eliminating the source of the risk entirely or altering the consequences or likelihood (probability of something happening, see chapter 2.2. for more info on probability) to one's advantage. The last part of this process includes monitoring and review, which is a phase where the objective is to monitor and review the effects of the risk treatment. Monitoring and review is also deeply imbedded in every step of the process.

2.2. Deep uncertainty

2.2.1. Definitions of deep uncertainty

As stated above, one way of defining uncertainty is as limits or gaps in knowledge about the future, past or current events (Walker, Lempert & Kwakkel, 2017; Walker, Marchau and Swanson, 2010; Aven, 2016). A notion worthy of remark is that uncertainty will not in all cases be reduced by the acquisition of new information; it may also increase. (Walker et al., 2017). This is demonstrated when additional information reveals understated or previously unknown uncertainties on an intricate system.

Knight made the distinction between risk and uncertainty in 1921, and consequently introduced one of the modern understandings of uncertainty as the lack of knowledge. (Walker et al., 2017). Knight argued that risk could be considered as the calculable and thus the controllable aspect of the unknown. Uncertainty constitutes the remaining share, this part being incomputable and uncontrollable. A resembling distinction was made in 1989 between stochastic and real uncertainty (Quade, 1989). Quade (1989) observed that stochastic uncertainty could be described by frequentist and subjective probability models. Real uncertainty would on the other hand describe future occurrences, making precise, long-term predictions practically impossible, examples being the financial markets and climate changes. Real uncertainty is now also commonly referred to as deep uncertainty (Walker et al., 2017).

Deep uncertainties are in Cox's analysis defined in the following manner (Cox, 2012, p. 1607):

Well-validated, trustworthy risk models giving the probabilities of future consequences of alternative decisions are not available; the relevance of past data for predicting future outcomes is in doubt, experts disagree about the probable consequences of alternative policies-or, worse, reach an unwarranted consensus that replaces acknowledgement of uncertainties and information gaps with groupthink-and policymakers are divided about what actions to take to reduce risks and increase benefits.

This bears significant resemblance to the definition of deep uncertainty given by Lempert et al. (Walker et al., 2017, p. 5):

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.

Experts do seemingly concur that the concept of deep uncertainty is characterized by substantial model uncertainty, even up to a point where no model is available, thus making it difficult or even impossible to give any predictions as to what the outcome of a given situation might be.

2.2.2. Categorizations of deep uncertainty

For the purpose of simplifying the identification and categorization of deep uncertainty, Courtney (2001) presented a taxonomy that distinguish between two highly extensive levels of uncertainty; determinism and total ignorance (Courtney, 2001; Courtney, 2003; Walker et al., 2010; Cox, 2012; Aven, 2013). Ranging between them are four separate levels of uncertainty "defined with respect to the knowledge assumed about the various aspects of a policy problem" (Walker et al., 2010, pp. 918). These aspects are: the future of the world (context), the model of the relevant system for said future world (system model), the outcomes from the system (system outcomes) and the emphasis which the respective stakeholders will put on the outcomes (weights on outcomes) (Walker et al., 2010). The taxonomy in its entirety can be seen in table 1, and the levels of uncertainty are explained in detail below (Walker et al., 2010):

Determinism: the ideal situation. Here every aspect of a given situation can be described with 100% accuracy. In this taxonomy it is used as a limiting characteristic, since this level of accuracy is impossible to obtain.

Level 1 uncertainty: a clear enough future. The characteristic feature of this level of uncertainty is the possibility of making accurate models and description of the related uncertainty. This is owed to the fact that the range of possible outcomes is narrow and the existence of a large amount of relevant data. Although this should not be confused with a perception of a perfectly predictable future, this level of uncertainty provides decent opportunity to assess the associated risk(s). Perfect measurements and models are impossible to make and are the stated reason for the overall uncertainty in this level.

Level 2 uncertainty: alternate futures. The characteristic features of this level of uncertainty is that a limited number of outcomes (but larger than in level 1) can be determined, where one of them will occur. It is also possible to adequately describe and model the probability of occurrence of each outcome, meaning that they can be ranked and the risk(s) involved can be assessed.

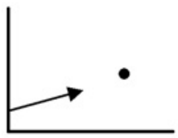
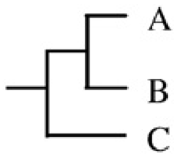
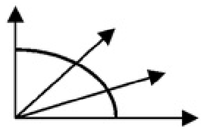
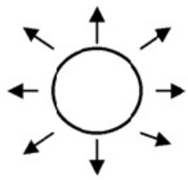
Level 3 uncertainty: a range of futures. This is the first level in the deep uncertainty category. The characteristic features of this level of uncertainty is that a limited number of outcomes can be determined, but the final outcome may not be among the ones that are identified. It is not possible to say anything regarding the probability of occurrence for the identified outcomes.

Level 4 uncertainty: an unknown future. This is the second level of deep uncertainty and it provides no opportunity to determine any future outcomes. This level is characterized by knowledge of one’s utter unawareness. Since the early 2000’s, a wide range of global scale level 4 uncertainty situations have occurred, e.g. the attack on the U.S. twin towers in 2001, the Indian Ocean tsunami in 2004 and the U.S. subprime mortgage crisis in 2007. These situations can be referred to as black swans: surprising extreme events relative to one’s own knowledge (Aven, 2014). Due to the extensive impact these events have had on society worldwide, ignorance as a contributing factor to increased risk have lately received an increased focus.

Total ignorance: One remains unaware of one’s own oblivion. This is a state of utmost uncertainty, leaving no possibility of insight into a given situation.

A similar distinction has also been made by Makridakis et al. (2009), where they called the first two levels of uncertainty for subway uncertainties and the last two levels of uncertainty for coconut uncertainties. Subway uncertainties refers to situations that can be modeled and where one can for all practical purposes say that the probability models that are introduced can be seen as an accurate representation of the "true" underlying probability (Makridakis et al., 2009). Coconut uncertainties refer to situations where it is impossible to introduce a probability model (Makridakis et al., 2009). The latter can also refer to rare and unique events that are difficult to imagine.

Table 1: The uncertainty taxonomy introduced by Courtney (Walker et al., 2010).

		Level 1	**Level 2**	**Level 3**	**Level 4**	
				Deep Uncertainty		
	Context	A clear enough future 	Alternate futures (with probabilities) 	A multiplicity of plausible futures 	Unknown future 	
Determinism	**System model**	A single system model	A single system model with a probabilistic parameterization	Several system models, with different structures	Unknown system model; know we don’t know	**Total ignorance**
	System outcomes	A point estimate and confidence interval for each outcome	Several sets of point estimates and confidence intervals for the outcomes, with a probability attached to each set	A known range of outcomes	Unknown outcomes; know we don’t know	
	Weights on outcomes	A single estimate of the weights	Several sets of weights, with a probability attached to each set	A known range of weights	Unknown weights; know we don’t know	

In 2013, Aven (2013) provided his reflections on some of the foundational pillars that previous work on deep uncertainty is based upon. This includes among others his reflections on the meaning of the deep uncertainty concept. Here he argues that the uncertainty taxonomy introduced by Courtney (2001), which also forms the basis for Cox’s presentation and analysis of deep uncertainty (Cox, 2012), can be challenged.

From the taxonomy above it is clear that if a system is not to be characterized by deep uncertainty it must fall within either level 1 or level 2. This means that the underlying “true” probability can be estimated or modeled (Aven, 2013). Probability models requires the introduction of frequentist type probabilities, which is to say that it must be possible to define or imagine a very large (in theory an infinite) population of similar situations to the one being considered. If x is a random variable that follows a probability model F, say a binomial distribution with parameters n and p, $F(x \mid n, p)$, it means that there must exist a frequentist type of probability that represents the “true” underlying P_f. The probability model F is a representation or a best estimate of the “true” underlying P_f, and it can never be 100% accurate (Aven, 2013). This framework presumes the following conditions to hold (Aven, 2013, p. 2084) (see the next page):

(i) *The existence of frequentist probabilities P_f.*

(ii) *The probability model introduced is an accurate representation of the "true" underlying frequentist distribution P_f.*

Aven argues that Courtney's taxonomy can be both simplified and made more precise by defining two categories or levels of uncertainty, Level A and Level B (Aven, 2013), see table 2. In Level A the relevant experts have full confidence in conditions (i) and (ii), meaning that the structure or model for the underlying P_f is considered known. This however does not mean that the "true" underlying P_f is known. Referring to a model as "correct" or "true" must be understood as a model that is an accurate representation of the real world. As previously mentioned, all models are just simplifications of the real world, meaning that they are incorrect or wrong if we are to use a precise language.

Level B includes every situation where conditions (i) and (ii) are not both met, in other words these situations are characterized by deep uncertainties (Aven, 2013). This category where we cannot justify frequentist probabilities P_f and/or accurate probability models, is large, as there are only a few situations where conditions (i) and (ii) are actually met. One could say that level A only refers to situations of controlled experiments where the experiments can be repeated a large number of times under similar conditions to confirm the hypotheses and the constructed probability model (Aven, 2013). In any other situation it will always be hard to know whether the data that has been collected is relevant and/or if the models in use can make accurate predictions about the future. Consider for instance a car manufacturing plant where they are concerned with the number of incidents N of a specific type A. To analyze and predict N we can collect data over time and make a Poisson probability model with the parameter λ that express the average number of times A occurs per unit of time in the long run. Despite the fact that this model is based on accurate historical data, is continuously updated as soon as new data has been collected, has been reasonably accurate until now, and we have good arguments for justifying its predictions, there is no way we can prove that future events will follow its predictions. In brief, we may be very confident that conditions (i) and (ii) are true, but this is impossible to prove, and even though we are confident that our model gives a good estimate of P_f it may in fact turn out to be wrong. This aspect of uncertainty is not reflected in Courtney's taxonomy, meaning that the knowledge dimension is not adequately taken into account (Aven, 2013).

Since Courtney's taxonomy does not reflect the strength of the background knowledge, which in turn leads to a lack in expression of confidence in condition (i) and/or (ii), Aven suggested to add the following notations to Level A in his categorization of uncertainty (Aven, 2013, p. 2085):

(1) Level A': Confidence in conditions (i) and (ii) (they have been justified), and the conditions hold (with reasonable accuracy).

(2) Level A'': Confidence in conditions (i) and (ii), but the conditions do not in fact hold.

Table 2: A simplification of the uncertainty classification taxonomy (Aven, 2013).

Level A • **Not deep uncertainties**	**Level B** • **Deep uncertainties**
Conditions (i) and (ii) are justified	Conditions (i) and (ii) are not justified
A' or A''	

When the decisionmaker makes his/her decision it will be impossible for him/her or the risk analysts to know for sure whether A' or A'' is true. For this reason, the decisionmaker should reflect on the possibility and likelihood that even though the risk analysts have strong confidence in both condition (i) and (ii) they may not hold. To account for this and to make an even more nuanced categorization of Level A, Aven (2013, pp. 2085) suggests to distinguish between:

(a) *Situations where we have strong evidence (we can for practical purposes conclude that A' applies);*

(b) *Other situations-characterized by some dominating explanations and beliefs.*

By following these ideas Aven made an alternative classification system, see table 3. This new system is based on the strength of knowledge and it includes the occurrence of black swans. In the first category where the uncertainties are low, the knowledge is strong, and the occurrence of black swans can for all practical purposes be ignored. In the second category where the uncertainties are moderate, the knowledge is based on some dominating explanations and beliefs, and one must prepare for the possible occurrence of black swans. In the final category which represents systems characterized by deep uncertainty the knowledgebase is poor. This makes it meaningless to refer to black swans, as there is no knowing what can happen. However, the occurrence of new types of events known as unknown unknowns may occur in this category as well as in the second (Aven, 2013).

Table 3: An alternative uncertainty classification taxonomy (Aven, 2013).

Low uncertainties	**Moderate uncertainties**	**Deep uncertainties**
Strong knowledge	Some dominating explanations and beliefs	Poor knowledge
No black swans	A black swan may occur	No black swans

2.2.3. Methods to assess and manage risk under deep uncertainty

In some cases, there is too little knowledge available to support a probabilistic representation of the uncertainties. In these cases, regular probabilistic risk assessments may not be satisfactory (Shortridge et al., 2017). Consequently, it might prove challenging to manage risk under deep uncertainty. Research has shown that in order to avoid presenting probabilities that can easily be considered untrustworthy, leading risk analysts are often hesitant to present subjective probabilities in situations where their background knowledge is limited (Chao et al., 1999). Assigning probabilities based on limited background knowledge may lead decisionmakers to believe there is a higher certainty than what is actually the case, thus giving an inaccurate conception of the true range of possible outcomes (Clark & Pulwarty, 2003). In an attempt to avoid this situation, some organizations restrict their application of probabilistic analysis under deep uncertainty to a limited number of cases (Shortridge et al., 2017). An example of this is the IPCCs' (Intergovernmental Panel on Climate Change) guidance on reporting climate impacts which requires high confidence (reliable evidence which are in general agreement with each other) before authors can use probabilities to characterize uncertainties (Mastrandrea et al., 2010). Consequently, a series of different methodologies have been proposed to provide a broader treatment of non-probabilistic uncertainty, which include “frequency of probability” approaches (Kaplan & Garrick, 1981), different numerical alternatives to probabilities like *imprecise probabilities* (Walley, 1991), *probability bounds analysis* (PBA) (Williamson & Downs, 1990 & Ferson & Ginzburg, 1996) and *possibility theory* (Dubois et al., 1988), qualitative measures for describing the weight of evidence on which probability assessments are made (Aven, 2008), robustness-based decision support frameworks that do not rely on probabilities like *Robust Decision Making* (RDM) (Lempert et al. 2006), *Info-Gap Theory* (Ben-Haim, 2000), and *Resilience Analytics* (Karvetski & Lambert, 2012 & Hamilton et al., 2016), and adaptive frameworks such as *Adaptive Risk Management* (ARM) (Holling ,1978) and *Dynamic Adaptive Policy Pathways* (Haasnoot et al., 2013). Adaptive Risk Management and Robust Decision Making will later be described in detail as these methods are often recommended and more commonly used then the others to assess and manage risk in situations characterized deep uncertainty (Lempert et al., 2004; Kasperson, 2008; Cox, 2012; Aven 2016; Walker et al, 2010; Kwakkel et al., 2016; Maier et al., 2016). Probability bounds analysis will also be described in detail as this method will be suggested as possible way to assess risk under a fifth level of uncertainty which will be introduced in chapter 3. The description of the rest of the previously mentioned methods for assessing and managing risk under deep uncertainty is beyond the scope of this thesis.

According to Shortridge et al. (2017) most of the currently existing research on the previously mentioned methods focus on their development, debate on their practicality and theoretical foundations, and applications to specific problems. This would have been a good thing had these methods not been developed in relative isolation from each other, which makes the relative advantages, limitations, assumptions and practical implications of each method compared to the others hard to grasp (Shortridge et al., 2017). Another downside of this is that it limits the degree to which scientists and other users can build upon the previous research in this field and apply these methods to solve problems where regular probabilistic analysis is thought to be insufficient or inappropriate (Shortridge et al., 2017). A set of systematic comparisons between the different methods could help in resolving some of these issues. Until recently, only a few such comparisons existed, and the ones that did tended to focus on numerical alternatives to probability without considering the semi-quantitative, robustness-based or adaptive methods mentioned above (Dubois & Prade, 1992 & Soundappan et al. 2004). This lack of direct comparison has fortunately been recognized, and in the last couple of years a few papers have been published on this subject. Shortridge et al. (2017) compare semi-quantitative uncertainty factors, probability bounds analysis and Robust Decision Making as methods to assess risk under deep uncertainty through the use of a stylized climate change adaptation problem related to flood risks in a riverfront city. Hall et al. (2012) compare Robust Decision Making and Info-Gap theory as methods to assess and manage risk and use the evaluation of greenhouse-gas emissions policies as an example. Kwakkel et al. (2016) compare Robust Decision Making and Dynamic Adaptive Policy Pathways as methods to assess and manage risk under deep uncertainty and illustrate it by using a flooding case in the Rhine Delta of the Netherlands. Systematic comparisons such as these help in clarifying and highlighting the fundamental differences between the different methods. They also open up for the use of more sophisticated examples and different forms of uncertainty as valuable tools to further distinguish the methods.

Cox (2012) preformed a thorough study on the topic of confronting deep uncertainties in risk analysis, reviewing ten tools to "help us to better understand deep uncertainty and make decisions even when correct models are unknown" (Cox, 2012, p. 1611). The tools he reviewed were: (subjective) expected utility theory; multiple priors, models or scenarios, robust control, robust decisions; robust optimization; average models; resampling; adaptive boosting; Bayesian model averaging; low regret online detection; reinforced learning; and model-free reinforced learning. Cox states that "they provide genuine breakthroughs for improving predictions and decisions when the correct model is highly uncertain" (Cox, 2012, p. 1607). The ten reviewed tools implement either one of two strategies (Cox, 2012, p. 1611):

finding robust decisions that work acceptably well for many models (those in the uncertainty set); and adaptive risk management, or learning what to do by well-designed an analyzed trial and error.

After reviewing Cox's study, Aven concluded that "deep uncertainties call for a managerial review and judgement that sees beyond the analytical frameworks studied in risk assessment and risk management contexts" (Aven, 2013, p. 2090). This was by no means stated in Cox's own paper (Cox, 2012 & Aven 2013).

Robust decision making

Robust decision making (RDM) can in many ways be seen as an inversion of a traditional optimum expected utility analysis (EU). In an EU analysis the first step is to characterize the uncertainties. This is followed by a ranking of the uncertainties, and finally a decision is made (Groves & Lempert, 2007). RDM on the other hand

is an iterative process that begins with decision options and then runs the expected utility machinery many times in order to identify potential vulnerabilities of these candidate strategies, that is, combinations of model formulations and input parameters where the strategy performs relatively poorly compared to the alternatives (Groves & Lempert, 2007, p. 76).

This process has three main goals (Groves & Lempert, 2007). The first one is to identify new and/or improved methods and strategies that can perform better than the ones that are currently in use. The second one is to describe the pros and cons of the different strategies relative to each other. This way we can reach the final goal, which is to identify the strategy that is most insensitive to the uncertainties. In other words, we have identified the most robust alternative. To identify the most robust alternative, RDM utilize the power of statistical cluster-finding algorithms like the patient rule induction method (Groves & Lempert 2007; Shortridge et al., 2017). These algorithms identify areas of probability space where the various alternatives have a large difference in performance (Groves & Lempert 2007; Shortridge et al., 2017).

In contrary to the regular RDM method, Lempert presented an alternative perspective on how to manage deep uncertainties (Aven, 2016). Instead of finding the action that is the most robust over all, he suggests that we identify which uncertainties matter the most, which matter the least, which present opportunities and which present threats, and why (Aven, 2016).

A *simplified* example of how robust decision making can be applied as a tool to manage climate change risk will be presented in the following. The given objective is preventing the average global temperature from surpassing the temperature measured at the start of industrial revolution by more than 2.0 degrees Celsius. If at all obtainable, one needs to identify to what extent the current greenhouse gas emissions needs to be reduced to avoid surpassing said temperature. A further analysis of how to achieve this reduction is beyond the scope of this example.

Initially one would gather numerous models of climate change. Secondly one would use similar input data of carbon and other greenhouse gas emissions, followed by an analysis of the different model predictions. In this analysis, one would first proceed to identify each model's recommended reduction of emissions, and then continue by drafting a comparison of all different recommendations. Ultimately, a team of qualified personnel would assess the justifications of the arguments which form the background of each of the different models. This renders them capable of giving more weight to the results of the well justified models, and less to the others. From these results the team would then be able to identify and recommended how much the emissions should be cut by.

Adaptive risk management

From the previous section on RDM it becomes clear that RDM works best if some data or models are available, and it will be hard or even impossible to complete with limited empirical data (Cox, 2012). It is in situations like the latter where the empirical data is limited we turn to adaptive risk management (ARM). This is no to say that ARM cannot be used in situations where relevant models and data are available. This method can be credited to Holling, as he developed and introduced it as a method assess and mange environmental risks in 1978 (Bjerga & Aven, 2015). This is an iterative and structured process that can be used to manage risk characterized by deep uncertainty (Bjerga & Aven, 2015). ARM normally consists of the following elements (Bjerga & Aven, 2015, p. 75):

- *Management objectives that are regularly revisited and accordingly revised.*
- *A model(s) of the system being managed.*
- *A range of management choices.*
- *Monitoring and evaluation of outcomes.*
- *A mechanism(s) for incorporating learning into future decisions.*
- *A collaborative structure for stakeholders' participation and learning.*

ARM is a transparent process where the analyst is aware of some or many of the possible futures that lay ahead, but it can be hard or even impossible to assign probabilities to the ones that are known (Bjerga & Aven, 2015 & Walker et al., 2010). Based on this limited knowledge an adaptable management action is implemented, and its effects on the system are monitored. New responses may be implemented based on the result from the monitoring. Responses like these can either be implemented manually or they can occur automatically. How these responses are implemented depends on how the system is set up. Walker et al. (2010) presented a description of the two types of adaptive responses together with a description of the possible timing these responses can have. For more information, see the detailed description on the next page (Walker et al. 2010):

1. How the adaptation is implemented:
   - *Planned adaptation:* this is a manual adaptation/response that is implemented because the decisionmakers know that the initial conditions have or are about to change. These adaptations are necessary to return the system to, maintain or achieve a desired state.
   - *Autonomous adaptation:* unlike the planned adaptations which are implemented manually, these adaptations are implemented automatically by the system itself as a response to changing conditions. These adaptations are necessary to return the system to, maintain or achieve a desired state.
2. The timing of the adaptation:
   - *Anticipatory adaptation:* these types of adaptations are automatic or manual responses that are implemented prior to a change in conditions to maintain or achieve a desired state.
   - *Reactive adaptation:* these types of adaptations are automatic or manual responses that are implemented after a change in conditions has been registered. These adaptations are necessary to return the system to, maintain or achieve a desired state.

The same simple case as the one that was used to illustrate how RDM works in practice will be used here to illustrate how ARM can be used as tool for managing risk under deep uncertainty. The goal that we want to achieve here is exactly the same as last time, namely to prevent the average global temperature from rising more than two degrees. Since models and data are available the analysis here would be performed in a similar manner as it was in RDM. What separates the two is the final step. Here it is suggested that the emission cuts that were found to be required should be reviewed every 5 years. A further reduction in emissions could be deemed necessary if the effects of the previous cuts have not had the expected results. By doing so we adapt our response to new information. A 5 year timeframe was chosen as climate change is a slow process, meaning that some time must go by for us to see the effects of the emission cuts.

Probability bounds analysis

The concept of bounding probability can be traced back to the middle of the 19th century and further development of these ideas has seen an increase over the last 40 years (Tucker & Ferson, 2003). In probability bound analysis (PBA) probability theory and interval arithmetic is combined to produce probability boxes or p-boxes (Tucker & Ferson, 2003). These p-boxes are structures that allow for the all-inclusive propagation of both *variability* (aleatory uncertainty as represented by frequentist probabilities) and *ignorance* (subjective or epistemic uncertainty) (Ferson & Ginsburg, 1996; Ferson et al., to appear; & Tucker & Ferson, 2003). This type of analysis is particularly useful when the analysts cannot specify one or more of the following (Tucker & Ferson, 2003):

1. Precise parameter values for the input distributions or point estimates in the risk model (min., max., mode, etc).
2. Precise probability distributions for some or all of the variables in the risk model.
3. The dependencies between the variables in the risk model
4. The exact structure of the risk model.

A p-box consists of a pair of distribution functions that are used to circumscribe an imprecisely known distribution function F (Tucker & Ferson, 2003). Say for instance that from previous knowledge it is assumed that a distribution is lognormal, but the precise values of the defining parameters μ (mean) and σ (standard deviation) are uncertain. However, it is known that the true value of these parameters must lie within the following intervals $\mu = [\mu_1, \mu_2]$ and $\sigma = [\sigma_1, \sigma_2]$. Here μ_1, μ_2, σ_1 and σ_2 represents the bounds on the mean and the standard deviation. To plot the bounds in a p-box one would simply have to compute the cumulative distribution function (CDF) enveloping the following four distributions: (μ_1, σ_1), (μ_1, σ_2), (μ_2, σ_1) and (μ_2, σ_2). To illustrate this a numerical example will be used. The distribution is still lognormal, and the bounds are: $\mu = [0.5, 0.6]$ and $\sigma = [0.05, 0.1]$. All that is known at this point is that the true distribution is lognormal with μ somewhere in the interval [0.5, 0.6] and σ somewhere in the interval [0.05, 0.1]. The p-box is displayed in figure 2. This was just an example with one of the commonly used distribution, it is just as easy with others such as normal, uniform, exponential, etc. This method has shown great promise as tool to manage risk in situations where when one or more of the previously listed elements could not be specified (Shortridge et al., 2017 & Flage et al., 2018).

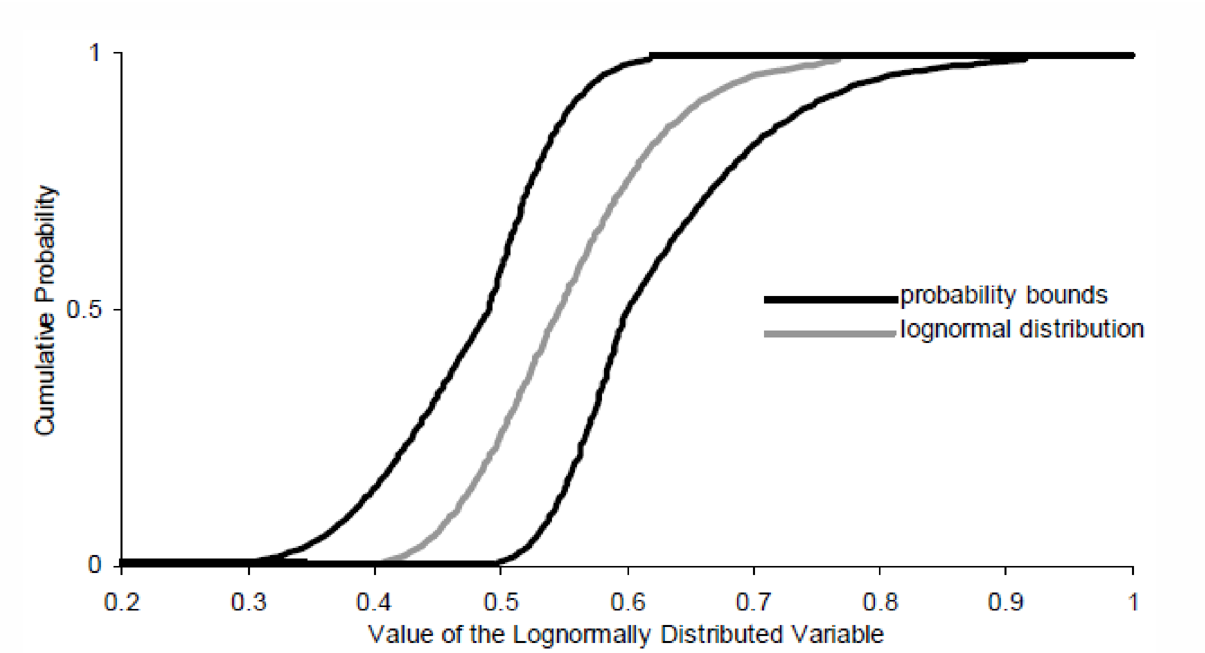
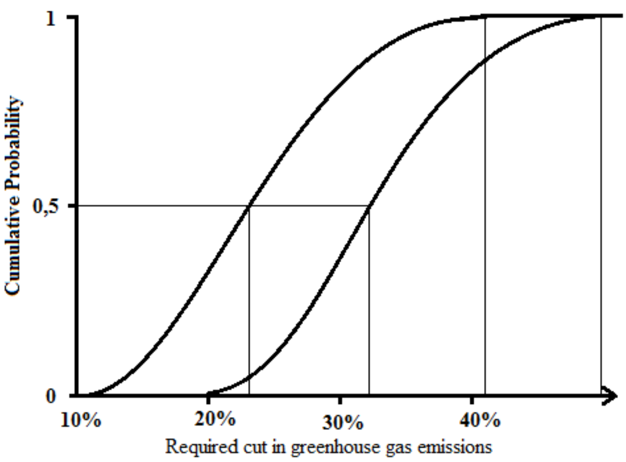


Figure 2: Illustration of the p-box from the previous example (Tucker & Ferson, 2003).

It has been argued by the EPA (U.S. Environmental Protection Agency) that combining frequentist and subjective probabilities to a single probability distribution should be avoided (EPA, 2001). This is because a single probability distribution must be interpretable either as an expression of aleatory or epistemic uncertainty (Tucker & Ferson, 2003). The p-boxes in probability bound analysis model them both, however, this is not a problem here as they are both clearly distinguishable in the end results. The aleatory uncertainty is represented by the CDF on the right and left side, while the epistemic uncertainty is represented by the space between them (Tucker & Ferson, 2003).

The same case that was used to illustrate how the two previous methods can be used to manage risk under deep uncertainty will also be used here. First off, all relevant models on future temperature development and required cuts in greenhouse gas emissions would be gathered and analyzed for data. Then we would use this data to form a p-box which would show the bounds on the required cut in greenhouse gas emissions. This p-box can be seen in figure 3. From this figure we can see that even the most optimistic estimates require emission cuts between [10%, 42%], and they estimate that there is at least a 50% chance that the emissions must be cut by more than 23%. The most pessimistic estimates on the other hand require emission cuts between [20%, 52%], and they estimate that there is at least a 50% chance that the emissions must be cut by more than 33%.

Say, that based on this we choose to cut the emissions by 35%. We can now go back and put this cut level into relevant temperature development models and analyze the data. This data can then be used to form a p-box which shows the bounds on temperature changes with a 35% cut in emissions. This p-box can be seen in figure 4. From this figure we can see that even in the most optimistic estimates the temperature will increase by at least [1.0, 2.5] degrees C and they estimate that there is at least a 20% chance that the temperature will increase by more than 2.0 degrees C. The most pessimistic estimates assume that the temperature will increase by at least [1.5, 3.5] degrees C and they estimate that there is at least a 65% chance that the temperature will increase by more than 2,0 degrees C. It is assumed that the "true" underlying distribution is somewhere within the bounds of these CDFs. However, it is impossible to say anything about which of these estimates is closest to the "true" underlying distribution.



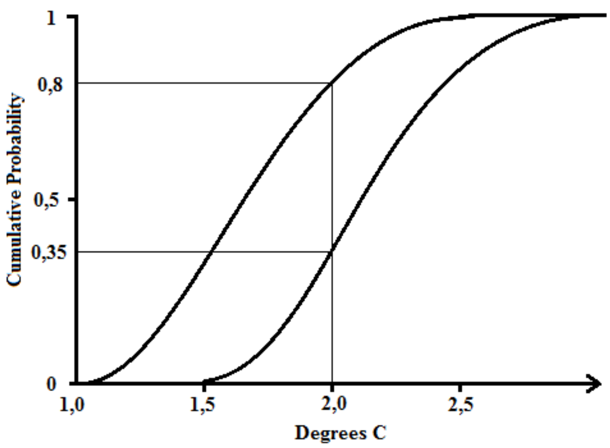


Figure 3(left): P-box with the bounds of estimated required cut in greenhouse gas emissions to prevent an increase in temperature of more than 2 degrees.

Figure 4(right): P-box with the bounds of estimated temperature increases with a 35% cut in greenhouse gas emissions.

Managerial review and judgement

The overall goal of a risk assessment is to provide the decisionmaker with a clear and detailed risk description (Aven, 2013). The results of the risk assessment will not explicitly tell the decisionmaker what to do, as there are often several aspects affecting a decision which are not reflected in a risk assessment (Aven, 2013). Before a final decision is made the decisionmaker has to take all of these aspects into account (Aven, 2013). These aspects may include but are not limited to other benefits related to the situation in question, and political/strategic concerns. Another important thing the decisionmaker must take into account before making a decision is the SoK behind the assumptions. This must be reviewed together with the results of risk analysis to justify whether the results are reliable or not (Aven, 2013). A decisionmaker also have to evaluate how relevant the results of the risk assessment are to the decision problem at hand (Aven et al., 2007). This process is known as a managerial review and judgement, and its goal is to bridge the gap between the risk assessment and the decision (Aven, 2013). The time and effort put into this process can in many cases be reduced by implementing measures like risk acceptance criteria (RAC). RAC represents the upper limit of risk that is acceptable in a given situation (Aven, 2013). The use of RAC as a means to reduce the time and effort put into this process should only be used if uncertainty is low (level 1), since the risk is not well reflected by probability numbers alone if the uncertainties involved are large (Aven, 2013). In a situation characterized by moderate or deep uncertainties the group involved in the managerial review and judgement process have to be able to see "beyond the narrow technical criteria when making judgements about the risk being acceptable or not" (Aven, 2013, p. 2086), and to pay more attention to the justification of the assumptions and the strength behind these assumptions (Aven, 2013).

To illustrate this process, and to link it to the previous methods we will continue with the same example, only in this situation we have the results of risk assessment. These results were as follows: to prevent the global temperature from increasing more than 2 degrees we have to cut the green house gas emissions by 25% within 2035, and by 50% within 2055. The decisionmakers will now have to take several aspects of this case into account and weigh them up against each other before a just decision can be made. First of all, they have to review the results of the assessment in light of the assumption, to assess whether the results of the assessment are reliable. Secondly, they have to weigh the pros (lower sea levels, less extreme weather, better protection of vulnerable ecosystems, less air pollution in the big cities, etc.) against the cons (lower productivity, higher restrictions, reduction in the use of transportations devices fueled by hydrocarbons, etc.). A few other aspects they have to take into account are the political implications, how it affects the world economy and if can cause zones of conflict (war zones). If they decide to go through with the emission cuts they will have to find a way to make it fair for everyone, as there are some nations who have contributed way more to the greenhouse gas emissions than others. As this is a simplification of a very complicated decision it does not include every factor the decisionmakers will have to take into account in addition to the risk assessment before a decision is made, but it demonstrate how the process is done.

3. Development of an extended deep uncertainty categorization taxonomy

The purpose of this chapter is to present and discuss the challenges that are present in the previously mentioned uncertainty taxonomies, and to develop an alternative taxonomy where these challenges are no longer an issue.

3.1. Challenges with the current categorization taxonomies

As pointed out by Aven, there are some clear challenges with the deep uncertainty categorization taxonomy Courtney introduced (Aven, 2013). He stated that "several critical questions can be raised regarding its foundation" (Aven, 2013, pp. 2082), including the use of probabilities and the lack of a specified interpretation of probability, and the lack of a knowledge dimension (SoK). To fix this he suggested a new and simplified taxonomy which he also argues increase precision. In this modified taxonomy he defines two uncertainty categories, Level A and Level B (Aven, 2013). Level A represent every situation where the risk analyst believes in the existence of frequentist probabilities P_f that perfectly describes the variation in a system, and that the probability model introduced by the risk analyst is an accurate representation of the "true" underlying frequentist distribution P_f. Level B represents every other situation and they are characterized by deep uncertainty (Aven, 2013). Following this path, he argues that Level A only refers to situations of controlled experiments that can be repeated a large number of times under similar conditions to test and verify suggested hypotheses and probability models (Aven, 2013). According to this definition of deep uncertainty every situation where there is a lack of knowledge is characterized by deep uncertainty. This is not in line with the definitions of deep uncertainty presented by Lempert et al. and Cox (Cox, 2012 & Walker et al., 2017). These authors argue that the concept of deep uncertainty is described by a situation of significant model uncertainty even up to the point where no model is available at all, rendering it hard or even impossible to say anything about the outcome of the given situation. There is clearly a significant gap between a situation where you have so much knowledge that it enables you to build a model that can predict nearly every outcome of an event with reasonably accurate probabilities and a situation where you have so little information that it is impossible to build a probabilistic model or even say anything about the outcome of an event. The author of this thesis would therefore argue that even though Aven's modified taxonomy may be simpler, Courtney's uncertainty taxonomy with its four levels of uncertainty is certainly both more precise and more in line with previously mentioned definitions of deep uncertainty. However, this does not mean that his taxonomy is without flaw. As pointed out by Aven it lacks in the specification of knowledge strength for each of the different levels as well as a specification for the interpretation of uncertainty.

Later on, in the same paper Aven (2013) presented a second alternative to Courtney's taxonomy, with three levels of uncertainty and the incorporation of the knowledge dimension, see table 3. This alternative implies that for a situation to be characterized by deep uncertainty the knowledge base of the assessor must be poor. This is more in line with the deep uncertainty definitions presented by Lempert et al. and Cox (Cox, 2012 & Walker et al., 2017). However, what this taxonomy gains in simplicity it lacks in detail as it is not possible to score high in both of these areas. It is basically just a light extension of the previous taxonomy introduced Aven, with one additional uncertainty level that covers the significant gap in the first taxonomy. This gap is still just as wide, so the author of this thesis would argue that this step only clarifies that situations where the background knowledge is moderate is no longer in the deep uncertainty category. There is clearly room for a splitting of or a more detailed moderate uncertainty level.

A similar argument can be used against Courtney's categorization taxonomy. The gap between level 2 and level 3 is quite wide. A system belonging to level 2 has several outcomes and each outcome has a point estimate of its probability. A system belonging to level 3 also has several outcomes, but nothing can be said regarding the probabilities. It is not like we go from having pin point accurate probabilities for every outcome to having no clue as to what the probabilities for each outcome can be. There is a step in between level 2 and 3 where imprecise probabilities can be assigned to the outcomes. These imprecise probabilities can be assigned in situations where we have limited data, poor measurements and subjective information (Beer et al., 2013). To account for the gap in the categorization system the author of this thesis suggest that a fifth level of uncertainty is added. This level is to be placed in between level 2 and level 3, and it will represent every system where the risk analyst does not have enough information to assign point estimates of the probabilities, but he/she still has enough information to say something regarding the probabilities. According to Lempert et al.'s definition of deep uncertainty (Walker et al., 2017, p. 5):

> *The condition in which analysts do not know or the parties to a decision cannot agree upon (...) the probability distributions to represent uncertainty about key parameters in the models*

it can be argued that this new uncertainty level can in some situations be seen as a part of the deep uncertainty category, while in other situations it will fall under the moderate uncertainty category. It comes down to the width of the probability intervals, and if the knowledge of the risk analyst is strong enough to justify a distinct probability distribution for the probability interval(s). Consequently, this fifth level of uncertainty will fall under both moderate and deep uncertainty, see table 4.

To better handle the previously mentioned challenges with the alternative taxonomy presented by Aven (2013) as well as the challenges with the original deep uncertainty taxonomy presented by Courtney (2001), the author of this thesis proposes a merger between the two taxonomies, with the addition of a fifth uncertainty level. Although this is a step away from the simpler approach taken by Aven, this new taxonomy will benefit from the incorporation of more nuanced representation of the varying degrees of uncertainty, as well as a more detailed explanation of each level.

3.2. An extended deep uncertainty categorization taxonomy

It is the opinion of the author that the only advantage of the two alternative classification taxonomies introduced by Aven (2013) compared to Courtney's (2001) original taxonomy is the incorporation of the judgement of the strength of knowledge. Yes, Aven's (2013) taxonomy is certainly much simpler, but the downside here is that it almost entirely ignores the nuances in the various degrees of uncertainty. For instance, he only recognizes one level of deep uncertainty in both of his alternatives, while Courtney has two. By merging Aven's ideas regarding knowledge strength with Courtney's taxonomy and adding a fifth level of uncertainty, we are left with a more detailed description of the varying degrees of uncertainty. A few other aspects like *degree of uncertainty*, *justification of conditions (i) and (ii)* (see p. 11 or 23), *the possible occurrence of black swans¸ probability interpretation* and *a system example* has been added to increase the level of detail. The fading color from green to red indicate an increase in uncertainty. The various levels of the new uncertainty taxonomy are described below, and the alternative taxonomy is presented in table 4.

Level 1 uncertainty: a clear enough future. In this first level of uncertainty, which falls under the low uncertainty category, there is a narrow range of possible outcomes and the uncertainty related to each outcome can be adequately modeled and described by frequentist probability models (aleatory uncertainty). These models are backed up by good evidence and a strong background knowledge. This does not mean that the future is perfectly predictable in every case, but we can for all practical purposes ignore black swans and conclude that conditions (i) and (ii) are justified (Level A', see page 11 or 23) and will hold with reasonable accuracy. The reason why we cannot say that the future in this level of uncertainty is perfectly predictable in every case is due to the possibility of minor errors in the probability models caused by the fact that perfect measurements and models are impossible to make. An example of a situation that falls under this level of uncertainty is flying with a commercial jet as a passenger.

Level 2 uncertainty: alternate futures with point estimates. In this second level of uncertainty, which falls under the moderate uncertainty category, it is possible to identify a limited range of possible outcomes where one of them will occur, except for few instances where the outcome is a black swan. The strength of the background knowledge falls somewhere between strong and moderate, and the uncertainty related to each outcome can in most cases be adequately modeled and described by frequentist probability models (aleatory uncertainty). Conditions (i) and (ii) are justified but will in some cases not hold (Level A' or A''). Two examples of situations that fall under this level of uncertainty are playing on a fair slot machine (Level A') and playing on a rigged slot machine (Level A'').

Level 3 uncertainty: alternate futures with imprecise probabilities. In this third level of uncertainty, which can fall under either the moderate or the deep uncertainty category, a limited range of possible outcomes where one of them will most likely occur can be determined. The probability of occurrence related to the outcomes cannot be given as a point estimate but rather as probability intervals. These probability intervals are thought to contain the true probability of occurrence, but this does not necessarily have to be true. The strength of knowledge that forms the foundation for the justification of the bounds on the probability intervals fall somewhere between moderate and poor and is based on some dominating explanations and belief. This is reflected through the width of the probability intervals: *Wider intervals* suggest that the analyst(s) have a poor understanding of the underlying causes and a poor strength of background knowledge. *Shorter intervals* suggest that the analyst(s) have a moderate understanding of the underlying causes, and a moderate strength of background knowledge. The risk analysts *may* believe that there exists a true underlaying frequentist distribution P_f (i), but the knowledge strength is too poor to come up with a probability model that can accurately describe each outcome, meaning that this group fall under level B' (see page 23 for a description of B'). Black swans are also relevant here since a limited range of possible outcomes can be suggested, but none of them are guaranteed to occur. This level can be seen as a bridge that fixes the gap between level 2 & 4. Two examples of situations that fall under this level of uncertainty are the spreading of a new disease, and the market reception of a new product.

Level 4 uncertainty: a range of futures. In this fourth level of uncertainty, which falls under the deep uncertainty category, a limited range of possible outcomes where one of them will most likely occur can be determined. It is impossible to describe the uncertainty related to each outcome with probabilities. The risk analyst(s) *may* believe in the existence of a true underlaying frequentist distribution P_f, justifying condition (i), but the knowledge strength is too poor to come up with any fitting probability model, meaning that this group too fall under level B'. The knowledge strength here is poor, but since it is possible to determine limited range of possible outcomes there must exist some dominating explanations and beliefs. Black swans can occur in this level, since none of the preidentified outcomes are guaranteed to occur. An example of a situation that fall under this level of uncertainty could be unstable macroeconomic conditions.

Level 5 uncertainty: an unknown future. In this fifth and final level of uncertainty, which falls under the deep uncertainty category it is impossible to determine any future outcomes. All that is known is that we don't know anything, and this ignorance is recognized. There is no confidence in either condition (i) nor (ii), which means that this group fall under level B''. Here it would be pointless to refer to surprising outcomes as black swans, since all outcomes will be surprising, and we are aware of that, so they are no longer black swans (Aven, 2013). Some examples of situations that fall under this level of uncertainty could be the outcome of major technological, economic or social discontinuities.

For convenience, all the notations that are used in table 4 are described below. The ones that are written in italic are direct citations from Aven (2013).

(i) *The existence of frequentist probabilities P_f.*

(ii) *The probability model introduced is an accurate representation of the "true" underlying frequentist distribution P_f.*

(1) *Level A': Confidence in conditions (i) and (ii) (they have been justified), and the conditions hold (with reasonable accuracy).*

(2) *Level A'': Confidence in conditions (i) and (ii), but the conditions do not in fact hold.*

(3) Level B': Some confidence in condition (i) only, but the condition may not hold. No confidence in condition (ii).

(4) Level B'': No confidence in either condition (i) or (ii).

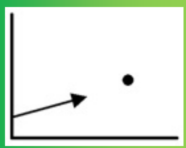
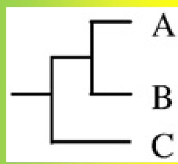
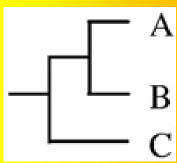
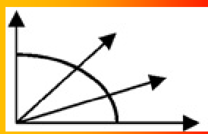
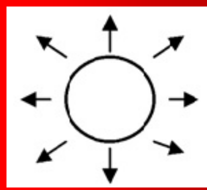
(a) *Situations where we have strong evidence (we can for practical purposes conclude that A' applies)*

(b) *Other situations-characterized by some dominating explanations and beliefs.*

(c) Situations where all we know is that we don't know.

To avoid unnecessary splitting of the alternative uncertainty taxonomy it will not start below, but rather on the next page.

Table 4: A modified uncertainty classification taxonomy. Compared to Courtney's taxonomy, newly added features are written in italic. The gradual shift in color from green to red represents the increase in uncertainty.

Level	**Level 1**	**Level 2**	***Level 3***	**Level 4**	**Level 5**
Type of uncertainty	*Low uncertainty*	*Moderate uncertainty*		*Deep uncertainty*	
Strength of knowledge	*(a)*		*(b)*		*(c)*
	Strong		*Moderate*	*Poor*	*Non existent*
Justification of conditions (i) and (ii)	*Conditions (i) and (ii) are justified*		*Conditions (i) and (ii) are not justified*		
	Level A'	*Level A' or A''*	*Level B'*		*Level B''*
Black swans?	*No black swans*		*Black swans may occur*		*No black swans*
Probability interpretation	*Aleatory*		*Both aleatory and epistemic*	*No use of probabilistic models*	
Context	A clear enough future 	Alternate futures (with probabilities) 	*Alternate futures (with imprecise probabilities)* 	A multiplicity of plausible futures 	Unknown future 
System model	A single binomial system model	A single system model with a probabilistic parameterization	*Several system models, with different structures*	Several system models, with different structures	Unknown system model; know we don't know
System outcomes	A Point estimate and confidence interval for each outcome	Several sets of point estimates and confidence intervals for the outcomes, with a probability attached to each set	*Several sets of point estimates and confidence intervals for the outcomes, with imprecise probabilities attached to each set*	A known range of outcomes	Unknown outcomes; know we don't know
Weights on outcomes	A single estimate of the weights	Several sets of weights, with a probability attached to each set	*Several sets of weights, with a probability attached to some of them*	A known range of weights	Unknown weights; know we don't know
System example	Flying with a commercial jet as a passenger	Playing on a fair slot machine (Level A'), or on a rigged slot machine (Level A'')	*Demand for new products or services*	Unstable macroeconomic conditions	The outcome of major technological, economic or social discontinuities

4. Adaptation of a method to assess risk characterized by level 3 uncertainty

The purpose of this chapter is to introduce the reader to a method that can be used to assess risk that fall under the previously introduced third level of uncertainty, and to illustrate how this method can be used by using it to assess risk in a hypothetical case.

4.1. Method adaptation

Probabilistic risk assessment has proven itself to be a great tool for evaluating risk in complex engineering systems. However, to obtain realistic result from these prediction models they require accurate information and appropriate mathematical modeling and quantification. This can be challenging in situations where the underlying knowledge to support probabilistic representation of uncertainties is limited. Such limitations may include imprecise measurements, sparse data and subjective information. Situations where there are limitations in available information has led to an increasing concern among risk analysts that traditional probabilistic risk assessment may not be sufficient for probabilistic modeling (Shortridge et al., 2017; Beer, et al. 2013). A number of alternative methods have been suggested to manage situations of deep uncertainty, see chapter 2.2.3. The main focus of this chapter is not to identify and develop a method which is applicable in every deep uncertainty situation, but rather to adapt previous thoughts and ideas from risk assessment to a method that can be used to assess risk in situations that fall into the third uncertainty level.

A situation that fall into the third uncertainty level is characterized by imprecise probabilities. Several recently published books and papers have argued that regular probability theory generate to precise results when the background knowledge supporting the probabilities is poor (Aven, 2010). To illustrate this, an example similar to the one used by Ferson and Ginzburg (1996) is introduced. A parameter σ_1 is a number somewhere between 0 and 1, and a parameter σ_2 is a number somewhere between 1 and 2, and no more information is given. What is the value of the sum $\sigma_1 + \sigma_2$?

One way to solve this is by using interval analysis. Here the total range of values the sum can have is identified by finding the smallest and largest possible sum. $(\sigma_1 = 0) + (\sigma_2 = 1) = 1$ is the smallest value the sum can have, and $(\sigma_1 = 1) + (\sigma_2 = 2) = 3$ is the largest. No number combination will fall outside this interval. Consequently, the answer is that the sum lies somewhere between 1 and 3. See a depiction of the interval [1, 3] in figure 5. It is important to note that this interval is not the same as a uniform probability distribution, since it is impossible to say anything about the probability distribution based on the available information.

Another way to try and solve this is by letting σ_1 and σ_2 be randomly varying numbers and sum them together using Monte Carlo simulations under the assumption that the parameters are independent. As no information is given regarding the probability distribution within the intervals it is easy to assume that every number within each interval is equally likely to represent the true value. Following this logic each parameter is set to have a uniform distribution within the given bounds and the simulation is run. The result of such simulations gives the impression that the extreme values 1 and 3 are much less likely to occur than the value 2. See the probability distribution illustrated in figure 6.

It is argued by Ferson and Ginzburg (1996) that the probabilistic distributions are incorrect, as they assume more information than what was given in the original question. This is clear from the probability distribution seen in figure 6. When no information is given regarding the probability distributions in parameters σ_1 and σ_2, the use of random numbers will give a false impression of the knowledge strength of the risk analyst. This method implies that the risk analyst knows more than he/she actually does, since the simulation gives the impression that some values are more likely than others. When no information is given regarding the probability distribution of parameters $\sigma_1 + \sigma_2$ it is wrong to assume that some values are more likely than others. Ferson and Ginzburg (1996) writes:

> *In this sense, they are the result of wishful thinking, rather than a careful analysis of what is actually known. This example illustrates what may be a widespread problem with applying classical probability theory in risk analyses where the relevant empirical information is sorely incomplete (as is usually the case).*

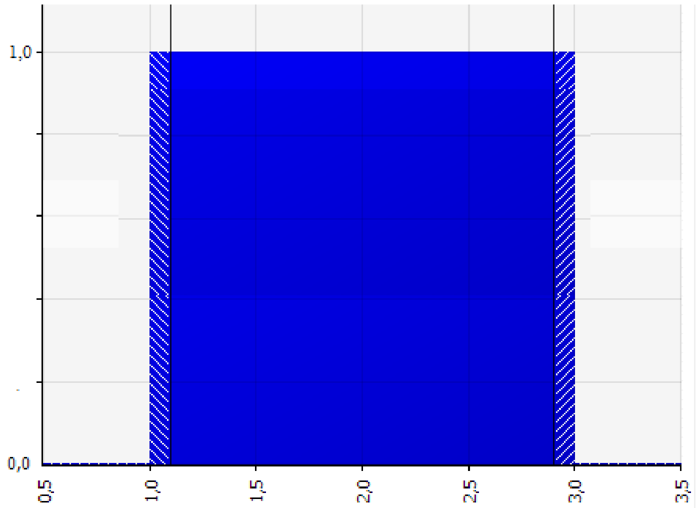


Figure 5: Illustration of the interval [1, 3] which is guaranteed to contain the sum of σ_1 = [0, 1] and σ_2 = [1, 2].

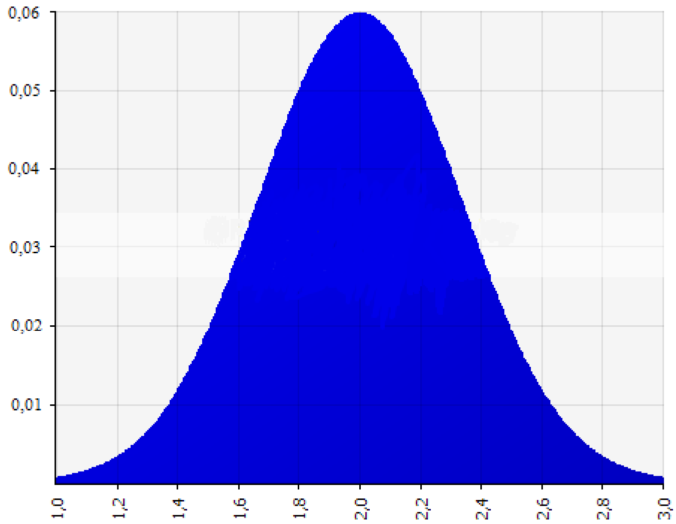


Figure 6: Probability distribution of the sum of the random variables σ_1 = uniform(0, 1) and σ_2 = uniform(1, 2) under the assumption that the parameters are independent.

Ferson and Ginzburg (1996) conclude that classical probability theory alone only provides methods appropriate for assessing and propagating random variability, and not for assessing and propagating epistemic uncertainties and ignorance. They suggest that probability bounds analysis should be used to propagate uncertainty in risk analysis in situations where both variability and ignorance is present (Ferson & Ginzburg, 1996). In a later article Beer et al. (2013) provides an overview on the developments on methods which involve the use of imprecise probabilities. From this overview it is clear that there is a pre-existing framework of imprecise probabilities that provide a mathematical basis to deal with such situations. This includes, but is not limited to: interval analysis, probability bounds analysis, bounds based on evidence theory and fuzzy probabilities (Beer et a., 2013; Aven, 2010). Beer et al. (2013) suggest probability bounds analysis as the preferred method of choice in situations where:

Parameters of a probabilistic model, the distribution type or, in a non-parametric description, the curve of the cumulative distribution function may only be specified within some bounds. This imprecision may arise, for example, when conflicting information regarding the distribution type is obtained from statistical tests, that is, when the test results for different distributions as well as for compound distributions thereof with any mixing ratio are similar. These test results do not provide grounds for assigning probabilities to the model options. If no additional information is available in such situations, the most suitable approach for modeling the cumulative distribution function is as a set of distributions. In the simplest form, this implies the use of intervals for the distribution parameters.

From these papers it is clear that probability bounds analysis is the preferred method of choice to analyse risk when facing a situation which is characterized by level 3 uncertainties. However, a probability bounds analysis alone may not be enough. Every parameter used to give a final numerical value in a standard quantitative risk analysis is based on a set of assumptions. To present an informative overview of the risk picture it is vital that these assumptions are justified, and that the strength of the assumption justification is clearly stated (Flage & Aven, 2009). To illustrate this a simple example will be used. An investor has asked two different risk analysts to analyze the likelihood, P(A), of him losing money on a given investment. The first analyst presents his/her results in the following way: P(A) = 0,3, while the second analyst present his/her like this: P(A|K) = 0,5, where K represents the knowledge the analysis is based upon. The second analyst has more than 30 years of experience with this type analyses and he judge the strength of his relevant background knowledge to be strong. The second analysis is clearly more informative as it gives some information on why the analyst came up with that exact result, and to what degree the analysis can be seen as credible.

Assumption justification and an assessment of the strength of the assumption justification is especially important in a situation characterized by level 3 uncertainty, to justify the setting of clear bounds on large uncertainties. This will aid in justifying to what degree the analysis can be seen as credible, and if it is of any value as a tool in decision support. It is therefore important that the decisionmakers are made aware of what grounds the assumptions are based upon and to what extent these grounds can be seen as solid. Consequently, the method used in this thesis will combine probabilistic bounds analysis with a justification of the assumptions and a validation of the strength of knowledge these assumptions are based upon. This way the final risk picture will be as broad, balanced and informative is it possibly can be under the given circumstances. A similar approach was taken by Flage et al. (2018) when they compared probabilistic bounds analysis with a subjective probability analysis. Along with their comparison they also presented a set of different ways to produce inputs a general risk assessment (Flage et al., 2018):

1. *Constraints (e.g. positive lifetime and restoration times, and probabilistic inequalities)*
2. *Hard data (on e.g. production volumes, lifetimes and restoration times)*
3. *Modeling of the system performance (e.g. of the availability of the system)*
4. *Bayesian updating*
5. *Aggregation (fusion) (combining data from several sources)*
6. *Judgements (degrees of belief of unknown quantities)*

The risk assessment method suggested in this thesis can be described as a stepwise process inspired by the risk assessment methods presented and used by other authors (Aven, 2014; Flage et al. 2018; Shortridge et al. 2017; Aven, 2013 & Aven, 2016), combined with the ideas of Ferson & Ginzburg (1996) and Flage et al. (2018).

Step 1: Problem conceptualization. In this step the problem is conceptualized, and the initiating event(s) and their causes and consequences are identified. This can be illustrated by a simple event tree and/or bow tie diagram. The problem conceptualization is done to get a better overview of the problem at hand.

Step 2: Identification of probabilities. In this step the relevant probabilities related to the causes and consequences are identified. The availability of accurate probabilities is limited as this is method is used for analysis situations characterized by large to deep uncertainties. If a set of probability intervals is identified it is important that the sum of the lowest values is below 100% and that the sum of the highest values is above 100%

Step 3: Setting up the p-boxes. In this step the probability bounds that are to be used to make the enveloped cumulative distribution functions are identified, followed by the computation of the p-boxes. This is done in accordance with the mathematics described by Tucker and Ferson (2003). If the probability bounds that are to be used in the making of the p-boxes cannot be easily identified, computational methods can be used to make every possible CDF and envelop the final result.

Step 4: Assumption justification. In this step the previously used assumptions must be justified, and the justification/strength of knowledge behind the assumptions is assessed. This assessment should follow the principles and guidelines presented by Flage & Aven (2009).

Step 5: Analyzing the result of the p-boxes in light of the assumption justifications. In this step the p-boxes are analyzed in light of the assumption justification to see whether or not the p-boxes can be used as a tool for decision support, or if the information laying the foundation of the p-boxes is too weak.

This method could be extended from method to assess risk under deep uncertainty to a method that can be used to manage risk under deep uncertainty by following the idea presented by Aven (2013) which is adding a final step, namely managerial review and judgement.

4.2. Case study

Here a case is presented to illustrate how probability bounds analysis can be used alongside a qualitative assessment of the assumptions to assess the risk in a situation characterized by level 3 uncertainties. Even though this example is simple, it is still realistic and interesting from a practical decision oriented perspective.

4.2.1. Case description

A small costal community is in risk of destruction by tsunami, and they wonder if their current safety systems should be upgraded, so a risk analysis is required as a tool to aid in the decision process. The focus of this analysis will be on the total fraction of deaths and survivors, and not on material destruction. This community consists of around 1000 people living at the end of a fjord, and they face severe injury or even death caused by a giant wave rolling over the settlement. This wave is in turn caused by a large landslide coming from the mountains 60 km out in the fjord. Several geologists have previously been consulted regarding the timeframe of this incident, and all they can say is that they assume there will be landslide sometime in the next 1000 years. They cannot say anything regarding the probability distribution of this timeframe.

Due to the composition of the landmasses and its geographic location in the mountains it is hard to say exactly how much of these landmasses will fall out and cause the formation of a wave. There is strong disagreement among the experts related to the size of the landslide, and from a risk assessment perspective this poses a challenge since the volume of the landslide is proportional to the height of the resulting wave, and the height of the wave is directly linked to the level of destruction it causes. Experts assume that the landslide will cause a wave that is between 10 and 30 meters high. For simplicity the wave heights are divided into four different sections, where each section represents a five meter interval, see table 5. These experts have also presented several different prediction models for a wave to fall within these intervals. The upper and lower probabilities for the occurrence of each of the wave height intervals found in these models is presented as probability intervals in table 5. It is important to note that nothing can be said about the probability distribution within these probability intervals. The only information they give is that the probability of occurrence is assumed to be somewhere within these intervals.

Table 5: Overview of probability of occurrence.

Wave height (m)	**Notation**	**Min. probability of occurrence**	**Max probability of occurrence**
10 to 14,99	W_1	10 %	20 %
15 to 19,99	W_2	15 %	40 %
20 to 24,99	W_3	25 %	60 %
25 to 30	W_4	10 %	50 %
Sum		**60 %**	**170 %**

From an energetic point of view, it is clear that the different wave heights will have different impacts on the community. Larger waves will reach further in on the mainland and cause more destruction than its smaller counterparts. For simplicity we assume that waves within the same height interval have the same reach and cause the same level of destruction. To prepare the community for the occurrence of a tsunami, tsunami drills are executed at random two times a year, and this has been done each year for the past 70 years. The fraction of survivors for the different wave height intervals is estimated after each drill and can be seen in table 6. These survival calculations are based on the location of the residents at the moment the wave would have reached land, and the actual inland reach of the different wave heights. To get a worst case scenario all residents that have not reached a shelter and are within the reach of the wave at the time it strikes are presumed dead. This survival data seems to follow four distinct normal distributions with the means, μ, and standard deviations, σ, given in table 6. The part of the distribution that reach above 100% is added to the fraction of times 100% of the population survives, and the part of the distribution that reach below 0% is added to the fraction of times 0% of the population survives. This addition to the fraction of times 0% or 100% of the population survives account for an addition less than 0,15% in both ends, due to the three sigma rule (Grafarend, 2006).

Table 6: Fraction of resident who survive at the different wave height intervals.

Wave height (m)	**Notation**	**Standard deviation (σ)**	**Mean (μ)**
10 to 14,99	FS_1	0,05	0,85
15 to 19,99	FS_2	0,10	0,70
20 to 24,99	FS_3	0,16	0,50
25 to 30	FS_4	0,12	0,35

The state is considering upgrading the safety systems by adding more shelters so that nobody that lives in the affected area is more than one kilometer from a shelter. It is estimated that this will increase the survival rate to somewhere between 90% and 100% for all wave heights. The total cost of this upgrade is estimated to be around 750 million NOK.

4.2.2. A probability bounds analysis approach

Step 1: Problem conceptualization. The problem and the situation at hand has been laid out in detail in the previous sub-chapter and will not done again to avoid unnecessary repetition. Instead an event tree will be presented to illustrate the possible chain of events, see figure 7. No focus will be given to the identification of possible causes of the landslide, as this is beyond the scope of this example.

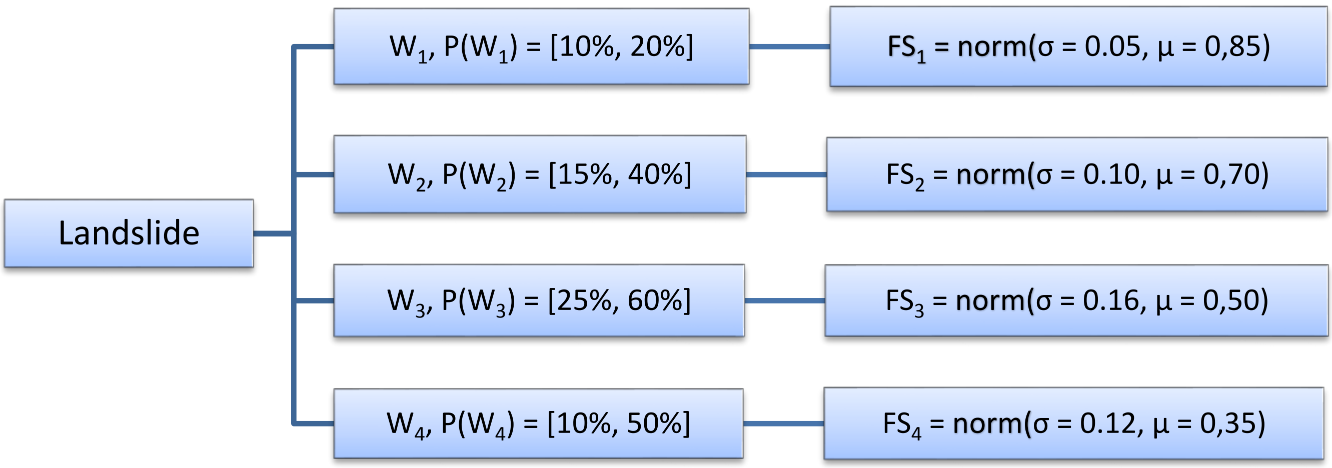


Figure 7: An overview of the situation in the form of an event tree.

Step 2: Identification of probabilities. As with the previous step, most of this step has also been done in the previous sub-chapter, and to avoid unnecessary repetition it will not be done again. The relevant probabilities will instead be presented as figures. The probability intervals that represent the uncertainty regarding the occurrence of waves belonging to the different height intervals can be seen in figure 8. These probabilities represent epistemic uncertainty, as their imprecision is caused by a lack of knowledge, or ignorance. They are based on the best estimate of several experts in the field of geology.

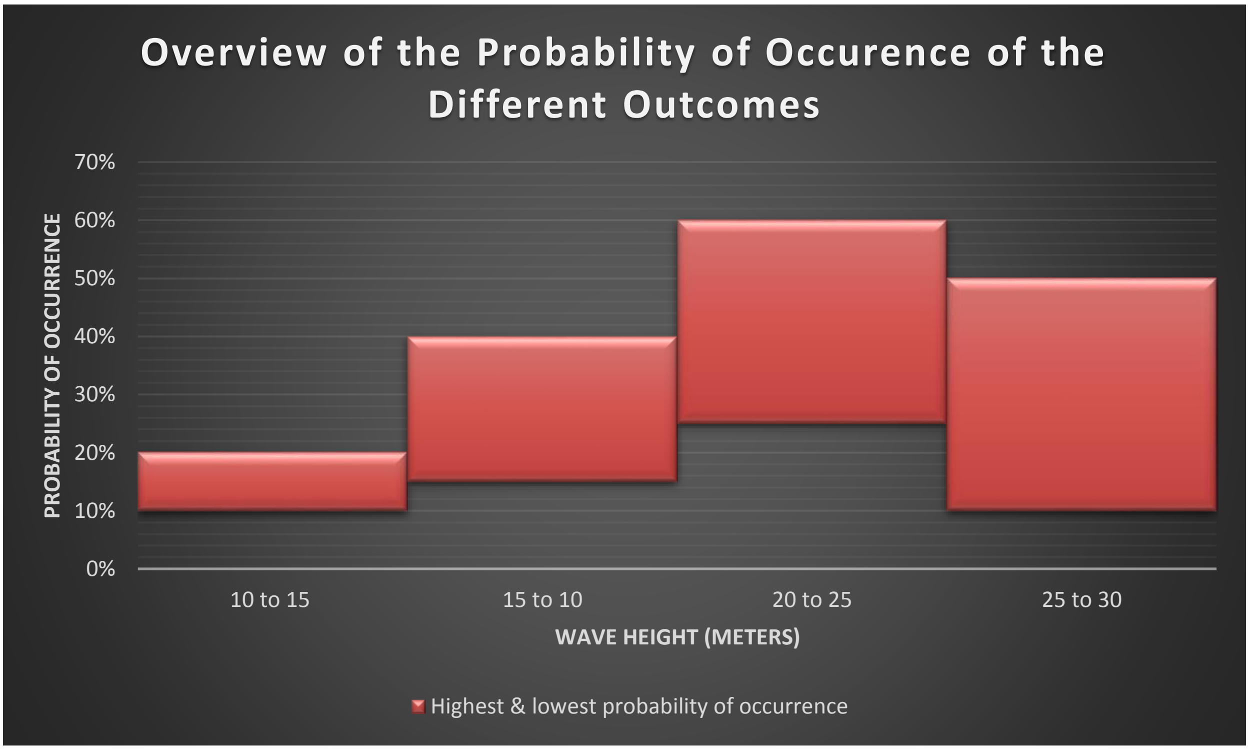


Figure 8: Overview of the probability of occurrence of the different outcomes.

The probability distribution of the fraction of survivors at different wave heights can be seen in figure 9. These probabilities are based on historical data and are not subjective judgements. They represent aleatory uncertainty as they stem from natural variations in the system, and they cannot be made more accurate by acquiring more knowledge.

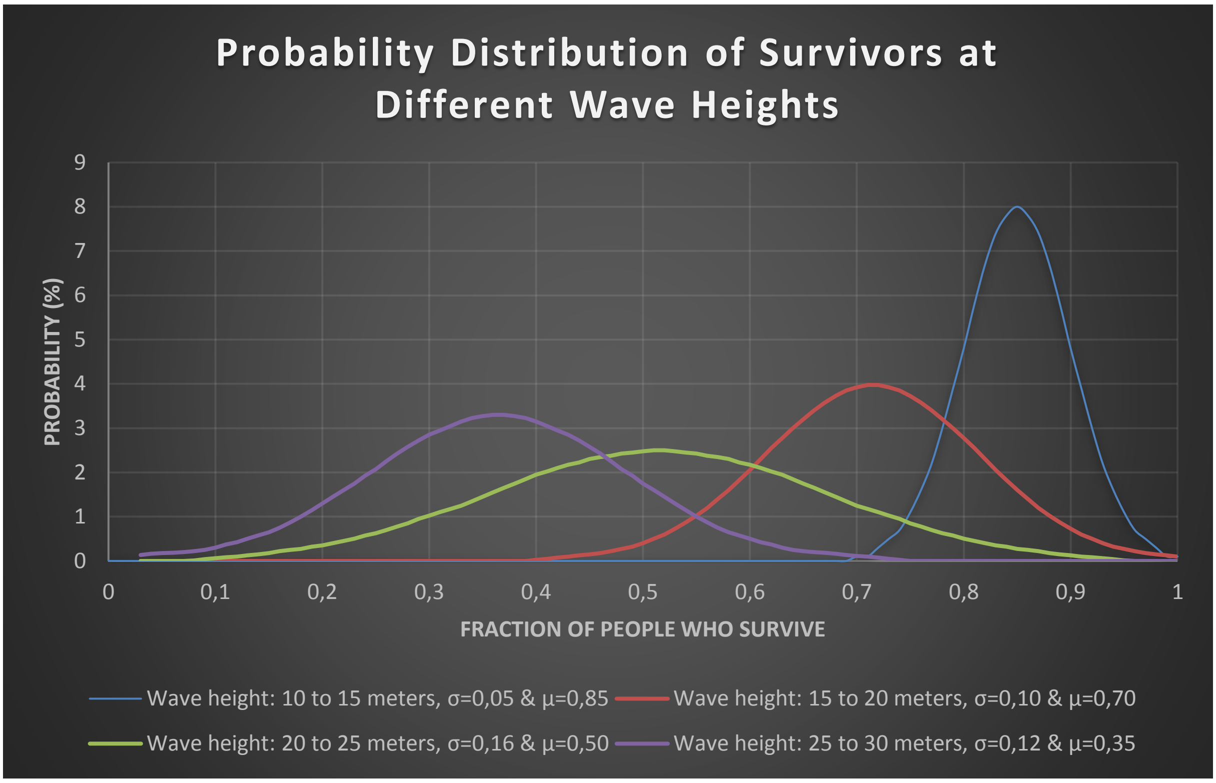


Figure 9: Probability distribution of survivors at different wave heights.

Step 3: Setting up the p-boxes. From the survival probabilities presented in figure 8 it is clear that in a best case scenario (highest fraction of survivors) the probability of the occurrence of a wave that is between 10 and 15 meters, $P(W_1)$, and one that is between 15 and 20 meters, $P(W_2)$, should be as high as possible, while the probability of the occurrence of a wave that is between 25 and 30 meters, $P(W_4)$, should be as low as possible. The probability of the occurrence of a wave that is between 20 and 25, $P(W_3)$, should take the residual value so that:

$$P(W_1) + P(W_2) + P(W_3) + P(W_4) = 100\%$$

The same logic only reversed can be used to find the probabilities required in a worst case scenario (lowest fraction of survivors). In this case $P(W_4)$ should be as high as possible, while $P(W_1)$ and $P(W_2)$ should be as low as possible. $P(W_3)$ should take the residual value so that the equation above is fulfilled. This reasoning leaves the following probability bounds, see table 7 for exact values and figure 10 for an illustration of the cumulative probabilities.

Table 7: Probability bounds on the worst and best case scenarios.

Wave height (m)	**P in the best case scenario**	**P in the worst case scenario**
10 to 14.99	0.20	0.10
15 to 19.99	0.40	0.15
20 to 24.99	0.30	0.25
25 to 30	0.10	0.50

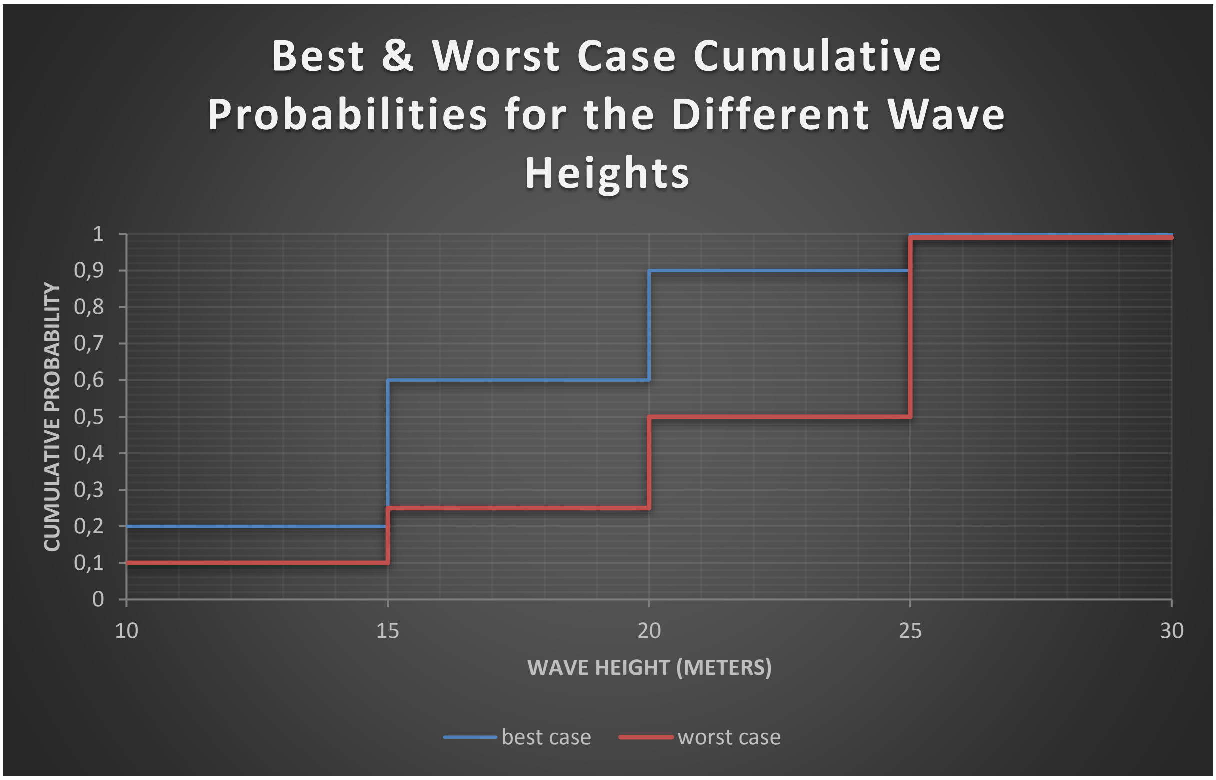


Figure 10: Best & worst case cumulative probabilities for the different wave heights.

By combining the probability values for the best and worst case scenarios, with the probability distribution of survivors at different wave heights the following p-box is revealed, see figure 11. The exact calculations behind this can be found in appendix A.

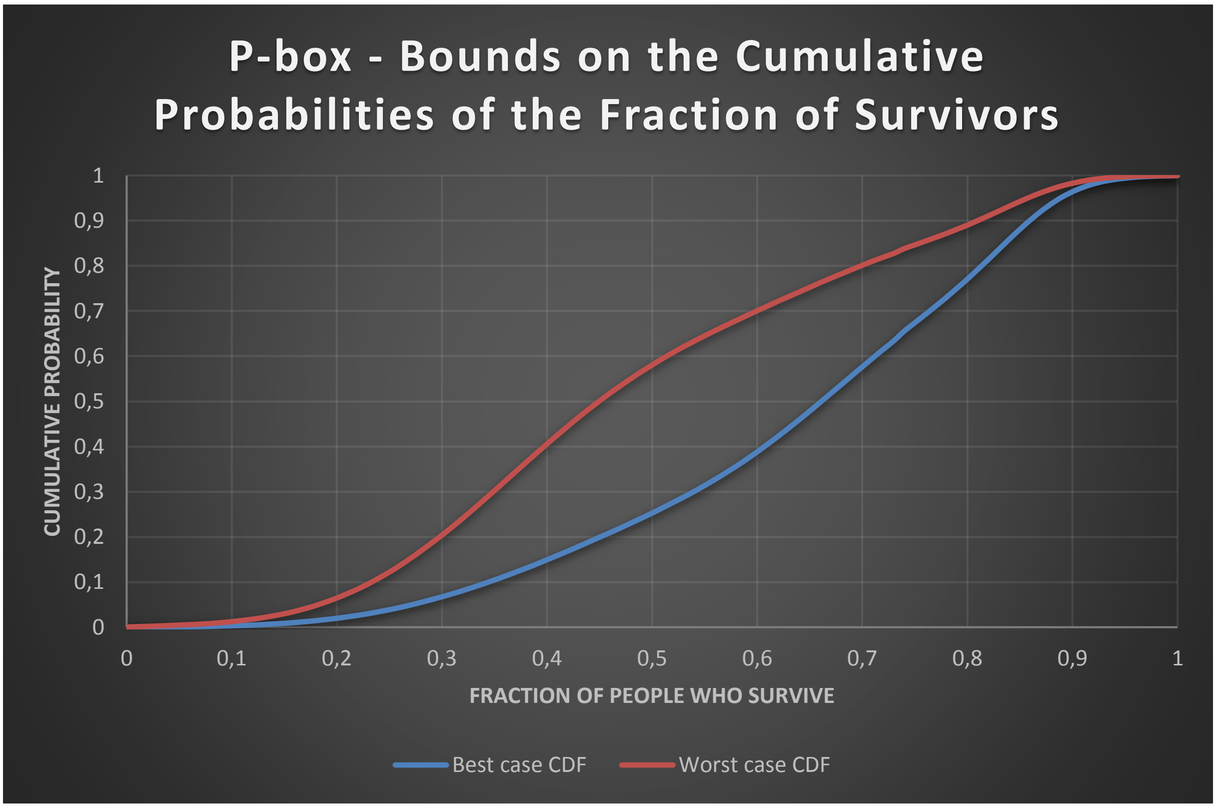


Figure 11: Bounds on the cumulative probabilities of the fraction of survivors.

If we assume that the probability intervals in figure 8 contain the “true” underlying probability of occurrence, and that the probability distribution of survivors at different wave heights in figure 9 is a correct representation of the “true” underlying probability of surviving at different wave heights, the p-box in figure 11 must contain the true cumulative probability of the fraction of survivors. However, we are unable to say anything regarding its shape, beyond that it must be within the bounds of the p-box.

Step 4: Assumption justification. The relevant assumptions, their justifications and the strength of the assumption justifications are presented in table 8 and 9 on the following pages. The judgements of these assumptions have followed the principles and guidelines presented by Flage & Aven (2009).

Table 8: An overview of the relevant assumptions, their associated justifications and judgement of the strength of these justifications. Part 1 of 2.

Assumption	Assumption justification	Strength of assumption justification
The bounds that are set on the probability of occurrence for the different wave height intervals contain the true probability of occurrence.	These bounds represent the upper and lower probability of occurrence found after going through several models made by experts. The experts claim that it is difficult to accurately predict the probability of occurrence for the different wave heights, since it is hard to say exactly how much of the in place landmasses that will fall out and cause a landslide.	Poor
The probability distribution of the fraction of survivors at the different wave height intervals is an accurate representation of the true underlying frequentist probability.	This is based on 70 years of continuously updated statistical data along with repeatedly proven models of how waves of different sizes behave after they hit land.	Strong
There will be a landslide sometime in the next 1000 years.	This is based on a combination of historical recordings of similar events, estimated annual precipitation and the surrounding geology. Essentially it is a subjective estimate made by a team of experts.	Moderate
All wave heights within the same interval will cause a similar degree of destruction.	It is clear that waves on the opposite side of each interval will *not* cause a similar degree of destruction. As previously mentioned, this assumption is only taken to simplify the example.	Poor

Table 9: An overview of the relevant assumptions, their associated justifications and judgement of the strength of these justifications. Part 2 of 2.

Assumption	Assumption justification	Strength of assumption justification
The maximum height a wave can reach in this situation is 30 meters.	This is based on a wave formation model of a worst case scenario where all of the in place landmasses fall out unhindered and crash into the ocean below. The experts believe they have an accurate estimate of the total volume and mass of the "available" landmasses, and the wave formation model has previously been proven to be accurate.	Strong
The cost of upgrading the safety systems will not surpass 800 million NOK.	Several contractors have been contacted regarding the job at hand, and all have given rice estimates that range between 700 and 800 million NOK, but costly unforeseen events can always occur, and especially relevant in projects of this magnitude.	Moderate
The estimated interval of the fraction of survivors if that the safety systems are upgraded contains the true fraction of survivors.	This is based on the location of the new shelters, and how quickly it is presumed that the people nearby can reach and enter them. The locations of the new shelters are chosen so that all of the residents (100%) should be able to reach and enter them from the time the warning alarm goes off and until the tsunami hits land.	Strong

Step 5: Analyzing the result of the p-boxes in light of the assumption justifications. As we can see from table 8, one of the most important assumptions: "The bounds that are set on the probability of occurrence for the different wave height intervals contain the true probability of occurrence" is poorly justified. The reason for its importance is that this assumption, along with: "The probability distribution of the fraction of survivors at the different wave height intervals is an accurate representation of the true underlying frequentist probability" lays the foundation for the bounds on the cumulative probabilities of the fraction of survivors, which in turn is the final quantitative result of the risk analysis. If the strength of this assumption justification truly is poor, the bounds on the cumulative probabilities of the fraction of survivors seen in figure 11 cannot be seen as reliable and must be extended to the bounds seen in figure 12, here represented by the purple and turquoise lines. Consequently, it is important to analyze why the strength of the assumption justification is judged as poor. If we follow the principles and guidelines suggested by Flage and Aven (2009) for assessing strength of

assumption justifications, the strength of an assumption justification can be considered poor if one or more of the following conditions are met:

- The assumptions made represent strong simplifications.
- Data/information is non-existent or highly unreliable /irrelevant.
- There is strong disagreement among experts.
- The phenomena involved are poorly understood, models are non-existent or known/believed to give poor predictions.

To find the bounds on the different probabilities of occurrence several expert made prediction models were analysed, and the highest and lowest probability of occurrence found for each of the wave height intervals were selected as the upper and lower bounds on probability. The assumptions that form the background of these models does not represent strong simplifications, nor are they seen as very reasonable due to the lack of relevant information. *The assumptions that form the background of these models represent some simplifications.* In this case it is not correct to say that data is non-existent or highly unreliable, nor is it correct to say that there exist large amounts of reliable and relevant data. There is some reliable data from local geological surveys along with data from similar situations. Again, we fall somewhere between the two, and a more correct statement would be that *some reliable and relevant data is available.* It is however true that there is strong disagreement among experts, and this is clearly expressed through the wide variety of probabilities predicted by the models. But this disagreement is in many ways reflected through the bounds of the intervals, which incorporate the variety in the probabilities, and this is consequently expressed by the width of the best- and worst-case CDF. It would therefore be more correct to say that *there is strong disagreement among experts, but this disagreement is reflected through the width of the probability intervals.* The phenomena involved in this case are neither poorly nor well understood, but understood to a certain degree, and consequently the existing models give various predictions. A more correct final statement would in this case be that *the phenomena involved are understood to a certain degree, and the existing models give varying predictions within a certain range which is believed to contain the true value.*

Since one of the conditions proposed by Flage and Aven (2009) is met, namely that there is strong disagreement among experts, the strength of the assumption justification is considered poor. It has previously been argued that this disagreement is reflected through the width of the probability intervals, and that all the other statements fall somewhere between strong and poor. From this simple analysis it is clear that even though the strength of the assumption justification has to be considered as poor since it meets one of the required conditions, one can still assume that the probability intervals contain the “true” underlying probability. Consequently, one can assume that the bounds on the cumulative probabilities of the fraction of survivors does contain the “true” CDF.

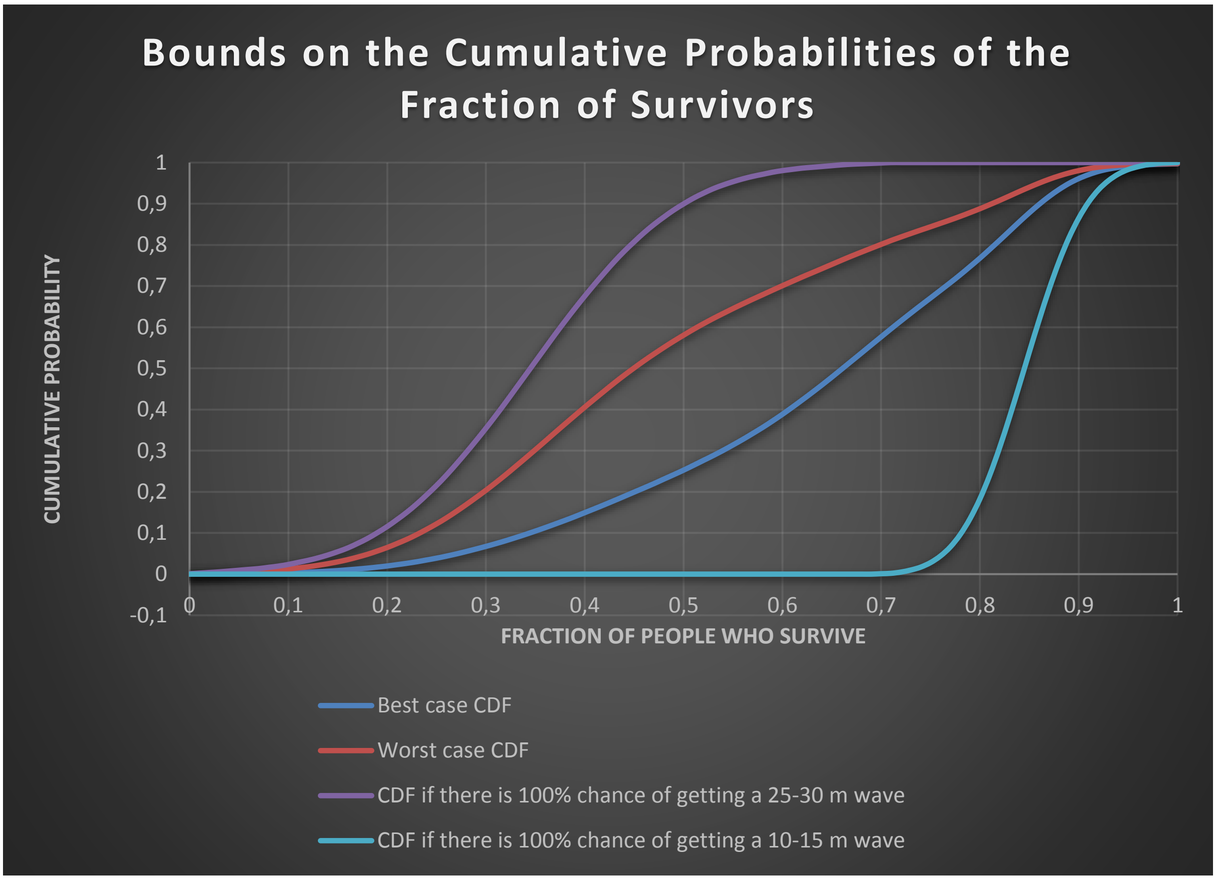


Figure 12: Bounds on the cumulative probabilities of the fraction of survivors.

From this point on it is assumed that the bounds on the cumulative probabilities of the fraction of survivors, seen in figure 11 does contain the "true" CDF. A quick numerical overview of this figure can be seen in table 10. We can see that both the best and worst case probabilities are similar for a fraction of survivors >0.9 and <0.2, but there is quite a big gap in between these values. For instance, in the best case scenario there is a 74.4 % that more than 50% of residents survive the tsunami, while in the worst case scenario there is only a 42% chance of having the same number of survivors. Again, it is important to note that it is impossible to say anything regarding which of the two scenarios is closest to the "true" value. The only thing that is assumed to be true is that the "true" value is found somewhere between the best and worst case scenarios.

Table 10: Numerical overview of the bounds on survival probabilities.

Minimum fraction of people who survive	Best case probability	Worst case probability
1.00	0.3 %	0.2 %
0.95	0.8 %	0.5 %
0.90	3.9 %	2.0 %
0.80	23.2 %	11.2 %
0.70	42.4 %	19.9 %
0.60	61.2 %	30.0 %
0.50	74.4 %	42.0 %
0.40	85.1 %	59.4 %
0.30	93.3 %	79.6 %
0.20	98.0 %	93.5 %
0.10	99.7 %	98.8 %
0.05	99.9 %	99.5 %

The best and worst case CDF can also be used to calculate a best and worst case estimate of the fraction of survivors in the long run. This only a hypothetical number that would be true if this situation could be repeated an infinite number of times under similar conditions. In the long run, the fraction of survivors is 0.63 in the best case scenario, and 0.49 in the worst case scenario. If we were to use these numbers in a cost-effectiveness analysis we would find that the implied cost of averting a fatality (ICAF) would lay somewhere between 1.5 million NOK and 2.8 million NOK. These numbers are way below the emerging consensus in the academic literature, which finds the statistical value of human life to range somewhere between 32 million NOK and 80 million NOK (Kniesner et al., 2012; & Robinson & Hammitt, 2015). It is still important to note that the implied cost of averting a fatality does not sufficiently reflect uncertainty. The exact calculations behind this can be found in appendix A.

If more detailed information regarding the probability distribution of the timeframe of the landslide, and the estimated population growth was available, it would be easier to give definitive recommendation about whether the state should invest in upgrading the safety systems or not. This upgrade has some clear advantages, where the main ones are that the probability of having at least 90% of the residents survive the tsunami is lifted from [2.0 %, 3.9 %] to 100%, and the worst and best case long run fraction of survivors are lifted from 0.49 and 0.63 to 0.90 and 1.0. See figure 13 for a comparison between the current bounds on the cumulative probabilities of the fraction of survivors and the bounds of the CDF if the safety systems are upgraded.

Combining probability bounds analysis with a qualitative judgement of the strength of the assumptions that lay the foundation for the numerical values used in the PBA, does in this case result in a detailed and informative risk picture in the event of a landslide. From the bounds on the cumulative probabilities of the fraction of survivors, and the expected number of survivors in the long run, it is clear that the residents living in the end of the fjord are at great risk of dying *if* there is a landslide. Currently the geologists cannot say anything regarding the probability distribution of this timeframe, only that they believe there will be a landslide sometime in the next 1000 years. Another factor that should be taken into account before the state choses whether to invests in new safety systems or not is how the population of this small community is going to change in the coming years. If the population is thought to grow a lot in the coming years the safety systems may need to be upgraded regardless of the timeframe of the landslide situation just to be able to accommodate everyone inside the shelters. An upgraded safety system may also have a positive impact on the population growth, which in turn can turn can contribute to economic growth in the community. These aspects are beyond the grasp of this risk assessment and should be a part of the managerial review and judgement, before a final decision is made.

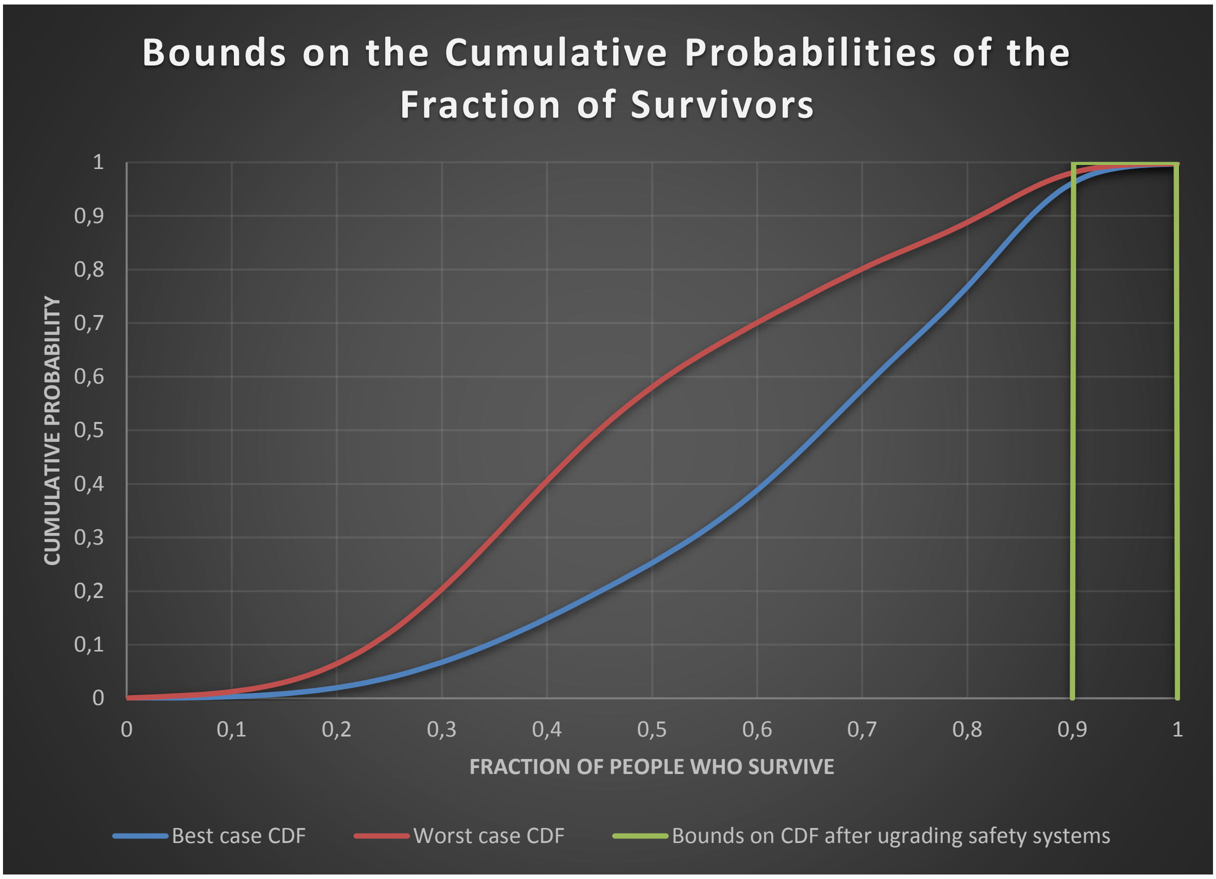


Figure 13: Bounds on the cumulative probabilities of the fraction of survivors.

5. Discussion

This chapter will introduce two separate discussion topics as well as suggestions for further work. The aim of the first discussion is to discuss why it is important to acknowledge an additional level of uncertainty and to what degree the new alternative uncertainty taxonomy eliminates the challenges that are present in the other uncertainty taxonomies. The aim of the second discussion is to discuss the previously introduced risk assessment method's applicability as a tool in risk management.

5.1. The importance of acknowledging the existence of an additional uncertainty level and how the introduction of this level solves challenges that are present in the other uncertainty taxonomies

As previously mentioned, uncertainty can be interpreted in two ways in a risk setting, either as aleatory or as epistemic uncertainty. Aleatory uncertainty represents the natural variation of a system and cannot be further reduced by acquiring more knowledge. If a person has the required background knowledge on a certain system, this type of uncertainty can be expressed by using a probability model that is an accurate representation of the underlying frequentist distribution. E.g. if you are to draw one card out of fair and properly shuffled deck of cards the probability of drawing hearts will always be 1/13. Epistemic uncertainty on the other hand stems from a lack of knowledge and can be expressed as a subjectively assigned probability which is based on some background knowledge (See chapter 2.1.1. for an explanation of how this probability can be interpreted). E.g. I assign the probability of me beating my best friend in chess to be 0.7. This is based on the knowledge I have gathered by knowing him the last 20 years. Subjective probabilities like this does not have to be expressed as point estimates. If the assignor has a moderate or poor understanding of the situation at hand he/she can opt to express his/her uncertainty as an imprecise probability, or a probability interval. This means that the assignor believes that the "true" probability will lay somewhere within the interval. If the assignor has the supporting background knowledge he/she may also attach a probability distribution to this interval. In those cases where no probability distribution is attached to interval, the assignor only expresses the he/she believes that the true probability lies somewhere within the interval, and nothing more. It is important to note that this is not the same as a normal distribution, where each value within the interval is equally likely. E.g. I assign the probability of me beating a random stranger in chess to be [0.3, 0.7]. This means that based on my knowledge and experience as a chess player I believe that the true probability of me beating a random stranger in chess lies somewhere within that interval, but I cannot say anything regarding the probability distribution of the values within the interval.

The uncertainty taxonomy introduced by Courtney (2001) totaly ignores this aspect of uncertainty. Level 1 & level 2 is represented by aleatory uncertainty that can for all practical purposes be said to be accurately described by frequents probabilities, while no probabilities can be used to describe uncertainty in level 3 and 4. This way of dividing different degrees of uncertainty would lead to many situations not fitting in anywhere in this taxonomy. There is a clear gap here that needs to be bridged. Aven (2013) did that to a certain degree with his alternative taxonomy, see table 3. Here he divides uncertainty into three distinct levels: low, moderate and deep uncertainty. There is no clear distinction between these levels, other than how strong one can judge the background knowledge to be, and if black swans can be present. The use of imprecise probabilities to express uncertainty could be used in situations that fall under moderate uncertainty where the background knowledge is based on some dominating explanations and beliefs. But what this alternative taxonomy gains in its simplicity, it lacks in detail and concrete information. A new taxonomy has been suggested in this thesis that incorporates the knowledge aspect that Aven (2013) found to be missing in Courtney's taxonomy into the taxonomy template that Courtney (2001) created and adds a fifth level of uncertainty that gives room for expressing epistemic uncertainty as imprecise probabilities. This way the level of detail that was introduced in Courtney's (2001) taxonomy is slightly increased, and there is made room for situations where the uncertainty can only be expressed as imprecise probabilities. Having a taxonomy that captures every nuance of uncertainty can aid risk analysts in finding the risk assessment and risk management method(s) that are best suited for handling a given situation.

There is however one minor shortcoming with this taxonomy which it also shares with the alternative proposed by Aven (2013). Neither of these two taxonomies give a clear and distinct separation between moderate and deep uncertainty. In Aven's taxonomy, moderate and deep uncertainties are separated by a judgement of the strength of knowledge. There are some dominating explanations and beliefs in the former category, while the background knowledge is judged to be poor in the latter. This leaves the exact separation of moderate and deep uncertainty open for debate. In the taxonomy proposed in this thesis the separation of moderate and deep uncertainty falls somewhere in the third uncertainty level. According to the deep uncertainty definition introduced by Walker et al. (2017) it could be argued that once imprecise probabilities are used to describe the uncertainties of a situation that situation will fall under the deep uncertainty category. By this account almost every situation where probabilistic point estimates are used to describe uncertainties would also fall under the deep uncertainty category, since point estimates can be seen as short intervals due to the natural imprecision of decimal numbers (Aven & Reniers, 2013). E.g. 0.7 = [0.650, 0.749].

Maybe the fact that it is hard to make a clear separation between moderate and deep uncertainty suggests that this is a natural grey area, and that the necessity for clear separation may not be that high? If, however it is deemed necessary to have a clear distinction between the two it could be done by introducing an additional level of uncertainty, namely high uncertainty. High uncertainty would cover every situation where imprecise probabilities are used to express uncertainties (Level 3), while deep uncertainty would cover every situation where it is impossible to use probabilities to express uncertainty (Level 4 & 5). Further work is needed to investigate this.

5.2. The newly proposed methods applicability as a tool for assessing risk

The three main objectives of a risk assessment are to identify risks, analyze the identified risks, and to evaluate these risks (ISO, 2018). This is done so that a clear and informative risk picture can be presented to the decisionmakers, thus aiding them in making a well informed decision. Once we enter the realm of level 3 uncertainty where it is no longer possible to use point estimates of probability to express uncertainty, and the only viable numerical option is to use imprecise probabilities, it is impossible to present a quantitative risk picture without using bounds on the probabilities. This means that in a situation like this it is impossible for the risk analyst to identify one clear numerical probability for each outcome. What can be done however, is to identify the best and worst case scenarios by using probability bounds analysis and use this as the quantitative basis for the risk assessment. A probability bounds analysis alone is not enough to present a clear and informative risk picture, because it does not involve any justification of the relevant assumptions. If we are to use a quantitate risk analysis method as the basis for evaluating risk in a situation where the underlying information is so poor that it is impossible to identify point estimates of the probabilities and imprecise probabilities are the only available option, it is of outmost importance that the bounds on these probabilities can be justified. If there is no form of assessment of the strength of the assumptions that lay the supporting foundation of a probability bounds analysis it would be very hard to say whether the bounds it presents in its final result truly represents the best and worst case scenarios of a given situation or not. The risk assessment method presented in this thesis is therefore a combination of a probability bounds analysis and a justification of the assumptions. In the latter all the relevant assumptions must be justified and followed up by a judgement of the strength of theses justifications. By doing so it becomes clear to the decisionmakers whether they can trust the bounds that are presented as the final result of the probability bounds analysis or not.

If the assumptions that lay the foundations for the bounds on the probabilities can be justified this method has some clear benefits as a tool for assessing risk. First of all, it lets the risk analysts make probabilistic risk assessments when the available empirical information regarding the input distributions is limited. Second, it can help narrowing down the bounds on the cumulative probabilities, as illustrated in the example in chapter 4.2.2., see figure 12. Third, it enables the risk analyst to combine both aleatory and epistemic uncertainty in a meaning full way. And finally, it allows the risk analyst to present a clear and informative risk picture to the decisionmakers by giving a quantitative expression of the risks, and a qualitative judgement of the assumptions. This is not to say that this risk assessment method is without limitations. The first, and most obvious one is its inability to produce reliable results *if* the assumptions that lay the foundation for the probability bounds cannot be sufficiently justified. The second one is that if the probability intervals that are used to identify the p-box are very wide, the resulting bounds on the cumulative probabilities will also be very wide, meaning that there is big difference in the best and worst case scenario. If this is the case, then using a probability bounds analysis will not help very much in narrowing down the bounds on the cumulative probability distribution, meaning that we barely gain new insight by doing the analysis. Say for example that the probability of occurrence for each of the wave height intervals in the previous example could be described by the same probability interval [0.05, 0.85]. Due to the width of this interval the best and worst case cumulative probabilities would be much closer to the purple and turquoise lines (see figure 12), which means that we can just barely narrow down the best and worst case scenarios, which in turn means that we just barely learn anything new from this analysis at all. The final limitation is how the PBA handles the introduction of a timeframe without a probability distribution. This limitation is clearly demonstrated in the case study in chapter 4.2.2., where it is impossible to say anything more about the timeframe of the incident other than that the experts believe it will occur sometime in the next 1000 years. This cannot be included in the PBA itself, which only gives a probabilistic overview of the best and worst case scenario if there is a landslide. This means that the results of the PBA have to be analyzed in the light of this information.

It is clear that this method has its limitations, but under the right conditions it has proved to be a great tool for assessing risk under moderate/deep uncertainty.

5.3. Further work

As previously mentioned some further work is needed to investigate the necessity of a clear separation between moderate and deep uncertainty, or if it is okay that this is a grey area. The combination of the probability bounds analysis and the qualitative assessment of the knowledge/justifications should also be tested on a real life case. This may uncover more strengths and/or weaknesses than what was found in this thesis.

6. Conclusion

This thesis highlights and discusses some of the challenges and limitations imbedded in the existing deep uncertainty taxonomies and develops an alternative taxonomy with a higher level of detail and an additional level of uncertainty to try and combat these challenges. Later on, a method that can be used to assess risk in situations that fall under the new uncertainty level is introduced and its applicability is illustrated through a stylized example.

The current deep uncertainty taxonomies were found either to be too limited in its reflections of the various nuances of uncertainty in the stages between low and deep uncertainty, and to be without a judgement of knowledge, or to have a level of detail that was too low to be precise and informative. To solve this, a new taxonomy has been introduced which expands the template of the former taxonomy to incorporate the knowledge dimension which was introduced in the latter and adds a fifth level of uncertainty so that every nuance of uncertainty is clearly reflected. This additional level of uncertainty incorporates the alteration from moderate to deep uncertainty, and it encompasses every situation where the only means of quantitatively expressing the uncertainty is by using imprecise probabilities. Simply put, the new taxonomy combines the best features from the former taxonomies with some new elements, which in turn gives it a more nuanced reflection of the various levels of uncertainty and a higher level of detail. By adopting this taxonomy, risk analysts may get a better understanding of the various levels of uncertainty, which could in turn aid them in the process selecting the most appropriate method to assess risk at hand in a given situation.

It can be challenging to do a comprehensive probabilistic risk assessment in a situation where the only means of quantitatively expressing the uncertainty is by using imprecise probabilities. Regular risk assessments methods are not capable of handling such situations properly, which is why a probability bounds analysis may be the better choice. Through the case study, the PBA was found to be a great tool in identifying and expressing risk quantitatively in situations where the probability intervals were not to wide. However, this analysis method alone is not capable of describing every aspect of the risk picture, as it lacks a qualitative judgement of the SoK. Without this type of judgment there is no way of knowing if the probability bounds that lay the foundation for the final results of this analysis can be trusted. To resolve this, every relevant assumption that is made during the analysis has to be justified and followed up with a judgement of the strength of these justifications, and finally the results of the PBA has to be reviewed in the light of these justifications. This combination has through the case study proven to be a valuable tool in assessing risk in a situation were the only quantitative expression of uncertainty is imprecise probabilities. It enables the risk analyst to present a clear and informative risk picture in situations where the empirical information is limited.

References

Aven, T. (2008). A semi-quantitative approach to risk analysis as an alternative to QRAs. *Reliability Engineering & System Safety. Vol. 93, pp. 768-775.*

Aven, T. (2010). On the Need for Restricting the Probabilistic Analysis in Risk Assessment to Variability. *Risk Analysis, Vol. 30, Issue 3, pp. 354-360.*

Aven, T. (2013). On How to Deal with Deep Uncertainty in a Risk Assessment and Management Context. *Risk Analysis, Vol 33, Issue 12, pp. 2082-2091*

Aven, T. (2014). *Risk, Surprises and black swans: fundamental ideas and concepts in risk assessment and risk management.* United Kingdom: Taylor & Francis

Aven, T. (2015) *Risk Analysis.* West Sussex, United Kingdom: John Wiley & Sons.

Aven, T. (2016) Risk Assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research. Vol. 253, pp. 1-13.*

Aven, T. & Reniers, G. (2013) How to define and interpret probability in a risk and safety setting. *Safety Science. Vol. 51, pp. 223-231.*

Aven, T., Vinnem, J.E. & Wiencke, H.S. (2007). A decision framework for management, with application to the offshore oil and gas industry. *Reliability Engineering and System Safety. Vol. 9, pp. 433-448*

Aven, T., Zio E., Baraldi P. & Flage, R. (2013). *Uncertainty in risk assessment: the representation and treatment of uncertainties by probabilistic and non-probabilistic methods.* West Sussex, United Kingdom: John Wiley & Sons.

Beer, M., Ferson, S. & Kreinovich, V. (2013) Imprecise probabilities in engineering analyses. *Mechanical Systems and Signal Processing. Vol, 37, pp. 4-29.*

Ben-Haim, Y. (2000). Robust rationality and decisions under severe uncertainty. *Journal of the Franklin Institute. Vol. 337, Issue 2, pp. 171-199.*

Bjerga, T. (2016). *Finn tittel* (Doktorgradsavhandling). Universitetet i Stavanger, Stavanger.

Bjerga, T. & Aven, T. (2015). Adaptive risk management using new risk perspectives – an example from the oil and gas industry. *Reliability Engineering and System Safety. Vol. 134, pp. 75 – 82.*

Courtney, H. (2001). *20/20 Foresight: Crafting Strategy in an Uncertain World.* Boston, Massachusetts: Harvard Business School Press

Courtney, H. (2003). Decision-driven scenarios for assessing four levels of uncertainty. *Strategy & Leadership, Vol 31 Issue: 1, pp. 14-22.* https://doi.org/10.1108/10878570310455015

Chao, P.T., Hobbs, B.F. & Venkatesh, B.N. (1999). How climate uncertainty should be included great lakes management: modelling workshop results. *J Am Water Resour Assoc. Vol. 199, pp. 1485-1497*

Clark, M.P. & Pulwarty, R.S. (2003). Devising resilient responses to potential climate change impacts. *Ogmius: News Cent Sci Technol Policy Res Vol. 5, pp. 2-3.*

Cox, L.A.(T.). (2012). Confronting deep uncertainties in risk analysis. *Risk Analysis, Vol. 32, Issue 10, pp. 1607-1629*

Dubois, D., Prade, H.M., Farreny, H. Martin-Clouaire, R. & Testmale, C. (1988). *Possibility theory: an approach to computerized processing of uncertainty, 2.* New York: Plenum Press.

EPA (U.S. Environmental Protection Agency). 2001. *Risk Assessment Guidance for Superfund (RAGS), Vol. 3, Part A: Process for Conducting Probabilistic Risk Assessment.*

Ferson, S. & Ginzburg, L.R. (1996). Different methods are needed to propagate ignorance and variability. *Reliability Engineering and System Safety. Vol. 54, pp. 133-144.*

Ferson, S., Ginzburg, L.R. & Akcakaya, R. (To appear). Whereof one cannot speak: when input distributions are unknown. *Risk analysis*

Flage, R. Aven, T. & Berner, C.L. (2018). A comparison between a probability bounds analysis and a subjective probability approach to express epistemic uncertainties in a risk assessment context – A simple illustrative example. *Reliability Engineering and System Safety, Vol. 169, Issue 169, pp. 1-10.*

Flage, R. & Aven, T. (2009) Expressing and communicating uncertainty in relation to quantitative risk analysis. *Reliability: Theory & Applications, Vol. 2, Issue 13, pp. 9 – 18.*

Grafarend, Erik W. (2006). *Linear and Nonlinear Models: Fixed Effects, Random Effects, and Mixed Models*. Walter de Gruyter.

Groves, D.G., Lempert, R.J. (2007). A new analytic method for finding policy-relevant scenarios. *Global Environmental Change. Vol. 17, pp. 73-85*

Haasnoot, M., Kwakkel, J.H, Walker, W.E & Ter Maat, J. (2013) Dynamic adaptive policy pathways: a new method for crafting robust decisions for a deeply uncertain world. *Global Environmental Change. Vol. 23, Issue 2, pp. 485-498.*

Hall, J.W., Lempert, R.J., Keller, K., Hackerbarth, A., Mijere, C. & McInerney, D.J. (2012). Robust climate policies under uncertainty: a comparison of robust decision making and info-gap methods. *Risk Analysis. Vol. 32, Issue 10, pp. 1657-1672.*

Hamarat, C., Kwakkel, J.H. & Pruyt, E. (2013). Adaptive Robust Design under deep uncertainty. *Technological Forecasting & Social Change, Vol. 80, Issue 4, pp. 408-418*

Holling, C.S. (1978). *Adaptive Environmental assessment and Management.* Chichester: John Wiley & Sons.

International Standardization Organization (ISO). (2018). *Risk Management - principles and guidelines, ISO 31000:2018*

Kaplan. S. & Garrick, B.J. (1981). On the quantitative definition of risk. *Risk analysis. Vol. 1, pp. 11-27.*

Kasperson, R.E. (2008). Coping with deep uncertainty: Challenges for environmental assessment and decision making. In Bammer, G. and Smithson, M. (Editors) (2008) *Uncertainty and Risk: Multidisciplinary Perspectives*. London: Earthscan, pp337-347.

Knieser, T.J., Viscusi, W.K. & Woock, C. (2012). The value of a statistical life: Evidence from panel data. *Review of Economics and Statistics, Vol. 94, Issue 1, pp. 74-87.*

Kwakkel, J.H., Haasnoot, M. & Walker, W.E. (2016). Comparing Robust Decision-Making and Dynamic Adaptive Policy Pathways for model-based decision support under deep uncertainty. *Environmental Modelling & Software. Vol. 86, pp. 168-183.*

Lempert, R.J., Groves, D.G., Popper, S.W. & Banks, S.C. (2006). A general, analytic method for generating robust strategies and narrative scenarios. *Management Science. Vol. 52, Issue 4, pp. 514-528.*

Lempert, R.J., Popper, S.W., Banks, S.C. (2004). *Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis.* Santa Monica, CA: RAND.

Løvås, G.G. (2004). *Statistikk for universiteter og høgskoler*. Oslo: Universitetsforlaget AS.

Macmillan Dictionary. (2018). *Definition of assumption.* From: https://www.macmillandictionary.com/dictionary/british/assumption Accessed: 5/03-2018

Maier, H.R., Guillaume, J.H.A.G., van Delden, H., Riddell, G.A., Haasnoot, M. & Kwakkel, J.H. (2016). An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together?. *Environmental Modelling & Software. Vol. 81, pp. 154-164*

Malcolm, D. G., Roseboom, J. H., Clark, C. E., & Fazar, W. (1959). Application of a technique for research and development program evaluation. Operations Research, 7, 646–669.

Makridakis, S., Hogarth, R.M. & Gaba, A. (2009). Forecasting and uncertainty in the economic and business world. *International Journal of Forecasting. Vol. 25, Issue 4, pp. 794-812.*

Mastrandrea, M.D., Field, C.D., Stocker, T.F., Edenhofer, O., Ebi, K.L. & Frame, D.J. (2010). Guidance note for lead authors of the IPCC fifth assessment report on consistent treatment of uncertainties. *Intergovernmental Panel on Climate Change (IPCC).*

Quade, E.S. (1989). *Analysis for Public Decisions,* Third Edition. New York: Elsevier Science Publishing Co., Inc.

Robinson, L.A & Hammit, J.K. (2015). Valuing reductions in fatal illness risks: Implications of recent research. *Health Economics, Vol. 25, Issue, 8, pp. 1039-1052.*

Shortridge, J., Aven, T. & Guikema, S. (2016). Risk assessment under deep uncertainty: A methodological comparison. *Reliability Engineering and Safety Systems. Vol. 159, pp. 12-23.*

SRA. (2015). Glossary society for risk analysis. From: http://www.sra.org/sites/default/files/pdf/SRA-glossary-approved22june2015-x.pdf Accessed: 10/01-2018

Taleb, N. N. (2010) *The black swan: The impact of the highly improbable (2nd ed.)* New York: Random House Trade Paperbacks

Tucker, W. T. & and Ferson, S. (2003). *Probability bounds analysis in environmental risk assessments*. New York: Applied Biomathematics.

Walker, W.E., Lempert, R.J. & Kwakkel, J.H. (2017). *Deep Uncertainty*. From: http://www.hau.gr/resources/toolip/doc/2012/05/10/deep-uncertainty_warren-e-walker.pdf Accessed: 05/03-2018

Walker, W.E., Marchau, V. A. W. J. & Swanson, D. (2010). Addressing deep uncertainty using adaptive policies: Introduction to section 2. *Technological Forecasting & Social Change. Vol.* 77, *pp. 917-923*

Walley, P. (1991). *Statistical reasoning with imprecise probabilities.* London: Chapman and Hall.

Williamson, R.C & Downs, T. (1990). Probabilistic Arithmetic. I. Numerical Methods for Calculating Convolutions and Dependency Bounds. *International Journal of Approximate Reasoning. Vol. 4, pp. 89-158.*

Winkler, R.L. (1996). Uncertainty in probabilistic risk assessment. *Reliability Engineering & System Safety. Vol. 54, Issue, 2, pp. 127-132.*

Appendix A

A.1. The calculations behind the bounds on the cumulative probabilities of the fraction of survivors

All of the calculations in this chapter follow the mathematics for calculating probability bounds as described by Tucker and Ferson (2003).

First, we use the σ_i and μ_i from table 6 as inputs for the normal distributions of the fraction of survivors at the different wave height intervals. Then we use these normal distributions to identify the fraction of times 1%, 2%, 3%, …, 100% of the residents survives for each of the different wave height intervals. To better illustrate this a sample of the results can be seen in table A1. This table is basically a numerical representation of the graphs seen in figure 9. The reason for the "-" in the fraction of times people survive the smallest wave height is that it is assumed that at least 70% of the residents survive if such a wave strikes.

Table A1: A numerical representation of the graphs seen in figure 9

	Percentage of residents who survive								
	45 %	46 %	47 %	48 %	49 %	50 %	51 %	52 %	53 %
Wave height (m)	The fraction of times the percentage of residents seen above survive the different waves								
10 to 14.99	-	-	-	-	-	-	-	-	-
15 to 19.99	0,002	0,002	0,003	0,004	0,004	0,005	0,007	0,008	0,009
20 to 24.99	0,024	0,024	0,024	0,025	0,025	0,025	0,025	0,025	0,024
25 to 30	0,023	0,022	0,02	0,018	0,017	0,015	0,014	0,012	0,011

These fractions, which are written in blue are then multiplied by the respective probability bounds on the best and worst case scenarios, and their values are summed up find the bounds on the cumulative probabilities of the fraction of survivors. See table 7 for the probability bounds of the best and worst case scenarios, and table A2 and A3 for a sample of the calculated bounds on the cumulative probabilities of the fraction of survivors.

Table A2: Bounds on the best case cumulative probabilities – A partial numerical representation of the best case CDF in figure 11.

	Percentage of residents who survive									
	≤ 44%	45 %	46 %	47 %	48 %	49 %	50 %	51 %	52 %	53 %
Wave height (m)	The fraction of times the percentage of residents seen above survive the different waves									
10 to 14.99	-	-	-	-	-	-	-	-	-	-
15 to 19.99	-	0,001	0,001	0,001	0,002	0,002	0,002	0,003	0,003	0,004
20 to 24.99	-	0,007	0,007	0,007	0,008	0,008	0,008	0,008	0,008	0,007
25 to 30	-	0,002	0,002	0,002	0,002	0,002	0,002	0,001	0,001	0,001
	The fraction of times the percentage of residents seen above survive									
Sum	-	0,010	0,010	0,010	0,011	0,011	0,011	0,012	0,012	0,012
Cumulative P	0,189	0,199	0,210	0,220	0,231	0,242	0,253	0,264	0,276	0,288

Table A3: Bounds on the worst cumulative probabilities – A partial numerical representation of the worst case CDF in figure 11.

	Percentage of residents who survive									
	≤ 44%	45 %	46 %	47 %	48 %	49 %	50 %	51 %	52 %	53 %
Wave height (m)	The fraction of times the percentage of residents seen above survive the different waves									
10 to 14.99	-	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
15 to 19.99	-	0,000	0,000	0,000	0,001	0,001	0,001	0,001	0,001	0,001
20 to 24.99	-	0,006	0,006	0,006	0,006	0,006	0,006	0,006	0,006	0,006
25 to 30	-	0,012	0,011	0,010	0,009	0,009	0,008	0,007	0,006	0,006
	The fraction of times the percentage of residents seen above survive									
Sum	-	0,018	0,017	0,016	0,016	0,015	0,015	0,014	0,013	0,013
Cumulative P	0,483	0,501	0,518	0,535	0,551	0,566	0,581	0,595	0,608	0,621

A.2. The calculations behind the best and worst case estimate of the fraction of survivors in the long run

The mathematics in this chapter are based upon the rules of statistics as presented by Løvås (2004).

To find the long run estimates of the number of survivors we simply multiply the fraction of times the different percentages of residents survive by that percentage and sum up the results, see the equation below.

$$\sum_{i=1}^{100} P(x_i)x_i$$

$$x_i$$

x_i = The fraction of residents who survive (Purple numbers in table A2 and A3)

$P(x_i)$ = The fraction of times x_i occurs / The probability of having x_i survivors (Red numbers in table A2 and A3)

To find the implied cost of averting a fatality you take the total cost of the operation and divide it by the expected number of lives it would save:

In the best case the scenario, the implied cost of averting a fatality would range somewhere between:

$$\frac{750\ million\ NOK}{[270,\ \ 370]\ lives\ saved} = [2.0, 2.8]$$

In the worst case scenario, the implied cost of averting a fatality would range somewhere between:

$$\frac{750\ million\ NOK}{[410, 510]} = [1.5, 1.8]$$