Faculty of Science and Technology

MASTER'S THESIS

Study program/ Specialization: Industrial Asset Management	Spring semester, 2019 Confidential
Writer: Amir Beiky	………………………………………… (Writer's signature)
Faculty supervisor: Idriss El-Thalji, PhD External supervisor: Arnaud Barre, MBA, MSc Eng, WoodGroup	
Thesis title: DIGITAL COLLECTION AND EVALUATION OF EXPERT OPINION IN VIRTUAL ENVIRONMENT	
Credits (ECTS): 30	
Key words: Expert judgment, Elicitation methods, Reliability assessment, Roulette method, Classical model, Digitalization, Excalibur, MATCH, Performance based weighting, Seed variables, target variables	Pages: …….34…… + enclosure: …8… Stavanger 06.06.2019

DIGITAL COLLECTION AND EVALUATION OF EXPERT OPINION IN VIRTUAL ENVIRONMENT

By

AMIR BEIKY

Thesis is submitted to the Faculty of Science and Technology

University of Stavanger

In Fulfillment of the Requirements for the degree of

Master of Science

(MSc)

Specialization: Industrial Asset Management

University of Stavanger

FACULTY OF SCIENCE AND TECHNOLOGY

University of Stavanger

Year 2019

Abstract

Using expert judgment as a source of information in engineering field is essential especially when there is few or no empirical or historical data available. Elicitation is a process to obtain expert information and use it as a dataset. There are different methods and techniques for eliciting the opinion of experts. The Expert judgment elicitation has been used for many years in oil and gas industry. However, expert judgment in a formal way is not systematically used in reliability engineering of oil and gas industry. Applying Digitalization which is a technological trend that is affecting a broad range of processes, businesses and industries, to the field of experts' judgment elicitation could result in some positive impacts.

The purpose of this thesis is to identify and suggest an "expert judgment elicitation procedure" which is suitable for obtaining experts' opinion about reliability of an equipment which is used in oil and gas industry and secondly, has the potential to be digitized. To accomplish the purpose, some of the available developed methods and procedures for expert judgment elicitation are reviewed and the criteria of a suitable method is extracted from a series of meetings and interviews with the manager of the case company. By using a 3 step approach a digital expert judgment elicitation method is suggested. Afterward, an illustrative example is used to show the application of the method. At the end, it is concluded that the suggested digital method has potential positive effects such as high efficiency and flexibility even though it can create some challenges. Moreover, if the goal of applying digitalization to the expert judgment elicitation process is to have an efficient and reliable elicitation process, we need to take in to account that analysts' expertise, technology and management commitment can affect the outcome significantly.

Acknowledgements

I would like to express my sincere gratitude to my advisor Prof. Idriss El-Thalji for the continuous support, for his patience, motivation, and valuable guidance. His advice helped me to find my way through the challenges that I was facing to accomplish my thesis. Besides my faculty advisor, I would like to thank Arnaud Barre, Integrity Operations Manager of WoodGroup for his support and dedication to this study.

Amir Beiky

Stavanger 2019

Contents

List of figures

List of tables

1. Introduction

Decision makers require relevant information to deal with uncertainty. The decision can be about huge investment in new market, introducing new species to an ecosystem or just purchasing the best family car. No matter what the decision is, the uncertainty is part of it. The uncertainty in an engineering system is because of non-cognitive sources, such as physical randomness, statistical randomness, lack of knowledge and modeling uncertainty, or cognitive sources such as parameter definition and human factors (Ayyub, 2001). If someone buys a lamp that is described on its package as “a reliable lamp”, there is uncertainty about how long the lamp is going to work and also uncertainty about the definition of “reliable” by the producer. The first uncertainty has non-cognitive source and the second uncertainty comes from cognitive sources.

Using expert judgment as a source of information in engineering field is essential especially when there is few or no empirical or historical data available (Ouchi, 2004). For example, in oil and gas industry, many risk assessment methods such as HAZOP (Hazard and Operability), HAZID (Hazard Identification) and FMEA (Failure Mode and Effect Analysis) are based on the experience and judgment of people, including engineers, technicians and managers, who are identified as experts within the scope of study.

Expert judgment in oil and gas industry is used for many years for different purposes such as assessing the oil and gas resources and risks of developing new fields. For example, when it comes to risk assessment, systematic used of HAZOP method in oil and gas industry for approximately 40 years is reported by Alain Leroy in his book “Production Availability and Reliability: Use in the Oil and Gas industry”. However, he claims in the same book that, expert judgment in a formal way is not systematically used in reliability engineering of oil and gas industry (Leroy, 2018). This statement is in accordance with the opinion of few experts in the field of reliability engineering who I talked with and interviewed them. This is somehow surprising because the reliability assessment of industrial assets has direct effect on the financial aspects of the investments on these assets. The required information and data for a proper reliability assessment is not always available due to lack of relevant historical data. New inventions and developments in equipment production or process design and exploring in new environments like Arctic conditions force the reliability engineers to rely more and more on the expert judgment. Therefore, a systematic, formal and robust methodology for using expert judgment which provides transparency and accountability in this field has obvious advantages. A possible explanation of absence of systematic and robust approach for using expert judgment in the reliability studies of oil and gas industry might come from the resource requirements for an expert study. A trained uncertainty analyst, appropriate number of experts, time and money are the resources for an expert study and they can vary significantly depending on the size and complexity of the study (Cooke and Goossens, 2004).

Digitalization is a technological trend that is affecting a broad range of processes, businesses and industries. Oil and gas industry is not an exception here and new digital capabilities are changing the decision making and asset management processes in this industry (Liyanage, J.P., Langeland, T., 2009). According to a study by Kimberlite, an international oil and gas market research and analytics company, offshore operators who use digitize data-based approach rather than a reactive approach for managing their maintenance, can reduce their unplanned downtime by 36%. This can be translated to $17 million saving each year (GE oil and gas, 2016). Digital data collection can provide great opportunity to increase the quantity, transparency and availability of data which is used for the decision-making process.

1.1. Thesis purpose and scope

As it is mentioned before, there are studies showing the benefits of digitalization in oil and gas industry and majority of these benefits come from the data collection. Elicitation is a process to obtain expert information and use it as a data set. Digitizing this process would be a continuation of a trend that has shown its positive potentials. Possible benefits of digital collection of experts' opinion are, saving time, reducing the cost of process, standardization of process, increasing accessibility and preventing errors in the process. The purpose of this thesis is to identify and suggest an "expert judgment elicitation procedure" which is firstly, suitable (based on the requirements that the industrial supervisors of this thesis have specified) for obtaining experts' opinion about reliability of an equipment which is used in oil and gas industry and secondly, has the potential to be digitized. In simple words, this thesis is trying to show how we can digitally elicit expert opinion about the reliability characteristics of an equipment.

The scope of this master thesis is to search for relevant literature that describe different elicitation methods, find the characteristics of a suitable method based on the requirements of the supervisors and suggest an expert judgment elicitation procedure which can be digitized. In addition, the suggested procedure is used for obtaining expert judgment about the reliability of an equipment as an illustrative example. In this thesis, the equipment that we use as an example is a Pipeline which is used continuously for oil and gas production. This equipment is chosen to just help the author have a clear scope of work and an illustrative example.

1.2. Methodology

It is needed to have a scientific methodology to accomplish the purpose of a Master thesis. An inductive scientific method is used for this thesis which means the conclusions are inferred from known premises. The thesis started with a literature study to understand the background of the topic and review some of the available developed methods and procedures for expert judgment elicitation. The next step was to obtain the needs of the industrial supervisor and analyze them to derive a series of requirements that are used to find suitable eliciting method among the available methods. The needs of supervisor were gathered during a series of face to face meetings, phone interviews and email communications. Afterwards, an expert judgment elicitation procedure which can be performed digitally is suggested. To verify the applicability and validity of the suggested procedure, an illustrative example is used, and the opinion of potential users are collected in a survey.

2. Theory

The first attempt to make a structured method for expert judgment elicitation was done by Research and Development corporation of Sante Monica, California also known as Project RAND in 1950s (Ayyub, 2001). The result of this project was developing two methods, Delphi method and Scenario analysis. After project RAND, many other researchers have suggested different methods for formal application of expert judgment for different applications such as Risk assessment and reliability analysis. For example, the researchers in Delft University of Technology have developed different methods that can be used for qualitative or quantitative assessments. A publication by European Commission called "A Procedure Guide for Structured Expert Judgement" also known as EUR 18820, presents the results of Delft's researchers' work (Goossens et al., 2008).

Before going further into the methods, it would be suitable to have a clear definition about some of the terms that are going to be used in next paragraphs. There is more than one definition for someone who has expertise in a topic, but the term "expert" can refer to someone who has appropriate level of knowledge about an issue and who can communicate this knowledge (Meyer and Booker 1990) Expert judgment in engineering can be defined as "informed opinion on a technical problem based on an expert's training and experience" (Kerkering, 2002). The formal procedure for obtaining information about specific questions about certain quantities such as failure rate, failure consequence or expected lifetime from experts is expert-opinion elicitation process (Ayyub, 2001). Therefore, the reliability of the information that is gathered from experts is directly related to the reliability of the expert judgment elicitation.

The reliability of the expert judgment elicitation in this thesis addresses the considerations for dealing with human biases and, having scientific approach towards the elicitation process. The judgments of experts are influenced by heuristics and cognitive biases. Heuristics give us experience-based ways to make decisions when we face complex situations with incomplete information. For example, when we use a rule of thumb to calculate something or when we make an educated guess about the weight of a piece of equipment, our heuristics are influencing our judgment. Even though, in many occasions, heuristics are helpful and positive, they can sometimes lead to bias. Cognitive biases can arise from different things such as heuristics, motivational circumstances and social factors. These biases can cause reasoning errors, wrong interpretations and subsequently irrational decisions (Baybutt, 2017).

When the expert judgment is used to assess a certain parameter, not only a suitable elicitation method is required but a proper aggregation technique is needed to create a single combined assessment of the parameter. For example, if the experts are judging the number of failures of an equipment per year and there are differences in their judgment, an aggregation method combines their assessment and give a single failure rate that can be used for analysis. The aggregation can happen in a group discussion among experts or by a mathematical method using single assessments. Even though, in a control environment where experts are rational and have all the necessary information, a group of experts generally perform better than individual ones, Clemen and Winkler in a study on 1999 showed that the best individual in the group can outperforms the group as a whole (Clemen and Winkler, 1999). According to Cooke, this phenomenon motivates "the elicitation of the assessments of individual experts without any interaction, followed by mathematical aggregation in order to obtain a single assessment per variable, thereby weighting the individual experts' assessments based on their quality" (Cooke, 1991 & Goossens et al., 2008).

2.1. Elicitation methods

There are different elicitation methods developed by different organizations and researchers. The first attempt to make a structured method for expert judgment elicitation was done by Research and Development Corporation of Sante Monica, California also known as Project RAND in 1950s. Researchers such as Cooke and Goossens also spent many years to develop guidelines for elicitation process. Here we introduce these different methods.

Delphi method:

Delphi method which is developed by RAND project is one of the most famous elicitation methods that has been used in a variety of fields. This method can be used for technological forecasting and it requires a group of knowledgeable experts within the scope of study and facilitators or analysts who plan the process and facilitate the group to create a consensus opinion around the questions of concern (Ayyub, 2001). A basic Delphi method has the following steps (Helmer 1968):

1. Issues or questions selection and development of questionnaires.
2. Selection of most knowledgeable experts within the issues or questions of concern.
3. Familiarization of experts with the issue by providing enough details
4. Elicitation of experts about the issues. The experts might not know who the other respondents are.
5. Presenting the results in the form of median values and an inter-quartile range (i.e., 25% and 75% percentile values).
6. Reviewing the results of elicitation by the experts and revision of initial answers by them. This step sometimes increases the accuracy of results. Experts who provide answers outside the inter-quartile range need to provide written justifications or arguments on the second cycle of completing the questionnaires.
7. Revision of results and re-review for another cycle. The process should be repeated until a complete consensus is achieved. Typically, the Delphi method requires two to four cycles.
8. A summary of the results is prepared with argument summary for out of inter-quartile range values.

There is no geographical concern for expert selection in Delphi method. The experts can answer the questions by email and after each round they receive feedback about the most to least mentioned answers (Avella, 2016). At the end of the last cycle usually the answers of experts are much closer to each other compare to the first cycle. The purpose of repetition is not achieving 100% agreement and typically there is a range from 55 to 100% agreement with 70% considered the standard (Vernon, 2009). The median value is usually taken as the best estimation for the issue of concern.

The analyst or facilitator who plan the Delphi process has to control the debate. Since there are different experts from different disciplines involved in this process and because of the anonymity provided by the process, individual panel members might present extreme and out of range opinions. The facilitator should avoid being judgmental and let the extreme opinions receive a level of attention that popular opinions receive. However, the Delphi process is designed to find those areas where the group of experts find consensus (Avella, 2016).

The advantages of the Delphi method are Flexibility and simplicity, knowledge sharing among experts from different disciplines, cost effectiveness (meaning except the time of participants and the analyst, there is usually no other costs), freedom of expression which is the direct result of anonymity, ease of communications, membership variations and lack of geographical limitations. However, this method like most other research designs has some disadvantages such as the influence of analyst bias and weaknesses on the result of the study and, lack of commitment from the participants because of anonymity or because of being among minority (Avella, 2016).

Figure 2-1 is an example that shows first part out of four parts of the questionnaire that was originally used by Helmer, one of the first developer of the Delphi method in 1968.

Questionnaire #1

- This is the first in a series of four questionnaires intended to demonstrate the use of the Delphi technique in obtaining reasoned opinions from a group of respondents.
- Each of the following six questions is concerned with developments in the United States within the next few decades.
- In addition to giving your answer to each question, you are also being asked to rank the questions from 1 to 7. Here "1" means that in comparing your own ability to answer this question with what you expect the ability of the other participants to be, you feel that you have the relatively best chance of coming closer to the truth than most of the others, while a "7" means that you regard that chance as relatively least.

Rank	Question	Answer*
☐	1. In your opinion, in what year will the median family income (in 1967 dollars) reach twice its present amount?	
☐	2. In what year will the percentage of electric automobiles among all automobiles in use reach 50%?	
☐	3. In what year will the percentage of households that are equipped with computer consoles tied to a central computer and data bank reach 50%?	
☐	4. By what year will the per-capita amount of personal cash transactions (in1967 dollars) be reduced to one-tenth of what it is now?	
☐	5. In what year will the power generation by thermonuclear fusion become commercially competitive with hydroelectric power?	
☐	6. By what year will it be possible by commercial carriers to get from New York to San Francisco in half the time that is now required to make that trip?	
☐	7. In what year will a man for the first time travel to the moon, stay at least one month, and return to earth?	

* "Never" is also an acceptable answer.

- Please also answer the following question, and give your name, (this is for identification purposes during the exercise only; no opinions will be attributed to a particular person).

Check one:

☐ I would like to participate in the three remaining questionnaires

☐ I am willing but not anxious to participate in the three remaining questionnaires

☐ I would prefer not to participate in the three remaining questionnaires

Name (block letters please): ____________________________

Figure 2-1: Delphi Questionnaire used by Helmer (1968)

Scenario Analysis:

Kahn and Wiener developed the scenario analysis for evaluating the uncertainty about future events in 1967. Kahn defines a scenario as “a set of hypothetical events set in the future constructed to clarify a possible chain of casual events as well as their decision points”. There are two questions that scenario analysis tries to answer. First, how a hypothetical situation might happen, and second, what are the available alternatives for each stakeholder for preventing, diverting, or facilitating the process? Event tree analysis and decision tree analysis are representatives of these two questions (Ayyub, 2001).

An example of scenario analysis can be seen in HAZOP studies where a group of experts go through a chemical process and evaluate different scenarios such as over pressure, high or low temperature, high or low flow and suggest the possible outcomes of each scenario. The experts also try to assess the probability of each scenario and come up with an action that needs to be taken to prevent or mitigate the undesired outcomes. Kahn and Wiener did not use scenario probabilities, because they were interested in realizing the long-term trends and the scenarios that they considered had very small likelihood (Ayyub, 2001).

Paired comparison:

In this method, experts rank alternatives pairwise according to some defined criterion like Safety, feasibility, etc. For example, with 10 alternative items, 45 comparisons must be made. There are methods suggested for ranking the items and based on the chosen method and availability of some measured or known values, there is a possibility to create an interval or even a ratio scale from the data. However, in general, the method of paired comparisons does not provide an assessment of uncertainty (Cooke and Goossens, 2004).

One of the popular models based on the paired comparison method is the Negative Exponential Life (NEL) model. The model follows the basic principal of paired comparison method where n components are compared pairwise and then ranked. The answers of experts are assessed to realize if there is a consistency in their answers or not. This is done by analyzing the count of circular triads in the expert's comparisons which means that if the expert says A>B and B>C then he or she should say A>C (Chinyamakobvu, 2015).

The roulette method

Gore in 1987 suggested an elicitation method which is called the trial roulette or sometimes chips and bins method. In this method the expert receives n chips and is asked to allocate them to a grid of m equally sized bins which cover the range of the variable that we are trying to get information about. (Morris, Oakley and Crowe, 2014). Afterwards it is possible to calculate the probability of our variable lying in a particular bin based on the proportion of chips allocated to that bin. For example, if the expert uses 20 chips then each of them present 5% of the mass of a probability distribution (Veen et al., 2017). Figure 2-2 shows an example of chips and bins method outcome.

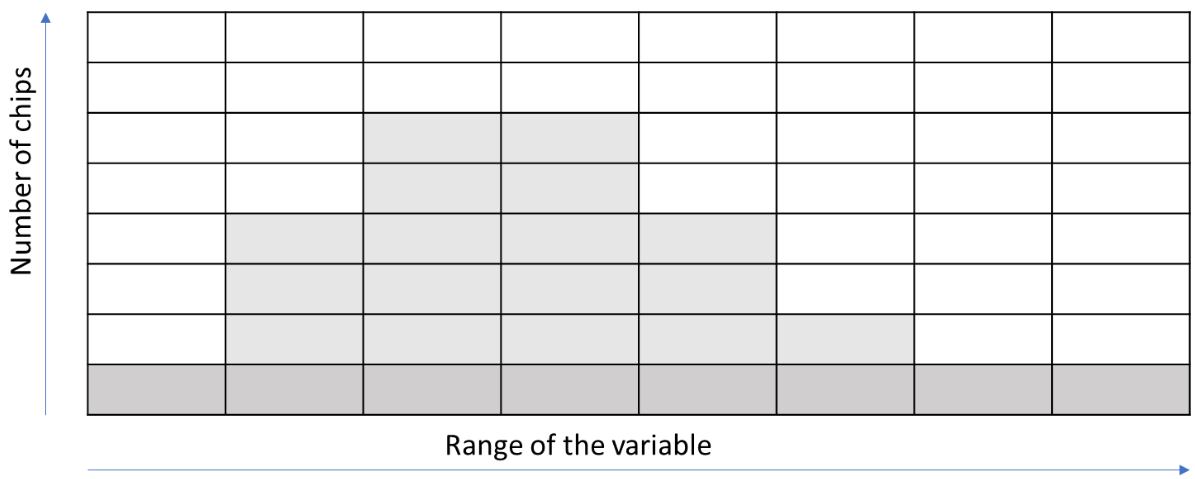


Figure 2-2: An example of the outcome of roulette method

One of the advantages of this method is that if done graphically, the expert can see the shape of his or her judgment's distribution while he or she allocates the chips. This can provide the opportunity for the analysts to make sure that they collect the true opinion of the judgment and wrong interpretation is not involved. After allocation of chips to bins by expert, the analyst can fit a distribution to the variable under investigation.

Classical Model:

Cooke is one of the known scientists in the field of expert judgment elicitation and he explains that expert judgments should follow the principles of the scientific process to be accepted as scientific data. These principals are accountability, neutrality, fairness, and the ability for empirical control (Cooke, 1991). Classical model in an eliciting approach which is performance-based and uses weighted averaging for aggregation. This model is trying to follow the scientific process as Cooke suggested. In this model, experts express their opinion about the uncertainty over the variable of interest and some seed variables. The variable of interests are the variables that the analysts need the expert judgments about them and seed variables are the variables within the field of experts which are uncertain for the expert but known to the analyst. The performance of experts would be scored based on their answers to the calibration questions about seed variables. The purpose of using seed variables is to evaluate experts' performance, combine the experts' judgments based on their performance and evaluate the combination of experts' judgments (Cooke and Goossens, 2004).

Classical model has three different weighting schemes for aggregating the elicited experts' opinions. The first Scheme is Equal weighting which assigns equal weights to each expert. Global weighting, and item weighting are the techniques that reflect the performance-based weighting aspect of classical model. The global weights are assigned to each expert based on their answers to the calibration questions and based the information score of an expert. The information score is based on the uncertainty intervals that the expert provides with regards to the available information (Cooke and Goossens, 2004).

Protocol and procedures guide:

The "protocol and procedures guide" is a document by Cooke and Goossens (2000) and it provides a guideline for a full expert judgement elicitation. This protocol is especially suitable for the situations that we need experts' judgment to make an uncertainty distribution of a variable. The European Commission has published this protocol as EUR 18820. The procedure has 15 steps (Cooke and Goossens, 2000):

1- Definition of case structure

2- Identification of target variables

3- Identification of query variables

4- Identification of performance variables

5- Identification of experts

6- Selection of experts

7- Definition of elicitation format document

8- Dry run exercise

9- Expert training session

10- Expert elicitation session

11- Combination of expert assessments

12- Discrepancy and robustness analysis

13- Feed back

14- Post-processing analyses

15- Documentation

Looking at these steps, it is clear that this protocol has the performance-based approach and has used some of the classical model characteristics.

2.2. summary

As it was mentioned before, there are other eliciting methods developed for different purposes (Rai, 2013 and Veen et al., 2017). The differences between methods can come from their behavioral or mathematical approach, aggregation techniques or weighing systems. However, the purpose of this thesis is not to review all the expert judgment eliciting methods.

3. Elicitation Method realization

The case company for this thesis is a consultancy company which provides wide range of services and one of which is developing RAM (reliability, Availability, Maintenance) models and studies for companies active in oil and gas industry. The focus of this chapter is to explain the status quo of expert elicitation process for RAM analysis within the case company and then suggest a suitable digital method for expert elicitation. To develop and suggest a proper expert elicitation methodology, it is essential to know what are the requirements for that method? I extracted these requirements from the needs of the supervisors. The information provided in this chapter are gathered form the supervisors of this thesis in the case company by face to face interview, phone interview or email communications.

3.1. Status quo

To make a RAM model and provide a RAM study for a facility, the case company uses different approaches depending on the type of that specific facility, type of components and some other factors such as scope of study and resource availability. One of the required information for performing RAM analysis is the failure data of the components. A typical approach of the company to gather and use the failure data consist of the following steps:

1. As a starting point using the OREDA (Offshore and Onshore Reliability Data) handbooks to obtain the failure data stated in them.
2. Data sharing within the company and different offices. This step happens on a varying degree and the quality of it is very project and people dependent. (There is typically a degree of “silo thinking” and internal rivalry between offices. Also, there may be some hesitation to share data at the risk of triggering lengthy discussions and requirements to “defend” for a study that may be outside the core scope for peer engineers and offices)
3. Over the years, the company has gained access to several reference studies and reports which may serve as “benchmarking” or “reference data sets”
4. For some specific studies, the client team may have historical data available.
5. The assumptions made are captured and documented in the RAM model report write-up, including reference to OREDA, previous RAM studies, site / asset survey reports and

method statement to describe the process of forming assumptions and (on a high level) who has been involved in the process.

6. Often a FMEA / FMECA workshop or similar is conducted, where also failure data sets may be discussed and captured

As it can be observed from this approach, at the moment, the company does not put much emphasize on the experts' judgment of failure data. Only the FMEA/FMECA workshops can be considered as the elicitation process that uses scenario analysis method. This approach can be challenging for studies that target the newly developed equipment with rare and few historical data, or the equipment that operate in specific and harsh environments like the environment of arctic offshore (Naseri,2016).

The other problem with the current approach is that using a scenario-based elicitation method (FMEA) increases the cost of doing a RAM analysis for a small project significantly. As an example, according to the senior reliability engineer of the company, for running a FMEA workshop for a pump, a RAM analyst needs to spend around 1.5 weeks for preparation, facilitating the meeting and making the report, a group of approximately 8 experts need to gather for at least one day workshop and a scriber is needed for that day. This process is time consuming and subsequently costly.

3.2. Translating needs into requirements and criteria

To address the mentioned problems of the current situation in the company regarding to collecting experts' judgment for RAM analysis and using the potential benefits of digitalization, the integrity operations manager of the case company wants a suitable digital method for expert elicitation. To identify, suggest and develop a suitable method and process we must find the characteristics of a suitable method based on some requirements. These requirements can be extracted from the needs of the supervisors.

During a series of face to face meetings telephone interviews and email communications, the integrity operations manager of the case company explained that he needs a "general process" that can be used for "expert judgment collection" for different studies. He added that with the current approach, some of the assumptions regarding to experts' opinions (for example, more experience means more reliable opinion) “is not accurate” and it is very hard to find out “how the experts performed on a certain study”. Moreover, he thinks that many “different biases affect the result of scenario-based elicitation methods” and “the opinion of the experts who perform better in elicitation process should be weighted more in the conclusion”. Last but not the least, knowing about the benefits of digitalization, the manager wants to have a digital process that can be part of a bigger digitalization plan (combination of servers, database and software).

Table 3-1 shows the summary of the needs of the manager and the translation of them into the requirements. The requirements indicate criteria that should be considered for suggesting a suitable elicitation method.

Table 3-1: Translation of Needs to Requirements and Criteria

NEED	REQUIRMENT	CRITERIA
a “general process” that can be used for “expert judgment collection”	A structured method for expert judgment elicitation	Comprehensive
Assumptions regarding the experts’ judgment is not accurate currently and it should be addressed	A method with clear and strong basis	Scientific approach, Performance-based
To know how the experts performed on a certain study	A method that provides the opportunity for performance review of experts	Performance-based
Dealing with the biases	The method should eliminate or reduce the effect of biases on the result of study	Not asking for single estimate but try to elicit a distribution, Providing feedback

the opinion of the experts who perform better in elicitation process should be weighted more	The method should give more credit in the aggregation phase to the experts who perfume better	Performance weight rather than equal weight

3.3. Methodology suggestion

In the theory chapter of this thesis some of the different methodologies for expert opinion elicitation are introduced. The differences between methods can come from their behavioral or mathematical approach, aggregation techniques or weighing systems. However, by studying the description of different methods, I concluded that there are three general steps in each of these methods. These three steps that I have shown them in figure 3-1 are the basis of the method that I suggest for digital elicitation of expert judgment. In this figure, I have also illustrated some examples of possible issues affecting each step. The first step for elicitation is to design questions. It is important to know whether we want single point values for the variables under question, or we want a distribution representing the variables. This can be determined depending on the decision makers' preference. For example, if the decision is not critical and a comprehensive representation of uncertainties is not required, asking for single point values is simple and convenient.

Another important factor in the designing questions is the effect of biases on the experts. The questions shall not introduce or promote any type of biases to the study. For example, if a question presents an "example value", it will create anchoring effect and the answers of experts to that questions are most likely close to the "example value".

Asking the designed questions from the experts is the second step of elicitation process. In this step we must decide about the communication method for asking the questions and the communication between experts. Here the analyst should decide if he or she wants to perform this step as a group activity or run it individually. It is obvious that, an unbiased approach is required for asking the questions from the experts.

The last step of the process of eliciting the judgment of experts is to aggregate their judgments. The aggregation can happen in a group discussion among experts or by a mathematical method using single assessments. Therefore, communication has also impact on step 3. The other important factor is the weight that we give to the opinion of our experts. If their level of expertise is different and we need to reflect that in the aggregation, then we need to have a proper weighing system. There are different mathematical aggregation methods that can be used based on the required level of complexity and the advantages of each method (Ouchi, 2004).

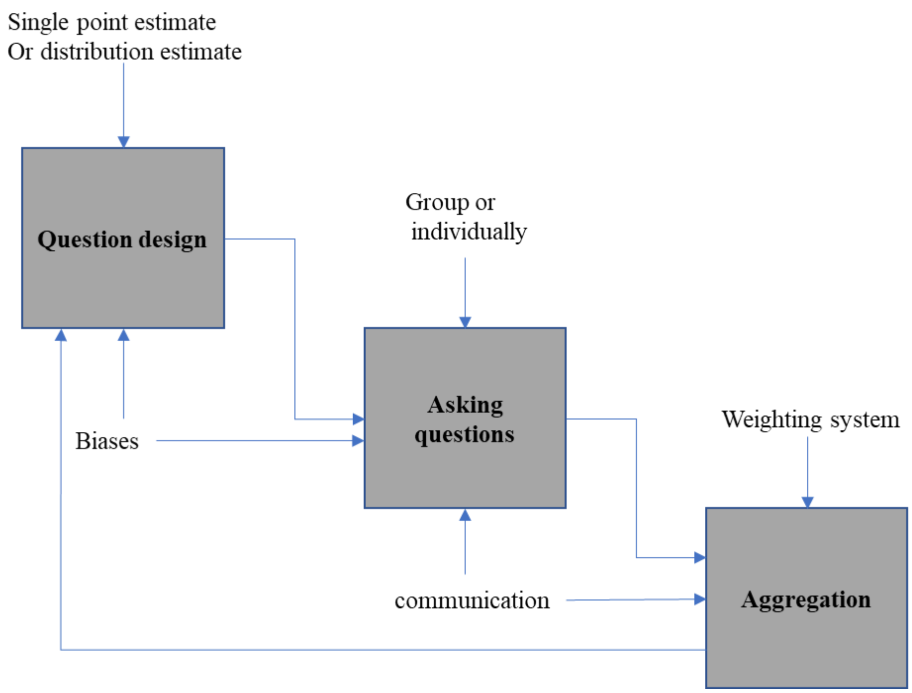


Figure 3-1: General steps in expert opinion elicitation and issues around them

To make a digital elicitation process we must digitize the steps that are shown in figure 3-1. The degree of digitalization in each step might be different depending on the methodology that we pick for the process. My suggestions for each step are as following:

Step 1

Designing the elicitation questions is part of the analyst tasks. There are many researches around the topic of biases and different methods to mitigate them. The book "thinking fast and slow" by Daniel Kahneman provides a comprehensive picture about this topic. My conclusion from this book is that the biases are very complex and there are some ways to mitigate their effect, but the most important tool is to aware the experts about them.

Asking about a distribution rather than a single point answers let the experts to express their uncertainty (Morris, Oakley and Crowe, 2014). The suggested methodology is going to be used for RAM analysis and presenting a clear picture of uncertainties is important for decision makers who use the results of analysis. Hence, my suggestion is to first provide a short introduction of biases and their influence on the human judgment for experts and then ask the questions which are targeting elicitation of distributions.

Step 2

When it comes to asking the questions, there are researches that motivates individual elicitations (Cooke, 1991 & Goossens et al., 2008). Considering the extracted criteria (table 2-1), the suitable method for asking questions should provide feedback to the experts, so they can express their exact opinion. Adding this point to the fact that for the first step we have decided to elicit a distribution rather than a single point, I suggest an elicitation method which is called the trial roulette method also known as chips and bins method. In this method, experts assign “chips” to “bins” of a histogram to express a probability. This method has been digitized by David Morris and Jeremy Oakley in a project called MATCH and I have decided to use their web-based tool (Morris, Oakley and Crowe, 2014). Using this tool helps to provide a real time visual feedback to experts. Figure 3-2 shows a picture of roulette method in MATCH. This software also provides other input modes but, all of them are following the same principal which is giving a visual feedback to the expert based on their input. The producers of this software have reported that, in their experience, users have found out the roulette method the easiest to use (Morris, Oakley and Crowe, 2014).

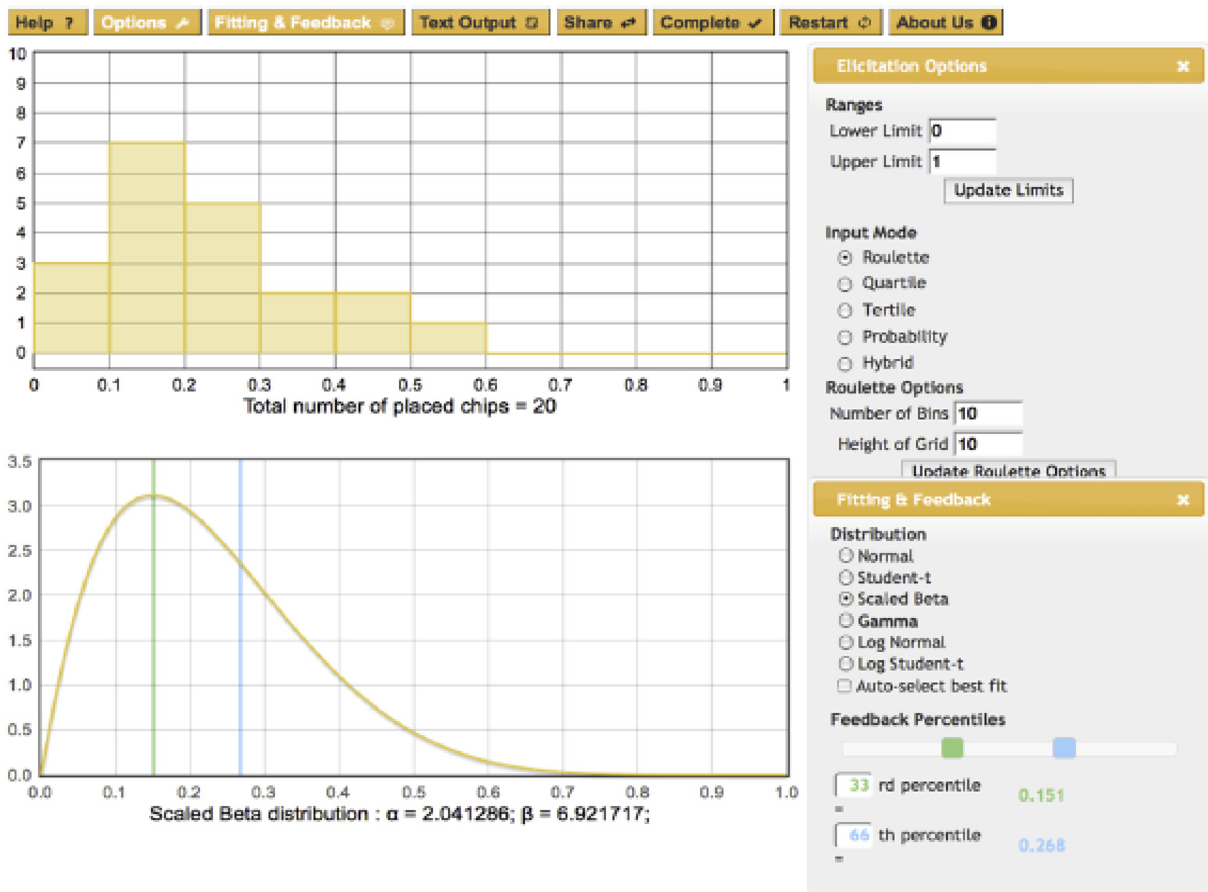


Figure 3-2: MATCH software output for eliciting a probability distribution using the roulette method. Available at: http://optics.eee.nottingham.ac.uk/match/uncertainty.php

Step 3

Looking at the table 3-1 and comparing the criteria with the available aggregating methods, classical model has the characteristics of a suitable method for step 3, since it follows the principles of the scientific process and it is a performance-based methodology. Excalibur is a software which is developed by TU delf and is based on this method. For the step 3, I suggest the use of this software which simplifies the calculations for different weighing systems which are introduced by classical model. Figure 3-3 shows the interface of this software. The main input for calculations in this software is different percentiles of experts' opinion distributions. This means the output of the step 2 can be easily used for this step.

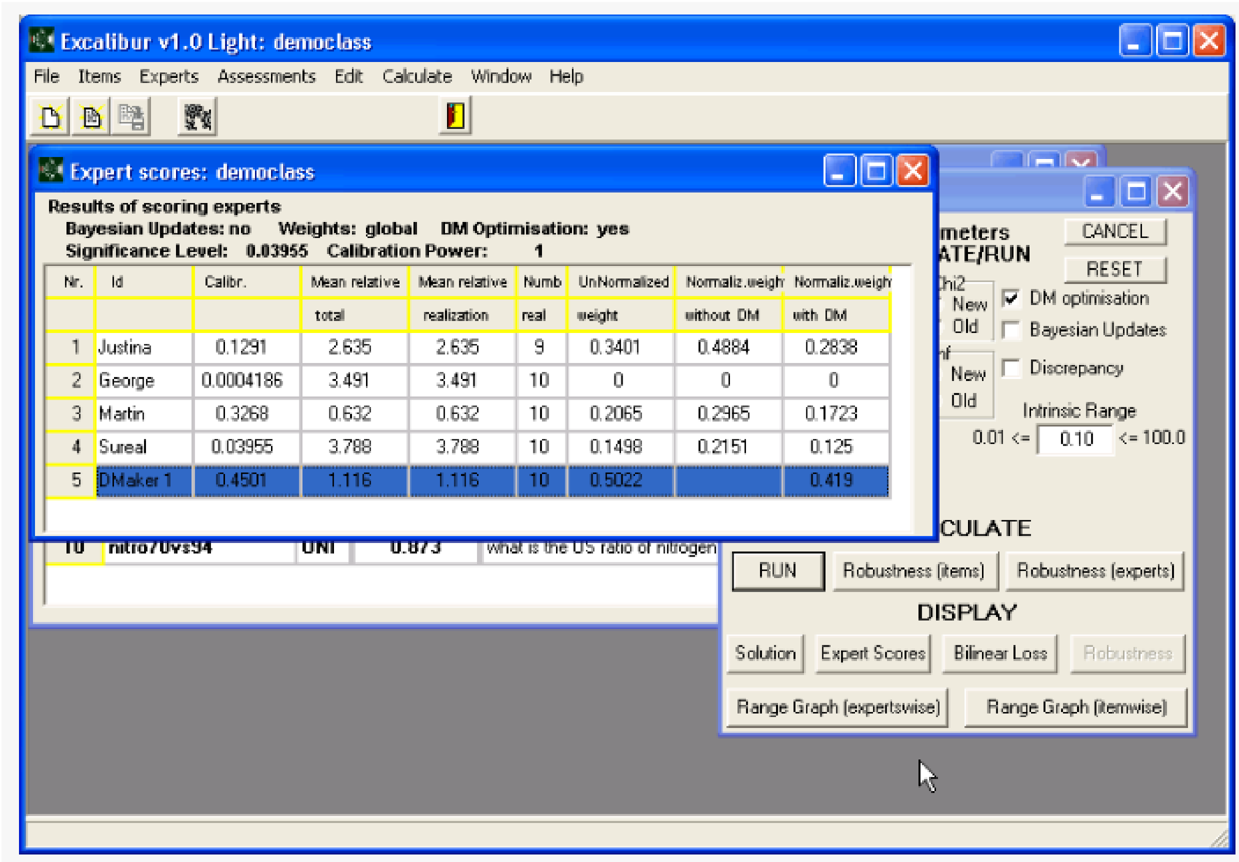


Figure 3-3: Excalibur interface Available at: http://www.lighttwist.net/wp/excalibur

Figure 3-4 concludes the suggested methodology. First, suitable questions to elicit a distribution for the target variable and the performance related questions are designed. After providing a short guideline about the methodology and warning the experts about the biases, they answer the questions individually by using the roulette method. Last but not the least, the analyst aggregates experts' judgments, using the classical method and global weighing.

Figure 3-4: Expert judgment elicitation procedure

4. Illustrative example

In this section the use of suggested method is illustrated by using a real-life example from the oil and gas industry. The purpose of this example is not providing accurate reliability data. This example is simply illustrating how the suggested method in section 3 can be used for expert elicitation.

4.1. Introduction

Oil and gas production from high temperature and high-pressure fields has its own challenges and one of which is transporting the production through pipelines. When the oil and gas with high temperature is transported through pipelines located at sea floor, the fluid lose its heat since the temperature is low. Hence the possibility of hydrate, wax, and asphaltene formation increases and there would be a risk of blockage. Since this solution is relatively new, there is not enough historical data about the failures of different components of this solution to make a proper RAM analysis. RAM study aims to ensure the reliability and the maintainability of the pipeline are well assessed, such inspection plan, monitoring, spare part philosophy and the manning are adapted, cost effective and optimized during the operation. For much of the deep-water equipment, little or no historical data is available. This shows the importance of expert opinion for deep-water systems' reliability analysis.

The case company has tried to perform a RAM analysis for a Pipe-in-pipe system and they found out that in this analysis the opinion of experts should often be used to adjust existing data. For collection of experts' judgments, a FMEA study has been done. In this section, I try to elicit the judgments of experts about the time to failure of two of the identified failures by FMEA study, using the digital methodology suggested in section 3. The two identified failures that I am targeting are "corrosion (internal or external) failure resulting loss of containment" and "impact failure resulting loss of containment".

4.2. Guideline and Questionnaire

To inform the experts about the purpose of the project and guide them how they should use the MATCH software, a brief guideline was is that explains the steps that they should take to enter their inputs to the software and get the feedback about their inputs from the software in the form of distributions. Appendix I shows the guideline.

There are two types of questions that experts should answer. First, the calibration questions which are asking about the seed variables. As it was explained in chapter two, classical model, seed variables are the variables within the field of experts which are uncertain for the expert but known to the analyst. I designed these questions based on the failure data in the Pipeline and Riser Loss of Containment report (PARLOC 2012) which contains the risk assessment data for generic loss of containment frequencies and covers pipelines and risers in the offshore oil and gas industry. The failure rates stated in this report are not applicable for the pipe-in-pipe technology, but they are a good indicator of the performance of the experts. Second type of questions are about the target variables. My target variables are the mean time to loss of containment due to two failure modes that are identified during a FMECA for pipe-in-pipe technology. Table 4-1 shows the seed and target variables.

Table 4-1: Seed and target variables

SEED	TARGET
Number of observed “impact failures resulting loss of containment” during 12 years among 1372 steel pipelines in UK continental shelf.	Mean time to loss of containment due to “Internal or external corrosion” for pipe in pipe pipelines
Number of observed “material failures resulting loss of containment” during 12 years among 1372 pipelines in UK continental shelf.	Mean time to loss of containment due to “impact failure” for pipe in pipe pipelines
Number of observed “operation and maintenance failures resulting loss of containment” during 12 years among 1372 pipelines in UK continental shelf.	

Number of observed "construction failures resulting loss of containment" during 12 years among 1372 pipelines in UK continental shelf.	

Since probability might not be a familiar topic for the experts I have decided to use natural frequencies for asking the questions. There are researchers supporting the fact that using natural frequencies can improve the result of elicitation (Burgman et al., 2006). The following are the assumptions and the questions for the elicitation process. The complete questioner is presented in appendix II.

Assumptions:

- *Impact failure*: failures due to anchoring, trawling and other impacts.
- *Material failure*: failures due to corrosion (internal or external) and other material causes.
- *Operations and maintenance failure*: incorrect operation (such as over-pressurization) or error during maintenance or removal (e.g. release resulting from incomplete flushing of the line prior to opening the line).
- *Construction failure*: failures during commissioning tests or occurring as inadvertent by-product of construction.

Questions part 1:

1. Using the MATCH software please express your opinion about the number of observed "impact failures in midline part of pipeline resulting loss of containment". To express your opinion, you can use the X-axis as the number of impacts and 20 chips to show your uncertainty (more chips for the number that you are more certain about it).
2. Using the MATCH software please express your opinion about the number of observed "Material failures in midline part of pipeline resulting loss of

containment”. To express your opinion, you can use the X-axis as the number of material failures and 20 chips to show your uncertainty.

3. Using the MATCH software please express your opinion about the number of observed “operation and maintenance failures in midline part of pipeline resulting loss of containment”. To express your opinion, you can use the X-axis as the number of material and maintenance failures and 20 chips to show your uncertainty.
4. Using the MATCH software please express your opinion about the number of observed “Construction failures in midline part of pipeline resulting loss of containment”. To express your opinion, you can use the X-axis as the number of construction failures and 20 chips to show your uncertainty.

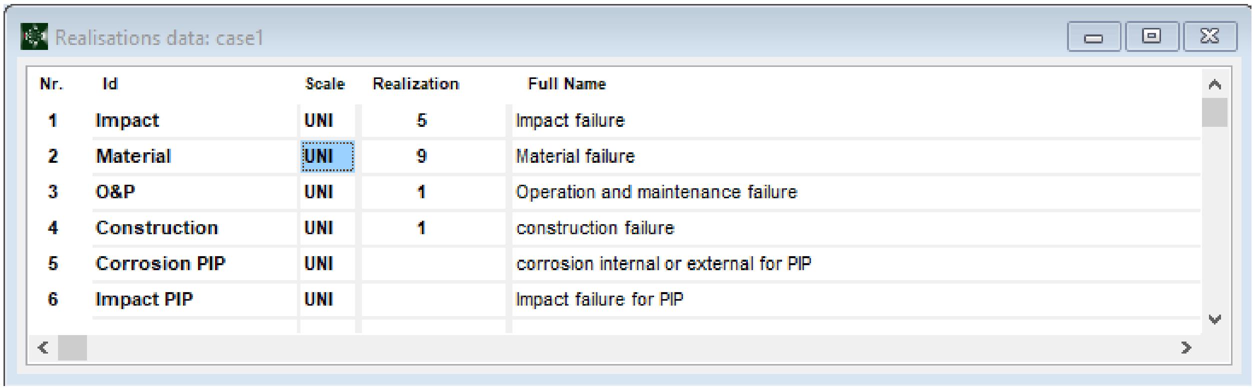
Questions part 2:

For the following questions assume we have 20 identical Pipe in Pipe pipelines at the sea bed for transporting oil (oil properties, temperature and pressure is constant in all the pipes). These pipelines (midline) are not protected and exposed.

1. Using the MATCH software please express your opinion about the time that it takes until the first “corrosion (internal or external) failure resulting loss of containment” happens in the pipelines. To express your opinion, you can use the X-axis as the time of failure in years.
2. Using the MATCH software please express your opinion about the time that it takes until the first “impact failure resulting loss of containment” happens in the pipelines.

4.3. Elicitation result

The communication with experts was through email and 4 experts participated in the study. By using MATCH software, they provided their judgment about each question in the form of a distribution. Then the 5%, 50% and 95% percentiles of these distributions were collected for aggregation. To aggregate the experts’ judgment, EXCALIBUR software was used. The first step to use EXCALIBUR in to define the variables and provide the known data for seed variables. Figure 4-1 shows the target and seed variables in the EXCALIBUR.



Realisations data: case1

Nr.	Id	Scale	Realization	Full Name
1	Impact	UNI	5	Impact failure
2	Material	UNI	9	Material failure
3	O&P	UNI	1	Operation and maintenance failure
4	Construction	UNI	1	construction failure
5	Corrosion PIP	UNI		corrosion internal or external for PIP
6	Impact PIP	UNI		Impact failure for PIP

Figure 4-1: Seed and target variables in EXCALIBUR

As it can be seen in the figure4-1, the first 6 variables are seed variables and their true values (Realization) that we have extracted from PARLOC 2012 are in front of them. Variable number 5 and 6 are the target variables. I have chosen uniform scale (UNI) for all the variables which means I assumed these variables are uniformly distributed between quantiles. The second step is to enter the percentiles collected from experts' judgment distributions in the software. Table 4-2 present these percentiles.

Table 4-2: Experts' judgments

	Percentiles			
	5%	50%	95%	
Expert 1	24.6	34.6	48.7	Qestion1
	35	16	35	Qestion2
	5.45	8.08	9.95	Qestion3
	3.6	4.4	4.84	Qestion4
	9.51	12.9	14.6	Qestion5
	5.47	6.92	7.72	Qestion6
Expert 2	19.4	53.7	85.5	Qestion1
	5.41	27.2	62.4	Qestion2
	3.43	15.7	38.8	Qestion3
	3.39	15.9	39.6	Qestion4
	6.01	12.1	17.2	Qestion5
	3.38	13.7	30	Qestion6
Expert 3	2.14	10	26	Qestion1
	2.14	10	26	Qestion2
	1.64	10	23	Qestion3

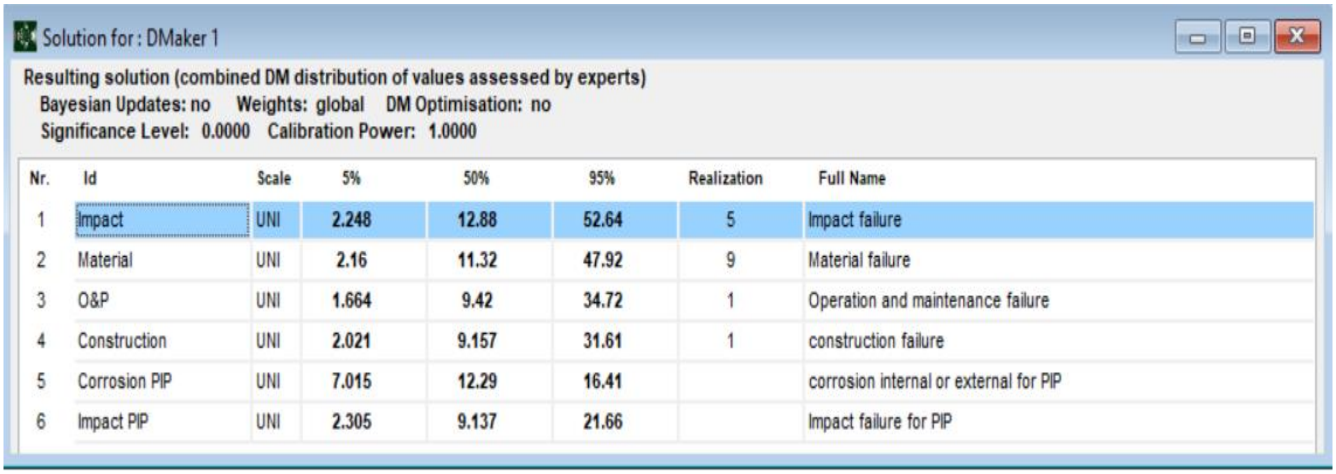
	2	10	23	Qestion4
	7.36	12.2	16.4	Qestion5
	2.25	10	20	Qestion6
Expert 4	13.8	21	26.5	Qestion1
	104	125	147	Qestion2
	69.4	133	182	Qestion3
	56.3	74.3	92.2	Qestion4
	1.32	3.74	6.16	Qestion5
	1.42	2.5	4.42	Qestion6

Excalibur provides different calculations and different weighing system. In this thesis however, only the Global weighing for simple calculation of the results for a decision maker in used. Table 4-3 shows the calculated global weights.

Table 4-3: Result of experts' global weights

Expert	Normalize Global Weight
Expert 1	0.11
Expert 2	0.03714
Expert 3	0.8523
Expert 4	0.0006329

The results for the aggregation are presented in Figure 4-2.


Solution for : DMaker 1

Resulting solution (combined DM distribution of values assessed by experts)
Bayesian Updates: no Weights: global DM Optimisation: no
Significance Level: 0.0000 Calibration Power: 1.0000

Nr.	Id	Scale	5%	50%	95%	Realization	Full Name
1	Impact	UNI	2.248	12.88	52.64	5	Impact failure
2	Material	UNI	2.16	11.32	47.92	9	Material failure
3	O&P	UNI	1.664	9.42	34.72	1	Operation and maintenance failure
4	Construction	UNI	2.021	9.157	31.61	1	construction failure
5	Corrosion PIP	UNI	7.015	12.29	16.41		corrosion internal or external for PIP
6	Impact PIP	UNI	2.305	9.137	21.66		Impact failure for PIP

Figure 4-2: Results of aggregation calculations

It is also possible to present the results for the target variables as a cumulative distribution function since the software generate the distributions. Figure 4-3 and 4-4 present the cumulative distribution function (CDF) for target variable. The X axis shows the "Time to failure" in years and Y axis shows the probability that the Time to failure takes a value equal or less than a certain point on X axis.

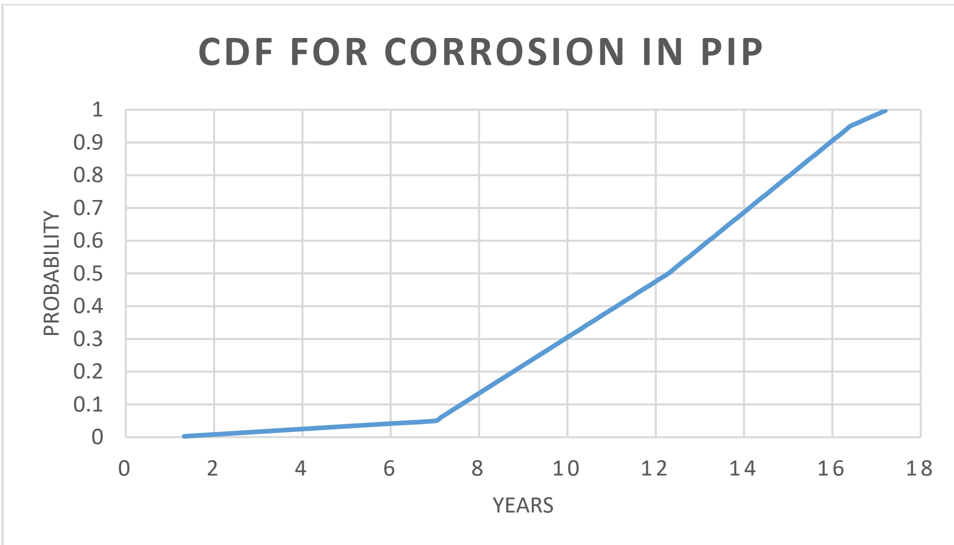


Figure 4-3: CDF for Time to failure (corrosion in PIP)

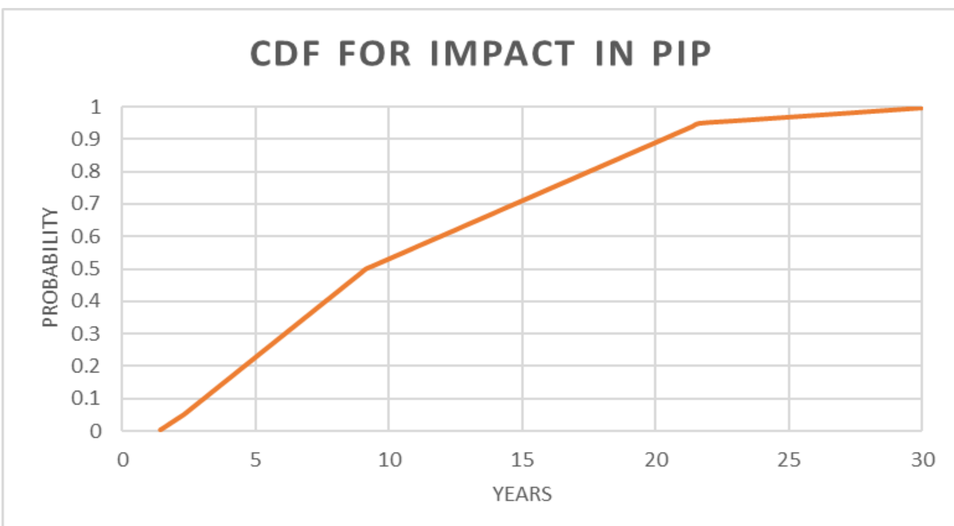


Figure 4-4: CDF for Time to failure (Impact in PIP)

The result of elicitation which is presented in figures 4-2, 4-3 and 4-4 can be used in a RAM analysis directly or combined with the other sources of available information. Based on the calculated CDF for time to corrosion failure, we can for example suggest inspection schedules or maintenance programs to reduce the risk of failure. It is also possible to determine the probability of interruption in the functionality of pipeline which subsequently can affect the production rate.

5. Discussion and conclusion

By reviewing different expert elicitation methodologies, I recognized that there are 3 main steps to perform an expert judgment elicitation (figure 3-1). To identify and suggest a digital expert judgment elicitation procedure which is suitable for the need of case company (to be used in RAM analysis), I investigated the available alternatives for each step based on the needs of the industrial supervisor of this thesis and the characteristics of the different available methodologies. Therefore, the suggested methodology might not be applicable for other purposes and companies, even though I tried to suggest a methodology which is as general as possible.

However, the 3 steps approach that I have used to conclude the suggested method can be used to make elicitation procedures for other purposes. For example, the suggested methodology in this thesis is mainly suitable for RAM studies (i.e. building data dossier) when a more quantitative results are desirable. For the studies which are looking for qualitative results, we need to use alternative methods and techniques which are suitable for qualitative studies for each step depending on the identified criteria and based on the needs of decision makers.

The illustrative example shows how expert's judgment can be collected in a virtual environment with the suggested methodology and supporting software. There was not any face to face meeting with experts for the elicitation and the communications were through email. This means there was not any cost related to traveling to perform the elicitation. Moreover, reading the guideline and answering questions by MATCH software did not required substantial amount of time and the process provided more flexibility for experts and they had the possibility to answer the questions whenever it suits them better. Reducing the time required for the process can be translated to reducing the cost of the process. Hence, one of the advantages of using digital elicitation method is reducing the cost of using the experts' judgment. A report from Harvard university (Diaz anadon et al., 2014) also support this conclusion when it states "An extremely conservative back of the envelope calculation of the monetary benefits (i.e. excluding benefits in future years, assuming that researchers travelling to interview experts do not need accommodation, and ignoring the time and effort savings to researchers and experts) indicates that online surveys with 11 experts are at least 40% cheaper than in-person elicitations with the same number of experts".

Even though there are potential benefits in the suggested 3 step digital expert elicitation, there are some challenges regarding to each step that need to be addressed. The first challenge is to design well defined and comprehensible questions and find suitable seed variables. The questions that are asked from experts should be complete and clear since a misunderstanding in the meaning of the question can reduce the accuracy and effectiveness of the elicitation significantly. Moreover, finding seed variables that are related enough to the target variables, and their true values are known, can be difficult. In this thesis, I used an older technology (steel pipeline) and historical data (PARLOC Report) for this purpose but we do not always have access to these resources for different studies. The next important point here is that we use expert's judgment to pick the seed variables and check the performance of other experts which means another source of uncertainty is introduced to our analysis.

The next challenge is to inform, train and encourage experts to use the digital tool to express their opinion. Even when the questions are comprehensive, if the experts are unable to use the tools properly, the elicitation reliability declines. In this study I asked experts to answer two questions about using the software. The questions are:

1. On a scale of 1 to 10, how easy did you find answering the questions using the software when 10 means extremely easy.
2. On a scale of 1 to 10, how useful you find the feedback process in the software when 10 means extremely useful.

Table 4-4 shows the answers of the experts to these questions. From these answers we can infer that the experts did not have a major difficulty to use this software. This is an indication of applicability of the suggestion method. However, most of the experts who participated in this study had knowledge about reliability and RAM analysis. That is not always the case when we ask questions from experts in different areas. In fact, the answers of the Expert number 3, who also got the highest global weight, is a vivid indicator of this challenge.

Table 5-1: Experts' opinion regarding to the methodology

Expert	Answers
Expert 1	Q1:7 Q2:7
Expert 2	Q1:8 Q2:8
Expert 3	Q1:5 Q2: "N/A did not use the feedback"
Expert 4	Q1:8 Q2: "I am used to this. I like to see the tails"

Last step in the suggested methodology is to aggregate the judgment of experts with the proper weighing. The proper calculation of weights can be challenging specially if we try to reduce the influence of "chance" factor in our analysis. Moreover, a surprising and unusual opinion from a wise and knowledgeful expert should be investigated by the analyst rather to avoid reducing the weight of that expert's judgment without a clear reason. In this thesis, I used EXCALIBUR to combine the results and this software provides some options to mitigate the factor of chance with some statistical methodologies. Since the main objective of this thesis was to show how digitalization can be used for elicitation, and because of complexity, I decided to not use these options. However, to benefit from these options we need to know our experts and their level of expertise to some degrees.

It is important to mention that there should be a clear purpose and goal behind the idea of digitizing the elicitation process to harness the potential benefits of it. It is a fair statement that by applying digital elicitation process we are trying to achieve a reliable and efficient elicitation process. However, to get the results that we are seeking, we need to have analyst expertise, suitable technology and management commitment which support the action that we are taking. The elicitation process should have a clear goal that analyst establishes and communicates with the management and experts. The questions that experts are supposed to answer should be designed according to this goal and the analyst must make sure that those questions do not introduce biases and provides enough information for experts. When it comes to suitable technology, a proper software which is user friendly and comprehensive improves the efficiency of the elicitation process. The software might be dependent on the internet connection which means a reliable connection for experts is essential. For example, if the experts are working on an offshore platform and they do not have access to the internet then, no matter how sophisticated our software is, the elicitation process gets interrupted. Last point about the suitable technology, but not the least, is the cyber security issue that creates other challenges for the user of the software. Management commitment to support the elicitation process by providing enough training for experts and analysts and suggesting suitable work processes is vital for having a reliable and efficient elicitation process. The management role is also important when it comes to the commitment of the experts to the process. Creating incentives for experts improve their commitment to the elicitation and therefore improve the efficiency of the process.

To show the relationship of digital elicitation, analyst expertise, technology and management commitment with a reliable and efficient elicitation process, I used a critical realist diagram (Pawson and Tilley,1998). Figure 5-1 shows the requirements that we have to take into account if we want to make a digital elicitation process which is reliable and efficient.

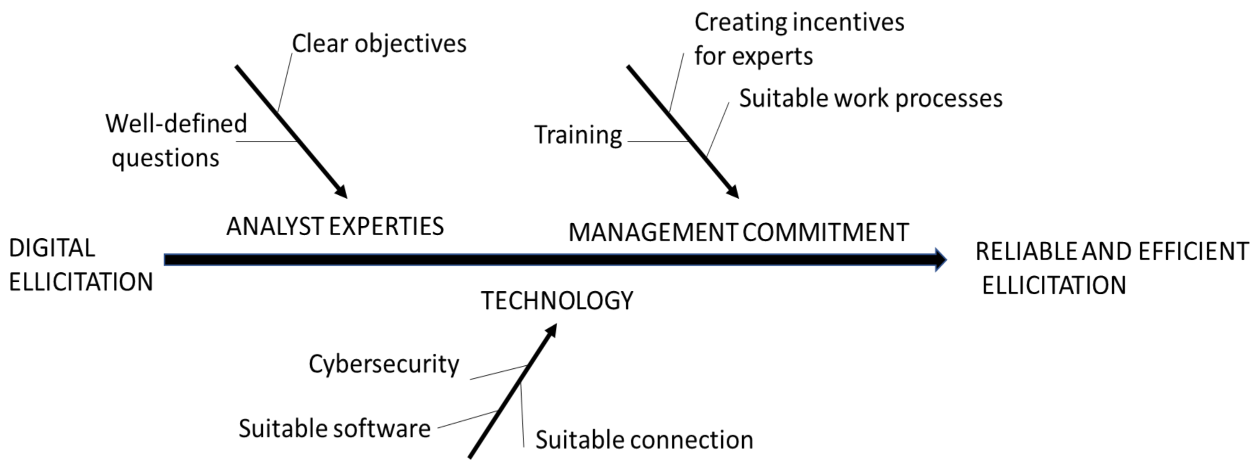


Figure 5-1: Causal relationship between requirements and the goal

References

Avella, J. R. (2016). Delphi panels: Research design, procedures, advantages, and challenges. International Journal of Doctoral Studies, 11, 305-321.

Ayyub, B. (2001). *A Practical Guide on Conducting Expert-Opinion Elicitation of Probabilities and Consequences for Corps Facilities*. IWR Report 01-R-01. Alexandria: Institute for Water Resources.

Baybutt, P. (2017). The validity of engineering judgment and expert opinion in hazard and risk analysis: The influence of cognitive biases. *Process Safety Progress*, 37(2), pp.205-210.

Burgman, M., Fidler, F., McBride, M., Walshe, T. and Wintle, B. (2006). *Eliciting Expert Judgments: Literature Review*. Melbourne,: Australian Centre for Excellence in Risk Analysis, pp.24-25.

Clemen, R.T., Winkler, R.L., 1999. Combining probability distributions from experts in risk analysis. Risk Analysis 19, 187–203.

Cooke, R., 1991. Experts in uncertainty. Opinion and subjective probability in science. Oxford University Press, New York/Oxford.

Cooke, R. and Goossens, L. (2000). Procedures guide for structured expert judgment. Brussels: European Commission.

Cooke, R. and Goossens, L. (2004). Expert judgement elicitation for risk assessments of critical infrastructures. *Journal of Risk Research*, 7(6), pp.643-656.

Diaz Anadon, Laura, Valentina Bosetti, Gabriel Chan, Gregory Nemet, and Elena Verdolini. “Energy Technology Expert Elicitations for Policy: Workshops, Modeling, and Meta-analysis.” Discussion Paper, Belfer Center for Science and Inernational Affairs, Harvard Kennedy School. October 2014.

GE oil and gas (2016). *The Impact of Digital on Unplanned Downtime, AN OFFSHORE OIL AND GAS PERSPECTIVE*. [online] GE oil and gas.

Goossens, L., Cooke, R., Hale, A. and Rodić-Wiersma, L. (2008). Fifteen years of expert judgement at TUDelft. *Safety Science*, 46(2), pp.234-244.

Gore, S. (1987). Biostatistics and the medical research council. Med. Res. Council News 35, 19 20.

Helmer, O., 1968. “Analysis of the Future: The Delphi Method,” and “The Delphi Method An Illustration,” in J. Bright (ed.), Technological Forecasting for Industry and Government, Prentice Hall, Englewood Cliffs, NJ.

J. Charles Kerkering (2002) Eliciting and Analyzing Expert Judgment, A Practical Guide, Technometrics, 44:2, 193-194, DOI: 10.1198/tech.2002.s716

Lee, J. (2016). *Pipe-in-pipe solutions*. [online] Offshore Engineer Magazine. Available at: https://www.oedigital.com/component/k2/item/13163-pipe-in-pipe-solutions [Accessed 18 Apr. 2019].

Leroy, A. (2018). *Production Availability and Reliability: Use in the Oil and Gas industry*. united states: Wiley, pp.163-164.

Liyanage, J.P., Langeland, T. (2009) *Smart assets through digital capabilities, in Khosrow-Pour, M., (ed), Encyclopaedia of Information science and technology*. IGI Global, 3480-3485

Meyer, M.A. & Booker, J.M. (1990). Eliciting and analyzing expert judgment: a practical guide. Office of Nuclear Regulatory Research, Division of Systems Research, US Nuclear Regulatory Commission Washington, DC.

Morris, D., Oakley, J. and Crowe, J. (2014). A web-based tool for eliciting probability distributions from experts. *Environmental Modelling & Software*, 52, pp.1-4.

Naseri, M. (2016) RAM Analysis of Oil and Gas Production Facilities Operating in the Arctic Offshore, A dissertation for the degree of Philosophiae Doctor, The Arctic University of Norway.

Ouchi, F. (2004). *A Literature Review on the Use of Expert Opinion in Probabilistic Risk Analysis*. World Bank Policy Research Working Paper 3201. World Bank.

Pawson, R. and Tilley, N. (1998), Realistic Evaluation, Sage, London.

PIPELINE AND RISER LOSS OF CONTAINMENT 2001 – 2012 (PARLOC 2012). (2015). PARLOC REPORT. ENERGY INSTITUTE, Oil & Gas UK, pp.30 &45.

Rai, V. (2013). Expert elicitation methods for studying technological change under uncertainty. *Environmental Research Letters*, 8(4), p.041003.

Veen, D., Stoel, D., Zondervan-Zwijnenburg, M. and van de Schoot, R. (2017). Proposal for a Five-Step Method to Elicit Expert Judgment. *Frontiers in Psychology*, 8.

Vernon, W. (2009). A Delphi technique: A review. International Journal of Therapy and Rehabilitation, 16(2), 69-76. Retrieved from doi: dx.doi.org/10.12968/ijtr.2009.16.2.38892

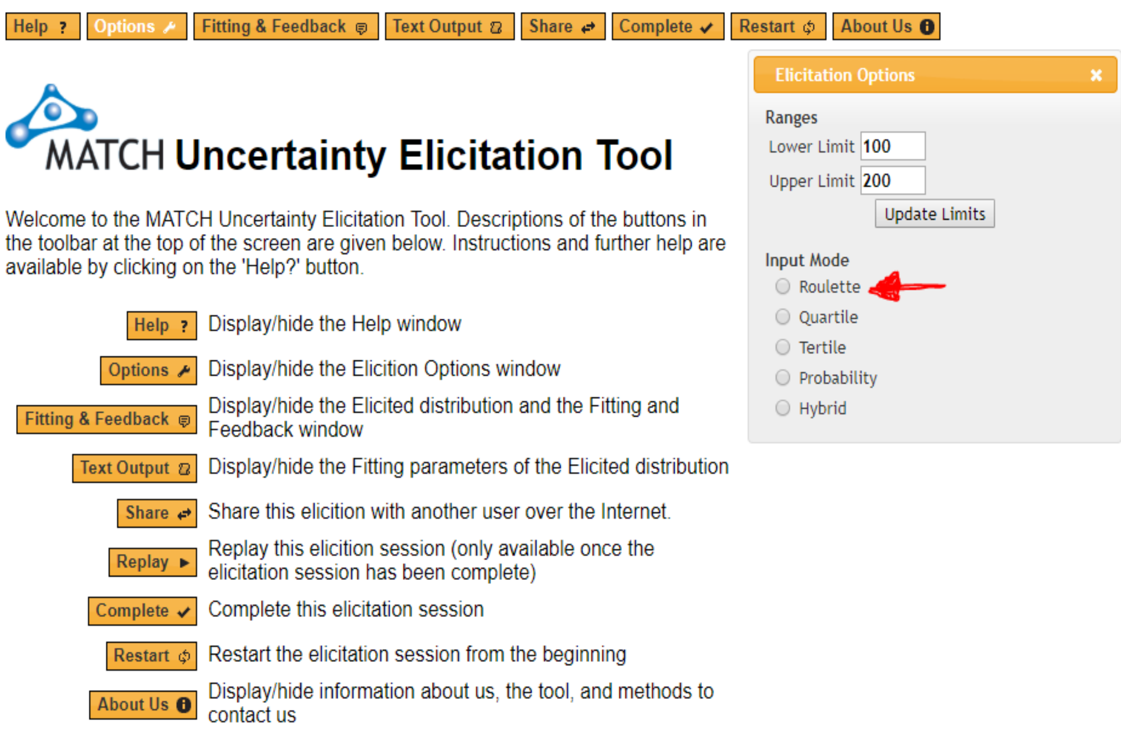
Appendix I

Guideline for using the MATCH software:

Note: Please be aware of cognitive biases that might affect your opinion. Biases such as statistical illusions, over confidence and underestimate, can reduce the accuracy of your judgment. Awareness about biases reduce the chance of falling in their trap.

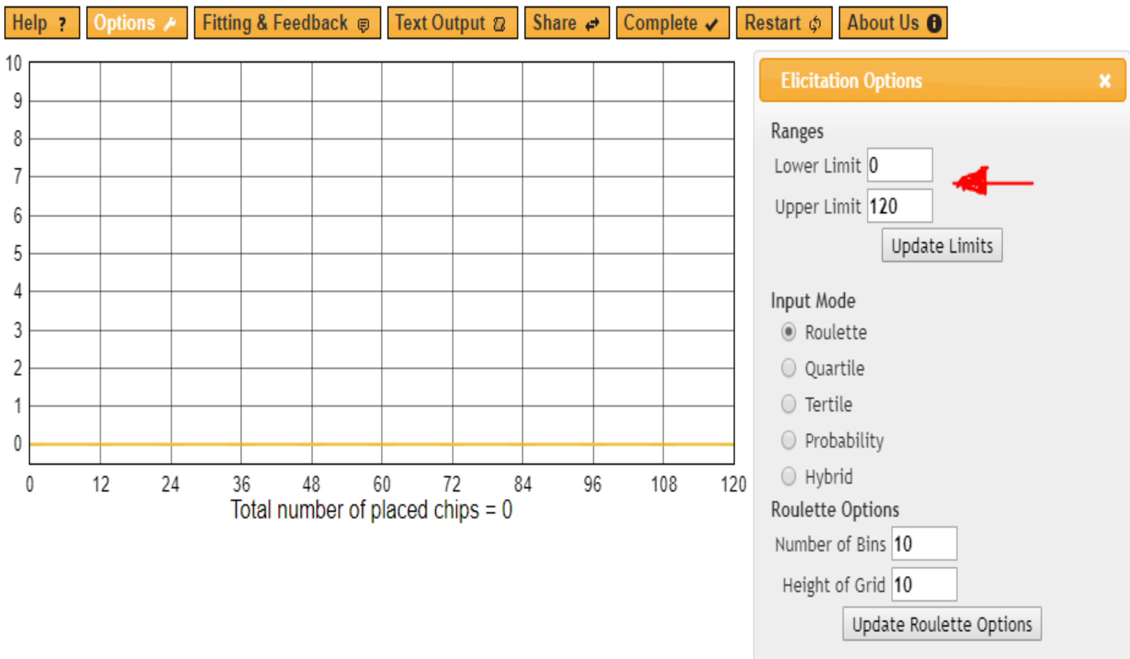
Please check the questions provided in the questioner. To answer each question, you need approximately 2 or 3 minutes.

1. To start the elicitation please read the first question and then go to the following address: http://www.match.ac.uk/uncertainty
2. On the right side of the page, pick Roulette as the input mode.

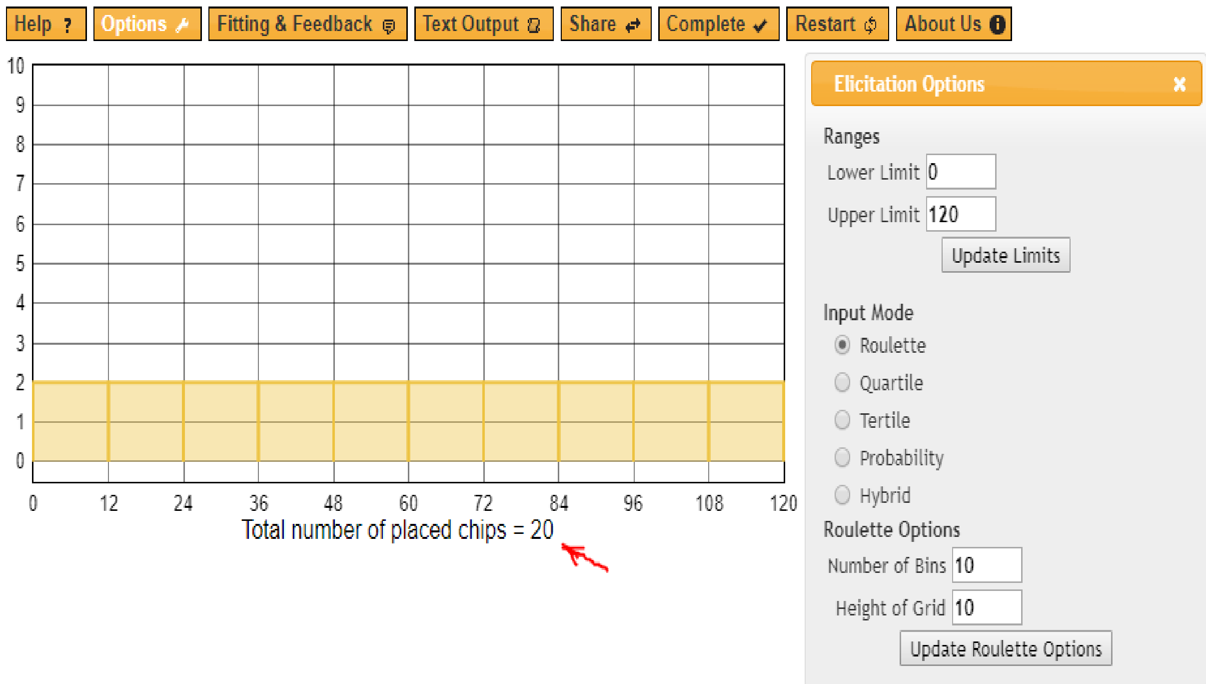


3. Use the horizontal axis to show the number of failures (for the part of questioner) or the time to failure in years (for the second part of questioner). The vertical axis is the number of chips for the first part and the number of pipelines. You can change the range of

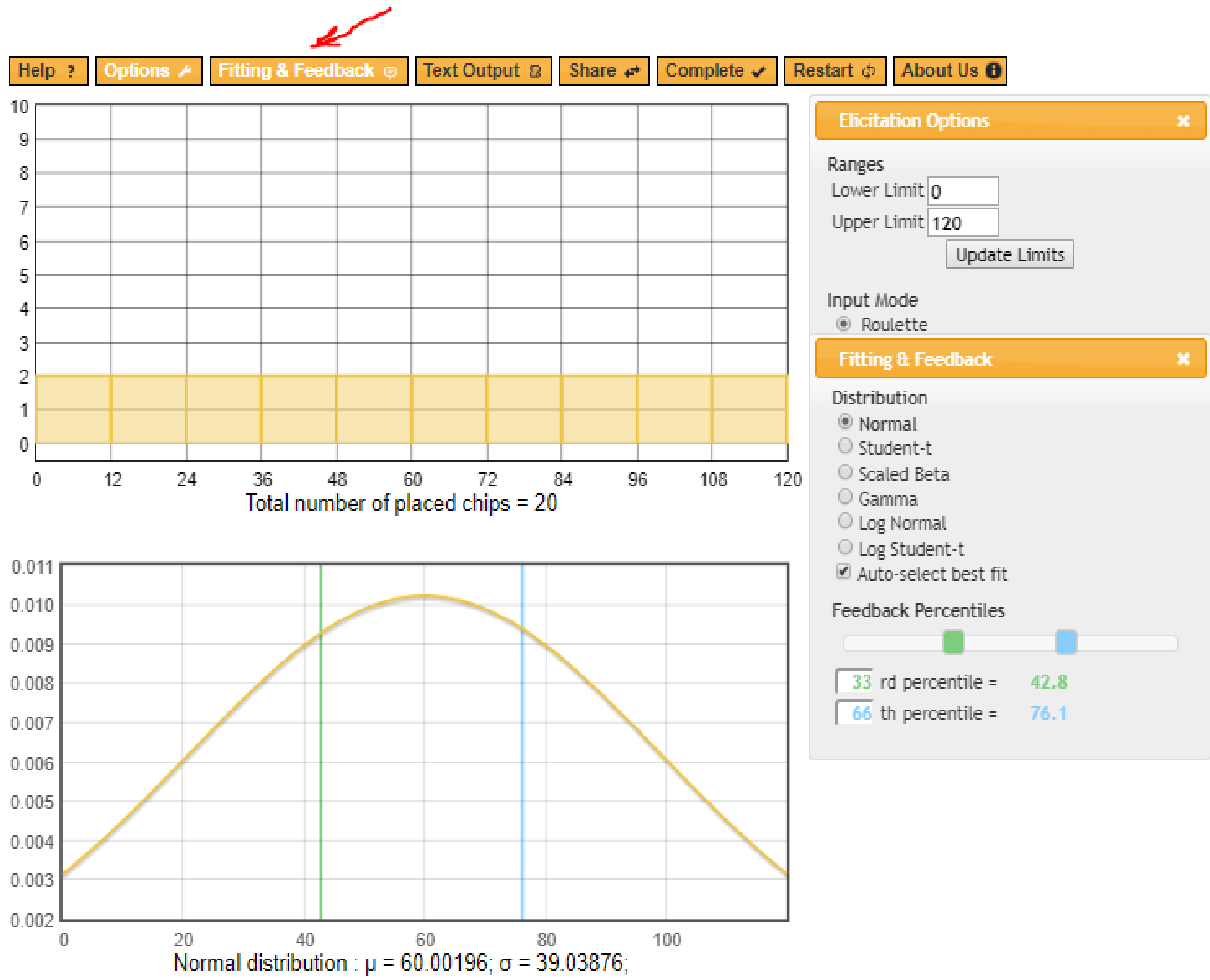
horizontal axis according to your opinion. For example, if I want to express my opinion about failure of heart in humans I will pick the lower limit 0 year and upper limit 120 years.



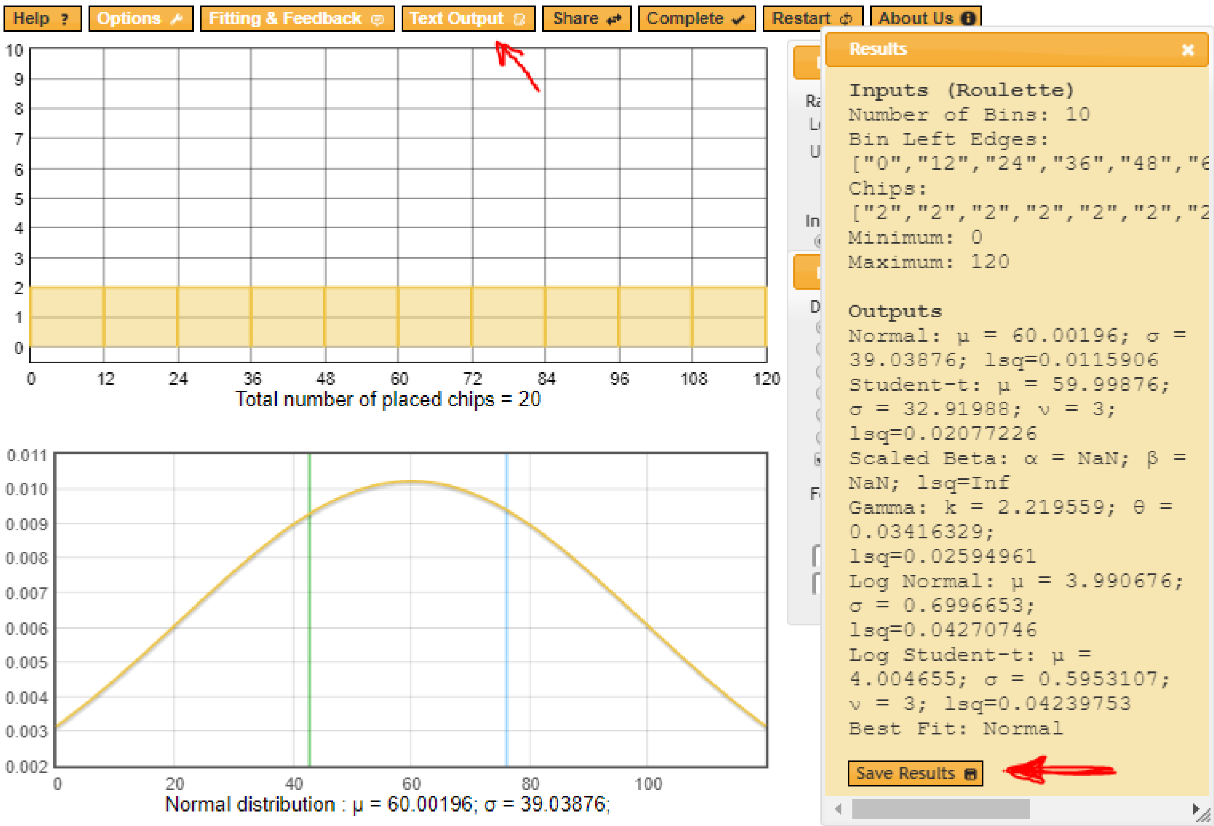
4. By clicking on each cell, you can fill it in and if you click on the cell again you can make it blank again. Please make sure after filling the cells, the total number of placed chips would not be more than 20.



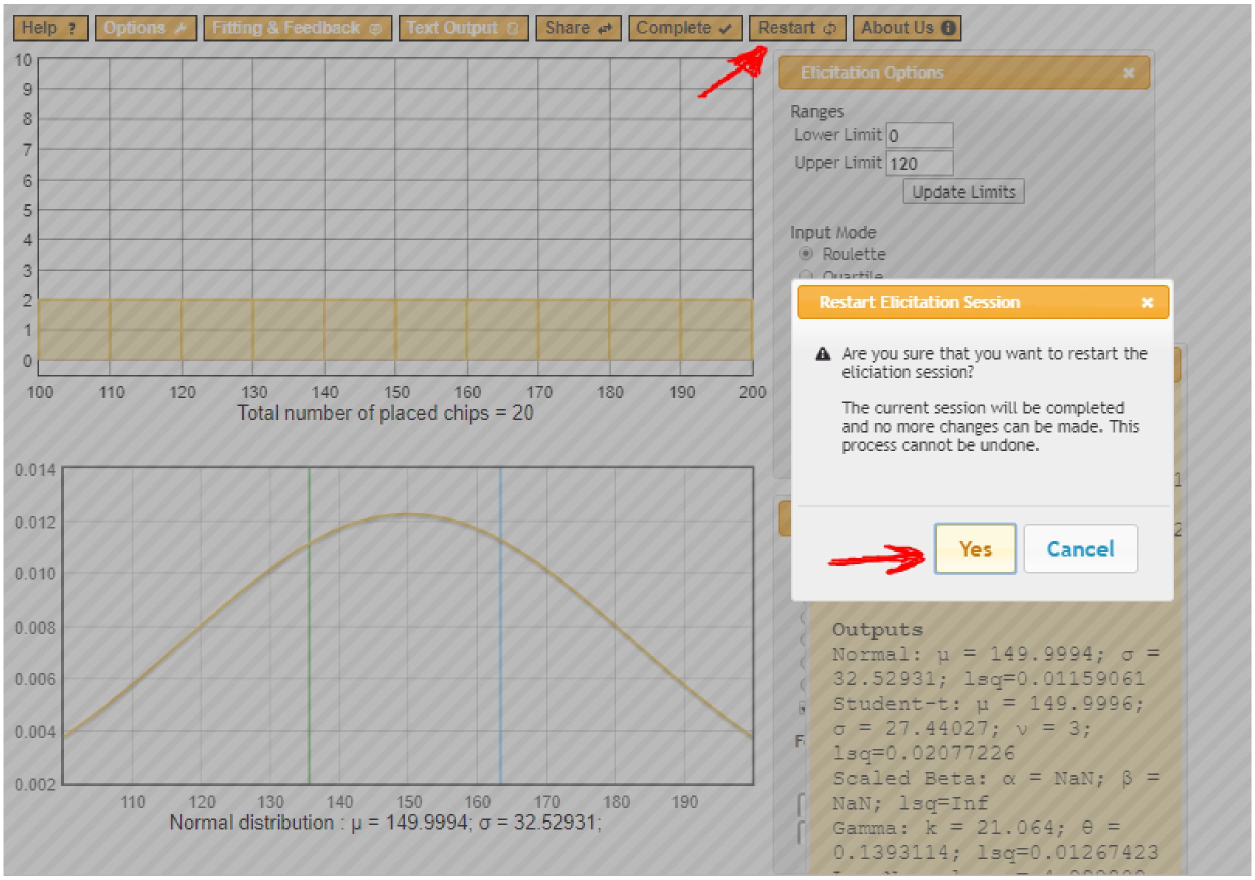
5. After filling the cell, you can go to the "fitting and feedback" page and see the distribution presenting your opinion. If you are not an expert in distributions and their interpretations, please ignore this step. If you need to change your inputs to get a more accurate distribution which represent your opinion you can refill the diagram.



6. Press the “text output” tab and save the results. Please save the results of each question according to the following name format: “Question number-your name”



7. After saving your results, to answer the next question please restart and repeat the process from step 1.



8. After answering all the questions, Please attach your answers for part 1 and 2 (text files) to an email and insert your answers for part 3 questions in the body of email and send it to: Amir.beiky@gmail.com , Arnaud.Barre@woodplc.com,

Appendix II

Questionnaire:

You need less than 15 minutes to answer the following questions.

For the following questions the descriptions of failures are as following:

- *Impact*: failures due to anchoring, trawling and other impacts.
- *Material*: failures due to corrosion (internal or external) and other material causes.
- *Operations and maintenance*: incorrect operation (such as over-pressurization) or error during maintenance or removal (e.g. release resulting from incomplete flushing of the line prior to opening the line).
- *Construction*: failures during commissioning tests or occurring as inadvertent by-product of construction.

Part 1:

We have randomly monitored 1372 rigid steel pipelines at the sea bed for transporting oil and gas in UK continental shelf during past 12 years. The pipelines are connected to the well and the platform.

1. Using the MATCH software, please express your opinion about the number of observed "impact failures in midline part of pipelines resulting loss of containment". To express your opinion, you can use the X-axis as the number of impacts and 20 chips to show your uncertainty (more chips for the numbers that you think are more probable).
2. Using the MATCH software please express your opinion about the number of observed "Material failures in midline part of pipeline resulting loss of containment". To express your opinion, you can use the X-axis as the number of material failures and 20 chips to show your uncertainty.
3. Using the MATCH software please express your opinion about the number of observed "operation and maintenance failures in midline part of pipeline resulting loss of containment". To express your opinion, you can use the X-axis as the number of material and maintenance failures and 20 chips to show your uncertainty.

4. Using the MATCH software please express your opinion about the number of observed “Construction failures in midline part of pipeline resulting loss of containment”. To express your opinion, you can use the X-axis as the number of construction failures and 20 chips to show your uncertainty.

Part 2:

For the following questions assume we have 20 identical Pipe in Pipe pipelines at the sea bed (north sea typical condition) for transporting oil (oil properties, temperature and pressure is constant in all the pipes). These pipelines (midline) are not protected and exposed.

1. Using the MATCH software please express your opinion about the time that it takes until the first “corrosion failure (internal or external) resulting loss of containment” happens in the pipelines (midline). To express your opinion, you can use the X-axis as the time of failure in years.
2. Using the MATCH software please express your opinion about the time that it takes until the first “impact failure resulting loss of containment” happens in the pipelines (midline). To express your opinion, you can use the X-axis as the time of failure in years.

Part 3:

3. On a scale of 1 to 10, how easy did you find answering the questions using the software when 10 means extremely easy.
4. On a scale of 1 to 10, how useful you find the feedback process in the software when 10 means extremely useful.