NTNU – Trondheim
Norwegian University of
Science and Technology

Combined Risk Indicator for Major Accident Precursors and Barriers in the Trends in Risk Level Project

Marius Stensland

Marine Technology
Submission date: May 2013
Supervisor: Jan Erik Vinnem, IMT

Norwegian University of Science and Technology
Department of Marine Technology

Preface

This report is a presentation of the master thesis work performed during the last semester at Marine Technology, NTNU. The scope of the master thesis has been to develop a new risk estimation method for the “trends in risk level” project.

During the master thesis work there have been certain issues with gaining access to the necessary data. As a result the new model has only been tested for one type of major accident precursor. It is however not my impression that this has resulted in a poorer model.

I would like to thank Professor Ingrid Bouwer Utne for guidance on the project thesis work during the autumn semester.
It would not have been possible to test the model without data from Safetec. I would therefore like to thank Principal Engineer Terje Dammen, Safetec, for taking his time to send me the necessary data. I would also like to thank my father, Principal Engineer Morten Stensland, DNV, for always answering my questions and giving valuable feedback on how the offshore oil and gas industry manage technical safety.

At last, but not at least I must thank my supervisor Professor Jan Erik Vinnem for his guidance during this spring. Quick responses, good feedback and good ideas. This master thesis could not have been written without this guidance.

Marius Stensland

Trondheim
21th of May 2013

(Intentionally left blank)

Summary

Since the “trends in risk level” project first was started, the weights used to estimate the risk level on the Norwegian Continental Shelf has more or less been the same; five sets of generic weights, one for each type of facility, and not facility specific and time dependent. The aim of this master thesis has been to develop a new weight estimation model.

The new model estimates time dependent and facility specific weights, based on risk analyses and observed barrier functions’ performances. These weights are combined with occurrences of major accident precursors for each facility and then summed up for the entire Norwegian Continental Shelf. This estimates one risk indicator, combining both types of major accident related data reported in the “trends in risk level” project.

The developed model consists of two methods; method 1 and method 2. Both methods use the facility’s QRA and reported barrier functions’ performances when estimating new weights. There are however some differences between the two methods.

- Method 1 reproduces parts of the risk analysis, but instead of using assumed barrier functions’ performances, observed performance is used. Based on the updated risk analysis, weights are estimated.
- Method 2 adjusts the risk estimated by the original risk analysis by correction factors based on deviations between the observed barrier functions’ performances and benchmark values.

Both methods estimate facility specific weights based on barrier functions’ performances. Method 1 gives more correct weights, is more time-consuming and requires more information from the available data. Method 2 will be less specific to the facility, as it uses general calibration factors established based on other facilities.

The test cases show that method 1 will give good weight estimates and that weights can be estimated by method 2. It can also be seen that the new weight estimates are significantly different from the generic weights and that the estimated weights changes from year to year depending on barrier functions’ performances.

Sammendrag

Siden "Risikonivå i Norsk Petroleumsvirksomhet" prosjektet først ble startet, har vektene som benyttes til å estimere risikonivået på norsk sokkel mer eller mindre vært de samme; fem sett generiske vekter. En for hver innretningstype, ikke innretningsspesifikke og uavhengige av tid. Målet med denne masteroppgaven har vært å utvikle en ny vektestimeringsmodell.

Den nye modellen estimerer tidsavhengige og innretningsspesifikke vekter, basert på risikoanalyser og observerte ytelser av barriere funksjoner. Disse vektene kombineres med tilløpshendelser for hver innretning og deretter summers det opp for hele den norske sokkelen. Dette estimerer en risikoindikator som består av begge typene storulykkerelatert indikatorer rapportert i "Risikonivå i Norsk Petroleumsvirksomhet" prosjekt.

Den utviklede modellen består av to vektestimeringsmetoder; metode 1 og metode 2. Begge metodene benytter innretningenes QRAer og de innrapporterte ytelsene av barriere funksjoner til å estimere nye vekter. Det er imidlertid noen forskjeller mellom de to metodene.

- Metode 1 oppdaterer deler av risikoanalysen, men i stedet for å bruke antatt ytelse av barriere funksjoner, brukes de observerte ytelsene. Basert på den oppdaterte risikoanalysen estimeres nye vekter.
- Metode 2 estimerer vekter fra den originale risikoanalysen og justerer disse ved hjelp av korreksjonsfaktorer. Korreksjonsfaktorene er basert på avvik mellom observert ytelse og benchmark ytelser.

Begge metodene estimerer innretningsspesifikke vekter basert på ytelse av barriere funksjoner, men metode 1 gir mer korrekte vekter, er mer tidkrevende og krever mer av den tilgjengelige dataen. Metode 2 gir mindre innretningsspesifikke vekter på grunn av at den bruker generelle kalibreringsfaktorer etablert basert på andre innretninger.

Test casene viser at metode 1 gir gode vektestimater og metode 2 kan estimere vekter. De viser også at det kan være signifikante avvik mellom estimerte vekter og de generiske vektene og ved å inkludere ytelse av barrier funksjoner vil vekter som varier fra år til år kunne estimeres.

Contents

List of Figures

List of Tables

List of Equation

Acronyms

BDV – Blow Down Valve

BOP – BlowOut Preventer

DFU – Defined hazards and accident situations (Nor: Definerte Fare- og Ulykkessituasjoner)

DHSV – DownHole Safety Valve

ESDV – Emergency ShutDown Valve

FAR – Fatal Accident Rate

GM – Metacentric height

HC – HydroCarbons

HSE (1) – Health and Safety Executive

HSE (2) – Health, Safety and Environment

Kg/s – Kilo Gram per Second, flow rate

KooN – K out of N items, where K is smaller than N

LTI – Lost Time Injury

NCS – Norwegian Continental Shelf

PFD – Probability of Failure on Demand

PLL – Potential Loss of Life

POB – Persons OnBoard

PSA – Petroleum Safety Authority Norway

PSV – Pressure Safety Valve

QRA – Quantitative Risk Analysis

RNNP – Trends in Risk Level (Nor: Risikonivå i Norsk Petroleumsvirksomhet)

1. Introduction

1.1. Background

The oil and gas industry has a high potential for major accidents. To monitor the risk level the Petroleum Safety Authority Norway (PSA) has an annual project called “Trends in risk level” (RNNP). The objective of this project is to measure the impact of Health, Safety and Environmental (HSE) work, identify critical areas for HSE and gain more insight in potential causes of accidents.

To achieve this goal, the RNNP-project analyses annual reports of accident precursors and observed performances of barrier functions from operators on the Norwegian Continental Shelf (NCS). In addition to this, a comprehensive questionary survey is performed amongst everyone working offshore within PSAs legal jurisdiction every second year.

Today’s model estimates a total risk indicator based only on the reported occurrences of major accident precursors. Barrier functions’ performances are analyzed and assessed separately from the incident reports. A more precise risk estimate could be achieved if barrier functions’ performances were included in the total risk indicator.

1.2. Master Thesis Objective and Problem Formulation

To improve the total risk indicator used by the RNNP-project today and to investigate the effect variations in observed barrier functions’ performances have on the estimated risk level, this master thesis will develop a model that combines major accident precursors and observed barrier performances into one indicator.

To examine whether the model can be used, it will be tested on two test cases. This will allow weaknesses, limitations and the potential of the developed model to be identified.

The objective of the master thesis has been as follows:

- Develop a model that combines major accident precursors and observed barrier functions’ performances into one indicator as an alternative to the model used by the RNNP-project today.
- Test if the developed model can be used to estimate the risk level.
- Determine if necessary information and data for the model is available.

1.3. Limitations

Certain limitations have been made in order to make the scope of the master thesis work more manageable. These are as follows.

- Only data from year 2003 to and including 2012 shall be used. This limitation is made because of certain issues with the data quality of the reported barrier functions' performances for year 2002.
- The focus of this master thesis is only major accidents. Furthermore it only focuses on risk to humans and ignores risk to environment and assets. The focus has been narrowed further by only considering DFUs with barrier functions assigned by "Metoderapporten" plus DFU 9. This means that the developed model only focuses on DFU 1, DFU 2, DFU 3, DFU 8 and DFU 9. DFU 10 is considered a precursor to DFU 9, so no weights will be developed for it.

  No weights are developed for DFU 2 as there have been no occurrences of this major accident precursor as long as the RNNP-project has been running.
- The model will be tested on two facilities. Due to lack of data it is only tested on DFU 1.
- For some of the barrier functions assigned to DFU 3 and DFU 8 no benchmark values are available. No effort has been put into establishing benchmark values for these barrier functions, as this not is considered to serve the intended purpose.
- Only offshore facilities are dealt with, no attention is paid to onshore facilities.

1.4. Structure of Report

A presentation of the master thesis problem formulation, objective of the master thesis, background and acronyms are given in the first chapter. The next part consists of general theory, a presentation of the RNNP-project and a description of where the project stands today. After this a description of the new model and testing of it is presented. When the new model has been tested, two discussion chapters follows; the first discusses assumptions, limitations and model development, the second discussion chapter is a general discussion around the developed method, its use, results of the test cases and further work. This is followed by a conclusion. The last part of the report consists of references and appendices.

2. Theory

2.1. Barriers

There are nearly as many interpretations of the term barrier as there is literature on the subject, so no unified definition on the term can be found. In this chapter a brief description of the term barrier will be given.

Some of the definitions found in the literature are as follows.

- "Technical, operational and organizational elements which individually or together shall: a) reduce the possibility of occurrence of specific errors or hazards, or b) reduce or prevent damage if they occur." (PSA, 2013a). This definition corresponds well with the one used by Sklet (Sklet, 2006). The main difference being that Sklet uses the term "safety barrier" and not only "barrier". In this master thesis barriers will only be referred to in relation to safety, so the two terms will be equal.
- "Physical or engineered system or human action (based on specific procedures or administrative controls) that is implemented to prevent, control, or impede energy released from reaching the assets and causing harm" (Rausand, 2011).
- "A barrier is, generally speaking, an obstacle, an obstruction, or a hindrance that may either: (1) prevent an event from taking place, or (2) thwart or lessen the impact of the consequences if it happens nonetheless" (Hollnagel, 2004).

In the consultation paper "Prinsipper for barrierestyring i petroleumsvirksomheten" (PSA, 2013a), PSA attempts to specify what they mean the regulation requires of barriers and what measures that should be in place on facilities under Norwegian jurisdiction. PSA defines barriers according to the energy and barrier model. This model describes the isolation of energy from assets by one or more barriers. See figure 1 for an illustration.

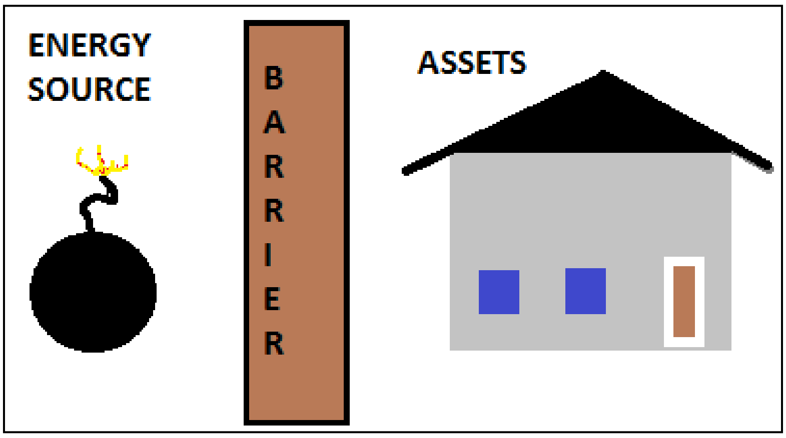


Figure 1 Energy and barrier model

2.1.1. Classification of Barriers

PSA (2013a) classifies barriers into three categories, technical, operational and organizational barriers. A description of these categories can be as follows. A technical barrier element is a feature of the design, implemented to for example contain energy. Organizational barrier elements have to do with personnel and their tasks and role in the execution of certain functions. While operational barrier elements are actions personnel performs.

PSA's classification of barriers can also be characterized as physical and non-physical. Technical barriers will then be physical, while operational and organizational barriers are non-physical (Sklet, 2006).

NORSOK Z-013 defines barrier elements, functions and systems. Respectively as components of the barrier system, functions planned to prevent, control or mitigate undesired or accidental events and systems designed and implemented to perform one or more barrier functions (NORSOK, 2010).

Barriers can be classified into other categories than those given by PSA as well. One way is to look at the different barriers' functions and divide into proactive or reactive barriers. That is barriers that prevent a hazardous event from occurring or barriers that stop or mitigates consequences after an unwanted event has occurred. Another classification is active and passive barriers. An active barrier is a barrier that for example needs an operator, a control system or some sort energy to function. A passive barrier on the other hand is built into the design and does not need anyone or anything in order to be activated or to perform its function.

Sometimes one barrier is not sufficient to stop an unwanted event from occurring. When more than one barrier is needed, they can be classified as partial barriers. If it is enough with one barrier, it is classified as a full barrier.

Barriers can also be classified according to at which "depth" or which protection layer they are assigned in the defence in depth strategy. The first barrier activated is the primary barrier, followed by the secondary barrier and so on until all defence layers has been classified (Rausand, 2011).

Hollnagel (2004) suggest a four group classification of barrier systems. The four groups are physical or material, functional, symbolic and incorporeal barrier systems. This classification scheme uses the nature of the barrier as a basis for classification and does not say anything about how or when the barrier functions are activated.

2.1.2. Barrier Performance Standard

Barrier performance can be measured. This can be done in several ways. PSA (2013a) specifies that performance can include capacity, reliability, accessibility, efficiency, ability to withstand loads, integrity and robustness. And that it is not essential what way barriers are categorized, but that the performance standards are related directly to the barrier element they describe.

The management regulation's article number 5 (PSA, 2011a) demands that personnel shall be aware of defined performance requirements for all barriers.

NORSOK S-001 (NORSOK, 2008) states that the safety performance standards, shall ensure that barriers are suitable, effective, have sufficient capacity, availability and adequate response time for all identified hazards and are suitable for all relevant operating conditions.

2.2. Indicators

2.2.1. Types of Risk Indicators

This master thesis use a categorization of risk indicators based on the nature of the barrier it reflects. The categories used are technical, operational and organizational indicators. Different categorizations are used by other sources, but the three given categories will be used as they are the same as described by PSA (2013a).

2.2.1.1. Technical Risk Indicators

A technical risk indicator says something about the condition of a technical barrier. They are also called "direct risk indicators" or "physical risk indicators" since they often measure a physical property directly on the monitored subject. Technical risk indicators are often easy to measure and connect to a risk model. This makes them well suited for prediction of future risk level (Øien, 2001b). An example of a technical risk indicator can be rate of failed gas detectors during testing.

2.2.1.2. Organizational and Operational Indicators

Many sources do not distinguish between organizational and operational risk indicators as they have several properties in common (Sklet, 2006). The most obvious similarity is that they both can be characterized as non-physical risk indicators.

The following distinction between organizational and operational indicators will be based on PSA's (2013a) distinction between types of barriers. Organizational risk indicators attempt to predict the performance of organizational barriers. I.e. how well personnel fulfill their tasks and roles. Operational risk indicators will attempt to predict performance of the actions personnel performs. I.e. how well operational barriers perform.

2.2.2. Lagging versus Leading Inidicators

The difference between lagging and leading indicators are essential for the way indicators are used. Some examples found in the literature are as follows:

- "Event-based indicators are lagging indicators which reflect experience in the past.", and "Leading indicators are proactive indicators,…" (Aven and Vinnem, 2007).
- "The lagging indicator reveals failings or 'holes' in that barrier discovered **following an incident or adverse event.** The incident does not necessarily have to result in injury or environmental damage and can be a near miss, precursor event or undesired outcome attributable to a failing in that risk control system." And "Leading indicators are a form of active monitoring focused on a few critical risk control systems to ensure their continued effectiveness. **Leading indicators require a routine systematic check that key actions or activities are undertaken as intended.** They can be considered as measures of process or inputs essential to deliver the desired safety outcome." (Health and Safety Executive, 2006).

Health and Safety Executive's (HSE) definition of lagging indicator points out that it reveals failure *after* the occurrence of a barrier failure. Lagging indicators will therefore say something about what has happened, i.e. the past. This can be used to avoid the same situation appearing again. Leading indicators on the other hand actively monitors critical risk control systems *predicting* if they will be able to perform when they are needed. I.e. they try to predict the future.

This master thesis will use the following definitions:

- **Lagging Indicator:**

  "A lagging indicator measures how well, one or more, barriers have performed after an incident or an accident has occurred."
- **Leading Indicator:**

  "A leading indicator will be based on observations or measurements done during daily operation, and attempt to predict how well, one or more, barriers will perform if they are needed in the future. I.e. state of the barrier."

In a similar way as there is a distinction between lagging and leading indicators, there is often a distinction between safety and risk indicators. A safety indicator will usually express how safe a system *has* been (Rausand, 2011). I.e. safety indicators are lagging indicators. Risk indicators are mostly indicators connected to the risk through a risk model and used to predict the *future* safety performance of a system (Øien and Sklet, 2001). I.e. risk indicators are leading indicators.

2.2.3. Dual Assurance

The challenge with using only lagging indicators is that an unwanted event has to occur before there is an indication of failure in one or more of the barrier functions. For offshore oil and gas facilities with the potential to cause major accidents, learning only from mistakes and failures are not an accepted strategy for risk management (Khan et al., 2010). To facilitate early warning, HSE (2006) argues for the use of both lagging and leading indicators. They call it dual assurance.

The benefit of using dual assurance is the possibility of having a continuous monitoring of indicator performance. If the lagging and leading indicator score is significantly different, it should be interpreted as the indicator system needing to be revised. Diagnosis of dual assurance system can be done as described in table 1.

Table 1 Comparison of leading and lagging indicator performance, dual assurance

Leading / Lagging	Good score	Bad score
Good score	Leading and lagging indicators correspond	Leading indicators do not represent corresponding barrier in a good manner
Bad score	One or more barriers no longer functioning as intended	Leading and lagging indicators correspond

2.3. Risk Measurement

The RNNP-project expresses risk both as Potential Loss of Life (PLL) and Fatal Accident Rate (FAR) values (PSA, 2011b). In the following a short presentation of the two methods will be given.

2.3.1. Potential Loss of Life

PLL expresses expected fatalities per year (PSA, 2011b, page 24). A more precise definition is “PLL is the expected number of fatalities within a specified population (or, within a specified area A) per annum” (Rausand, 2011, page 602). For the RNNP-project the area is the NCS and the population is everyone spending time on these facilities. PLL is independent of time and personnel exposed to the risk. This enables PLL for several activities to be added together without any weighing being necessary (Rausand and Utne, 2009).

PLL can be expressed as in equation 1.

$$PLL_{tot} = \sum_{a=1}^{m} PLL_a = \sum_{a=1}^{m}\sum_{i=1}^{n} IRPA_{a,i}$$

Equation 1

Where PLL_a is expected PLL for activity a, $IRPA_{a,i}$ is Individual Risk Per Annum for activity a and person i (Rausand, 2011, Rausand and Utne, 2009).

2.3.2. Fatal Accident Rate

The FAR value is the expected number of fatalities in accidents during 100 millions exposed hours of activity (Rausand, 2011). FAR depends on time exposed to a risk, FAR values calculated for several activities will therefore have to be weighted according to time exposed in order to find a total FAR. I.e. in order to calculate FAR for an offshore worker, the flight out, work in process areas, work in accommodation area and so on has to be weighted according to time exposed (Rausand and Utne, 2009). FAR can be expressed as (Rausand, 2011, page 91, equation 4.9):

$$FAR = \frac{Expected\ number\ of\ fatalities}{Number\ of\ hours\ exposed\ to\ risk} * 10^8$$

Equation 2

Or as a function of PLL (Rausand and Utne, 2009, page 59, equation 4.9):

$$FAR = \frac{PLL}{\sum_{i=1}^{n} t_i} * 10^8$$

Equation 3

Where n is total people exposed and t_i is the time person i is exposed to the risk

2.4. Statistical formulas

This chapter will give a brief presentation of formulas used when processing the reported data and estimating new weights.

2.4.1. Basic Statistical Formulas

This chapter is based on the book “Probability & Statistics for Engineers and Scientist, 9th edition” (Walpole et al., 2012) and all equations, unless stated otherwise, can be found there. In the following X represents the sample space, and consist of [x_1, x_2, ...x_{n-1}, x_n].

2.4.1.1. Sample Mean Value

Sample mean or arithmetic average can be expressed as (Walpole et al., 2012, page 11, definition 1.1):

$$A = \frac{1}{n} * \sum_{i=1}^{n} x_i$$

Equation 4

2.4.1.2. Expected Value

The expected value, *E(X)*, can be expressed as (Walpole et al., 2012, page 112, definition 4.1):

$$E(X) = \sum_{i=1}^{n} x_i * f(x_i)$$

Equation 5

Where $f(x_i)$ is the probability of outcome x_i (Walpole et al., 2012).

E(X) expresses the result or outcome that can be expected, based on all outcomes and their corresponding probability or frequency of appearance.

2.4.1.3. Variance

Variance, *Var(X)*, can be expressed as (Walpole et al., 2012, page 120, definition 4.3):

$$Var(X) = E((X - E(X))^2)$$

Equation 6

Or as (Walpole et al., 2012, page 121, theorme 4.2):

$$Var(X) = E(X^2) - (E(X))^2$$

Equation 7

Variance expresses the distribution's dispersion about the sample mean. When *Var(X)* is small it means that the distribution has little variability from the sample mean. Correspondingly, when *Var(X)* is large it means that the variability about the sample mean is high.

In order to judge if the variance is high or not, it is not sufficient to only calculate the variance. It must also be compared with the input values. That is, *Var(X)* must be evaluated against the distribution of *X*. This will typically be done by looking at the standard deviation. The standard deviation is calculated as the positive square root of the variance (Walpole et al., 2012).

2.4.1.4. Covariance and Correlation

Covariance, *Cov(X,Y)*, can be expressed as (Walpole et al., 2012, page 124, theorem 4.4):

$$Cov(X,Y) = E(X * Y) - E(X) * E(Y)$$

Equation 8

Where *Y* is a sample space with joint probability distribution as *X* given by *f(x,y)*.

The covariance measures the association between *X* and *Y*. When *Cov(X,Y)* is positive, *X* and *Y* varies positively with each other. When *Cov(X,Y)* is negative, *X* and *Y* vary negatively with each other.

Covariance is not a scale free parameter; it does not indicate anything about the strength of association between two variables, i.e. how much variable *X* and *Y* relate to each other. The correlation indicates the strength of the relationship between two variables. The sample correlation coefficient, also called the Pearson product-moment correlation coefficient, is expressed as (Walpole et al., 2012, page 125, definition 4.5):

$$\rho_{xy} = \frac{Cov(X,Y)}{\sqrt{Var(X) * Var(y)}}$$

Equation 9

Where ρ_{XY} is the correlation coefficient.

The correlation coefficient will always have a value between -1 and +1. When the correlation coefficient is smaller than zero, there is a negative relationship between *X* and *Y*. When it is close or equal to zero there is no correlation. And when it is larger than zero there is a positive relationship between *X* and *Y* (Walpole et al., 2012).

To assess the significance of correlation, a critical value for the correlation coefficient is established. For this master thesis, a significance level of 90% is chosen to match the significance level already used by the RNNP-project (Oljedirektoratet, 2001). Critical values for the correlation coefficient are taken from statistical tables based on a two tailed *t*-distribution. In this master thesis the critical values have been taken from Ayyub and McCuen (2003, page 608, table A-5). When the critical value has been established, the absolute values of the calculated ρ_{XY} are compared with it. Whenever ρ_{XY} are found to be larger than the critical value, the correlation can be said to be significant.

2.4.2. Exponential Distribution

The exponential distribution is a one parameter distribution. It can be expressed as (Rausand and Høyland, 2004, page 26, equation 2.27):

$$f(t) = \lambda * e^{-\lambda * t}$$

Equation 10

Where λ is the constant failure rate, and both λ and t is larger than zero (Rausand and Høyland, 2004).

2.4.3. Probability of Failure on Demand

The Probability of Failure on Demand (PFD) expresses the long run average probability that an item not will work on demanded. It can be expressed as (Rausand and Høyland, 2004, page 427, equation 10.3):

$$PFD = 1 - \frac{1}{\tau} * \int_0^{\tau} R(t)dt$$

Equation 11

Where τ equals the time between each function test and *R(t)* is a cumulative probability distribution of the failure rate (Rausand and Høyland, 2004).

2.4.4. Significance

To asses if changes in a dataset are statistically significant, confidence intervals are used. Confidence intervals are established based on the probability distribution used. The RNNP-project uses a significance level of 90%. When assessing significance, this master thesis will use the same level. As the expression for confidence interval depends on the failure distribution used, no formulas will be given here.

3. “Trends in Risk Level”

3.1. General

The RNNP-project is divided in annual phases, 2012 being the 13th phase. Every end of April the last year’s risk level and trends for the NCS is presented. The objective of the project is to measure the effect of HSE work, identify critical areas for HSE and to gain more insight in what causes accidents (PSA, 2012a). Through this work PSA focuses on the entire oil and gas industry as a whole and not just at specific facilities.

The RNNP-reports can be divided into three main areas; risk indicators based on incidents, risk indicators for barriers related to major accidents and a questionary survey performed every other year.

3.1.1. History

In year 2000 the petroleum directory started the RNNP-project. The first year was a pilot project and had a limited scope. The pilot project’s objective was much the same as it is today, but it was also aimed at development, testing, adjusting and calibration of a model, suitable for assessing changes in the risk level on the NCS. The project was further supposed to identify any changes necessary for making the pilot project into an annual project (Oljedirektoratet, 2001).

Phase 2 of the RNNP-project continued most of the scope from the pilot project. It further included work on expanding the scope to relevant ship traffic outside of the safety zone, establish indicators for noise and chemical working environment, helicopter transport to and from the NCS and developing the model for major accident to include the performance of barriers. The two latter were included in phase 3 and indicators for noise and chemical working environment were implemented in phase 4. Including ship traffic outside of the safety zone was found to be too difficult and was left out. Phase 4 also included the creation of an overall barrier indicator, this indicator was left out of the project after phase 8 (Oljedirektoratet, 2002, 2003, PSA, 2004, 2008, 2009).

In phase 5, fieldwork was performed on three facilities that had performed well on HSE elements in the survey performed in phase 4 and showed decreasing amount of HSE critical incidents. The aim of this field work was to identify why some facilities had more improvements from HSE work than others. Phase 5 also included modification of the indicator for DFU 5 – Vessel on collision course, and DFU 1 and 2 – HydroCarbon (HC) leaks (PSA, 2005).

Work was started in phase 6 on how to include onshore facilities under PSA's legal jurisdiction. This work was implemented in phase 7 and is presented in a separate report. Phase 6 also included a series of workshops aiming at identifying if and how organizational, technological and social/cultural measures and their interaction affect risk. But also aiming at contributing to exchange of experience and suggest improvements. In phase 7, fieldwork was conducted on two facilities, with the focus area well service (PSA, 2006, 2007).
A study of risk exposed groups, comparing offshore and onshore locations was conducted and one of the indicators for chemical working environment was left out due to misunderstandings around reporting criteria in phase 8 (PSA, 2008). In phase 9 improved indicators were implemented.
Depth interviews were conducted again in phase 9. This type of interviews had not been performed since phase 4 and PSA wanted to see if there were any significant changes. A pilot study on well integrity indicators was performed and DFU5 was modified again (PSA, 2009). The well integrity indicator was further developed in phase 10 and fully implemented in phase 11.
In phase 10 indicators related to maintenance and ergonomics was implemented, these indicators received some modifications the year after. Changes were made to indicators for DFU 12 – Helicopter transport and it was conducted a study of the external conditions' influence on HSE (PSA, 2010c). The perhaps biggest change in phase 10 is the establishment of a separate report concerning risk related to acute emissions (PSA, 2010b). This effort has been continued in phase 11 and 12 (PSA, 2012a). In phase 11, a qualitative study analyzing causal relation and identifying measures in relation to HC leaks on NCS was performed (PSA, 2011b). This qualitative work was continued in phase 12 (PSA, 2012a).
Continuous improvement of performance data for barriers related to major accidents, noise indicators, chemical working environment, onshore facilities and acute risk to the environment has been a general aim for all phases after they were included in the project.

3.1.2. Categorization of Facility Types

The model used today categorizes offshore facilities into 5 categories. A description of the categorization can be found in table 2.
This categorization covers all facilities on the NCS, but it does not include any field related vessels.

Table 2 Categorization of facilities (PSA, 2011, page 30, table 10)

Facility type	Includes
Fixed installations	Gravity based structures and jacket based facilities
Floating production unit	All floating production facilities, including tensions-leg platform and storage vessels
Complexes	Two or more facilities connected together with bridges, except for fixed installations with flotel and tender vessels
Normally unmanned facilities	Also includes loading buoys
Mobile installations	For drilling and flotels

3.2. Defined Hazards and Accident Situations

In order to measure the safety level, PSA use what they call DFUs. For each DFU the operators report last year's incidents. These incident reports are used to establish risk indicators. The RNNP-project defines the following DFUs (PSA, 2012b).

- DFU 1. Non-ignited hydrocarbon leaks
- DFU 2. Ignited hydrocarbon leaks
- DFU 3. Well events/loss of well control
- DFU 4. Fire/explosion in other areas, flammable liquids
- DFU 5. Vessel on collision course
- DFU 6. Drifting object
- DFU 7. Collision with field-related vessel/installation/shuttle tanker
- DFU 8. Structural damage to platform/stability/anchoring/positioning failure
- DFU 9. Leaking from subsea production systems/pipelines/risers/flowlines/loading buoys/loading hoses
- DFU 10. Damage to subsea production equipment/pipeline systems/diving equipment caused by fishing gear
- DFU 11. Evacuation (precautionary/emergency evacuation)
- DFU 12. Helicopter crash/emergency landing on/near installation
- DFU 13. Man overboard
- DFU 14. Personal injury
- DFU 15. Occupational illness
- DFU 16. Total power failure

DFU 18. Diving accident

DFU 19. H2S emissions

DFU 21. Falling object

The first twelve DFUs are considered major accident precursors. DFU 14 and 18 are used to express statistical risk for occupational and diving accident. DFU 15 gives statistical risk for occupational illness. Whilst DFU 13, 16, 19 and 21 are called other condition affecting risk. DFU 17 – Control room out of operation and DFU 20 – Lost control of radioactive source are no longer included in the RNNP-project (PSA, 2006, page 171)

3.3. Barrier Function Performance

Performance are reported for the following barrier functions (PSA, 2012b).

1) Fire detection
2) Gas detection
3) Shutdown
    a) Riser Emergency ShutDown Valve (ESDV)
        i) Closure test
        ii) Leak test
    b) Wing and master valve (Christmas tree)
        i) Closure test
        ii) Leak test
    c) DownHole Safety Valve (DHSV)
4) Pressure relief valve, Blow Down Valve(BDV)
5) Safety valve, Pressure Safety Valve (PSV)
6) Isolation with BlowOut Preventer (BOP)
7) Active fire prevention
    a) Deluge valve
    b) Start-up test
8) Well integrity
9) Maritime systems
    a) Ballast system valves
    b) Watertight doors
    c) Deck height for jack-up facilities
    d) (Metacentric high) GM values for floaters at the end of the year
10) Mustering time

11) Maintenance management
   a) Tags and HSE critical equipment
   b) Preventive and corrective maintenance, both total time and outstanding

Description of which DFUs the different barrier functions are assigned to can be found in "Metoderapporten", table 9 (PSA, 2011b, page 22 and 23), except for barrier indicator 8 – Well integrity, which describes DFU 3 – Well events/loss of well control. The allocation of barrier functions is reproduced in appendix b.

4. Model development

4.1. Today’s Model

4.1.1. Generic Weights and Total Risk Indicator for Major Accidents

This chapter will describe how the total risk indicator for major accidents is estimated today based on reported incidents for DFU 1 to DFU 10.

Today’s model divides all facilities on the NCS into five categories. For each of the five categories, generic weights have been established. When an incident occurs, the potential contribution to the total risk is estimated by these weights.

The generic weights are mainly based on Quantitative Risk Analyses (QRA) for a selection of facilities in each category. The use of generic weights means that the contribution to the risk level estimated after an incident not will be specific to the facility. I.e. it will not take into account the barrier functions actually in place for that specific facility. It will only use what can be defined as general barrier functions for the facility type.

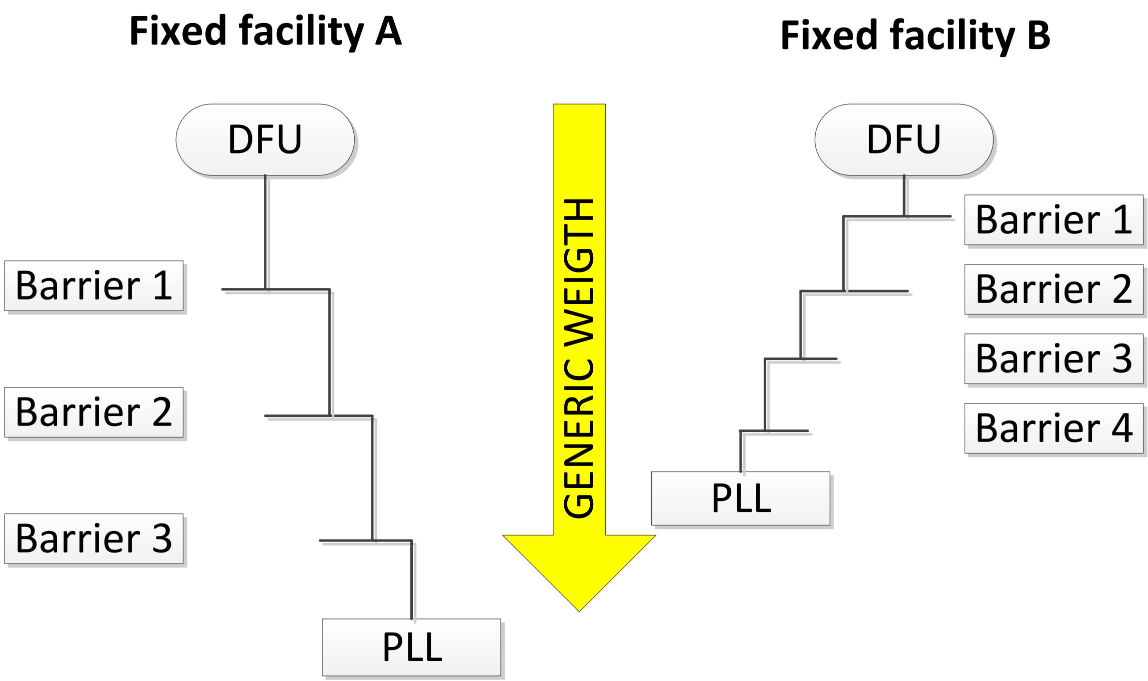


Figure 2 Connecting DFU directly to PLL, two different facilities

Figure 2 is a development of figure 7.2 from “Risk Management – With Application from the Offshore Petroleum Industry” (Aven and Vinnem, 2007). The yellow arrow demonstrates the principle used today. It is the generic weight for a given DFU, based only on facility type. The arrow demonstrates the independence between the generic weight and the barrier functions specific to the facility. The figure shows how the same generic weight is used to estimate the risk from an incident for two facilities of the same type, but with different barrier functions in place. The performance of the barrier functions is illustrated by the length between the barriers. For "Fixed facility A" the weight is too small and for "Fixed facility B" the generic weight is too large.

To calculate the total risk level, *R*, the following formula is presented in "Metoderapporten" (PSA, 2011b, page 24):

$$R = \sum_i \sum_j DFU_{ij} * v_{ij}$$

Equation 12

Where DFU_{ij} equals number of incidents of DFU i for facility type j and v_{ij} equals the generic weight of DFU i for facility type j.

Some of the generic weights are not established based on QRAs from the exciting facilities. DFU 7 and DFU 8 are established from a more general risk analysis method. This method will not be described any further here and it suffices to say that the weights estimated by it are generic and established for each of the five facility categories.

For some types of major accident precursors no generic weights are estimated. The RNNP-project simply states that they will be individually assessed if they occur.

R as expressed in equation 12 will give a total risk indicator expressed as PLL. The total risk indicator is however presented as a FAR value. This is achieved by normalizing the total risk indicator against hours worked (PSA, 2011b).

When using equation 12 to express the total risk indicator, there is no need to look at each facility separately. Just this, the simplicity of the method, is its main benefit. But as showed in figure 2, there are severe limitations that follow from it as well.

4.1.2. Barrier Functions Related to Major Accident

Major accident critical barrier functions' performances are measured and reported as a part of the RNNP-project. An overview of which barrier functions performance is reported for can be found in chapter 3.3.

A total barrier indicator was established as a part of the RNNP-project and included from phase 4 to and including phase 8. Professor J. E. Vinnem informed on a meeting 12th of February 2013 that it was left out of the RNNP-project as it did not communicate the desired message sufficiently well (PSA, 2004, 2007, 2008). Today's RNNP-model does not give an indicator for total barrier performance, but present performance for the separate barrier functions.

Indicators for barrier functions' performances are expressed in several ways. For most of the indicators it is presented as share of barrier function failed during testing. This presentation is used for barrier function 1 to 7 and 9 a) and b) in chapter 3.3. For barrier function 8 – well integrity, a four color classification scheme is used. Barrier function 9 c) and d) are given as metric values. Barrier function 10 is presented as total amount of musters performed and how many musters performed within the facility's specified time limit. Barrier function 11 is measured in hours (PSA, 2011b, 2012a).

To give an indication on how good the observed barrier functions' performances are, some of the barrier functions are compared with what can be considered as acceptable performances. This applies to barrier function 1, 2, 3 a) to c), 4, 5 and 7 a) and b), which are compared with Statoil's internal guidelines.

4.2. Performance of Barrier Functions

As previously mentioned, today's total risk indicator for major accidents does not include barrier functions' performances. The objective of a new model is to include barrier functions' performances when estimating the risk level. This will be achieved by estimating facility specific weights that includes performance of the facility's barrier functions.

Barrier functions' performances are reported on an annual basis. This allows for a time depended risk level estimate to be established.

4.2.1. Available Data

Reporting of observed barrier functions' performance was incorporated into the RNNP-project in phase 3. The first year of performance reporting resulted in certain data quality issues (PSA, 2004). These issues are however avoided as the limitations of this master thesis states that only data for year 2003 to and including year 2012 is to be used. The data quality of the reported performances for these years is considered to be sufficiently good. For further discussion of data quality, see chapter 6.4.1.1.

An important criteria when selecting what barrier functions performance was to be measured for was that it not would create too much extra work for the operators (Oljedirektoratet, 2003). This means that priority was given to barrier functions that are similar for most facilities on the NCS. Only reporting barrier functions' performances common for most facilities leads to certain limitations when the facility specific weights are estimated. Some of the barrier functions actually in place on the specific facility can therefore not be updated with observed performances based on the reported data. The model developed in this master thesis, will have to take this into consideration.

4.2.2. Failure of Barrier Functions

Exponential distribution will be used to describe failure of barrier functions. In the following some elements of failure estimation is described. Arguments for choice of methods can be found in chapter 6.1.

Failure Rate

For the exponential distribution, failure rate can be estimated in several ways. In the following a presentation of the identified alternatives is given.

1. Annual failure rate. For year *i*:

$$\frac{Failed\ tests_i}{Total\ amount\ of\ tests_i}$$

Equation 13

2. Average failure rate. For year *m* to *n*:

$$\frac{\sum_{i=m}^{n} Annual\ failure\ rate_i}{n-m}$$

Equation 14

3. Total failure rate. For year *m* to *n*:

$$\frac{\sum_{i=m}^{n} Failed\ tests_i}{\sum_{i=m}^{n} Amount\ of\ tests_i}$$

Equation 15

4. Three year rolling average. For year *i*:

$$\frac{\sum_{i=1}^{3} Failed\ tests_{i-2}}{\sum_{i=1}^{3} Total\ amount\ of\ tests_{i-2}}$$

Equation 16

Failure rate 2 and 3 are based on equations from "Metoderapporten" (PSA, 2011b, page 25).
Failure rate number 4, three year rolling average, will be used in this master thesis.
As a failure rate of zero not is considered as reasonable, minimum failures are defined as a half failure. I.e. if there is no failure for the three year rolling average, a half failure will be added before the failure rate is estimated.
Discussion and arguments for choice of failure rate estimation can be found in chapter 6.1.2.

Finding Probability of Failure on Demand for Barrier Elements

An important part of using data from barrier functions' performances consists of being able to estimate the availability of barrier elements based on their observed performances. This can be expressed by the PFD.

The formula for PFD is given in equation 11 and expresses the long run average probability of failure on demand, which is the average probability of having a failure on demand once between two inspections (Rausand and Høyland, 2004).

It is assumed that the time between inspections, τ, is one year. With exponential distribution and failure rate given by a three year rolling average, equation 11 can be expressed as follows.

$$PFD_{barrier\ element} = 1 - \frac{1}{1} * \int_0^1 e^{-\left(\frac{\sum_{i=1}^{3} Failed\ tests_{i-2}}{\sum_{i=1}^{3} Total\ amount\ of\ tests_{i-2}} * t\right)} dt$$

Equation 17

Equation 17 will be used to find PFD for barrier functions when using method 1 to establish facility specific weights.

4.2.3. Correlation between Barrier Indicators

When barrier functions are to be used as a part of the DFU incidents weighing process, correlation between barrier functions' performances could have a significant effect on the estimated risk level. It has however been decided not to take correlation into consideration at this stage. Further discussion around correlation can be found in chapter 6.1.4.

4.3. Development of a New Model

This chapter will give a description of the new weight estimation model. The model consists of two risk level estimation methods; method 1 and method 2.

4.3.1. General about the New Model

As described in chapter 4.1, incidents are weighted to express their contribution to the estimated risk level. The new weights will partly be based on the idea behind the generic weights; calculation of facility specific weights will follow the same estimation procedure as the generic weights, and as long as it is practical DFUs with more than one category will use the same categories as the generic weights.

But where the generic weights are established for each type of facility, the new weights are estimated for each specific facility. This allows each facility's barrier functions to be taken into consideration when estimating the weights. As each facility's barrier functions affect the estimated weights, the reported barrier functions' performances can be included, giving time dependent and facility specific weights. This is illustrated in figure 3, which is a development of figure 2.

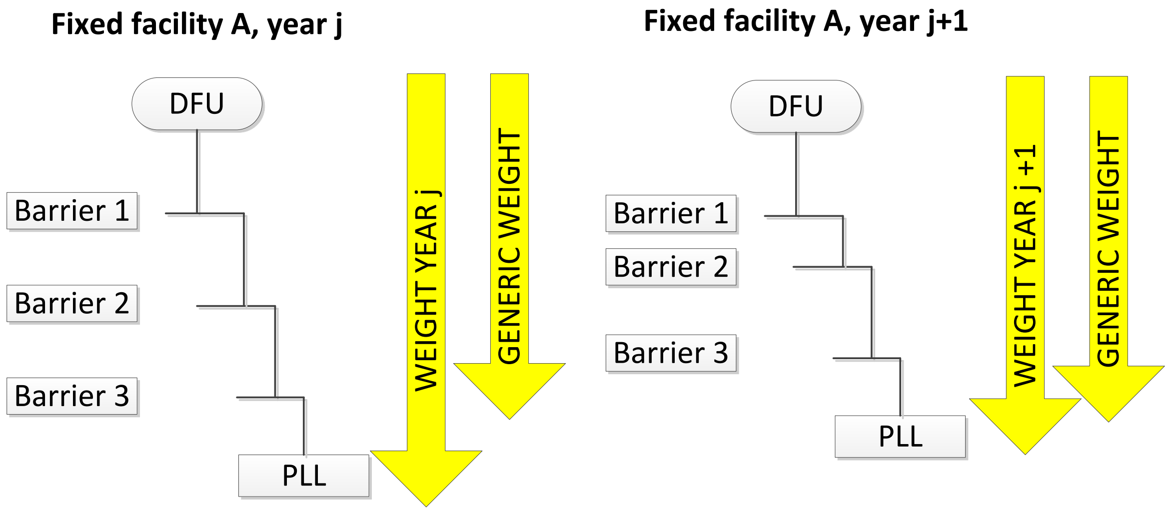


Figure 3 Time dependent weights, new model

By using the facility specific weights when estimating the total risk indicator for major accidents, a single indicator including both barrier functions' performances and actual incidents is estimated.

Facility specific weights will only be estimated for the DFUs with relevant barrier functions' performances reported. These are DFU 1, DFU 3, DFU 8 and DFU 9. In accordance with the master thesis limitations no weights are estimated for DFU 2 and DFU 10 even though there are reported relevant barrier functions for them. As the weights are estimated from QRA data, only weights for DFUs with corresponding cases treated in the QRA can be estimated and if any of the DFUs not are relevant for the facility, no weights are estimated.

For an overview of which barrier functions that are assigned to what DFUs, see appendix b and chapter 6.2.2.

As the observed barrier functions' performances change on an annual basis, the weights will be changed from year to year. This means that the weight estimation process have to be performed for every year since 2003 and each time new barrier function performances are reported.

4.3.1.1. Flow of the Weight Estimation Process

The process of establishing new weights is described in figure 4. Depending on the available data from the risk analysis, the process can follow two different paths. They are called method 1 and method 2 and will be described in chapter 4.3.2 and 4.3.3 respectively. In figure 4 the left path describes method 1 and the right path describes method 2.

When new weights have been estimated, the new risk level will be compared with the risk level obtained from the use of generic weights. A description of the comparison can be found in chapter 4.4.

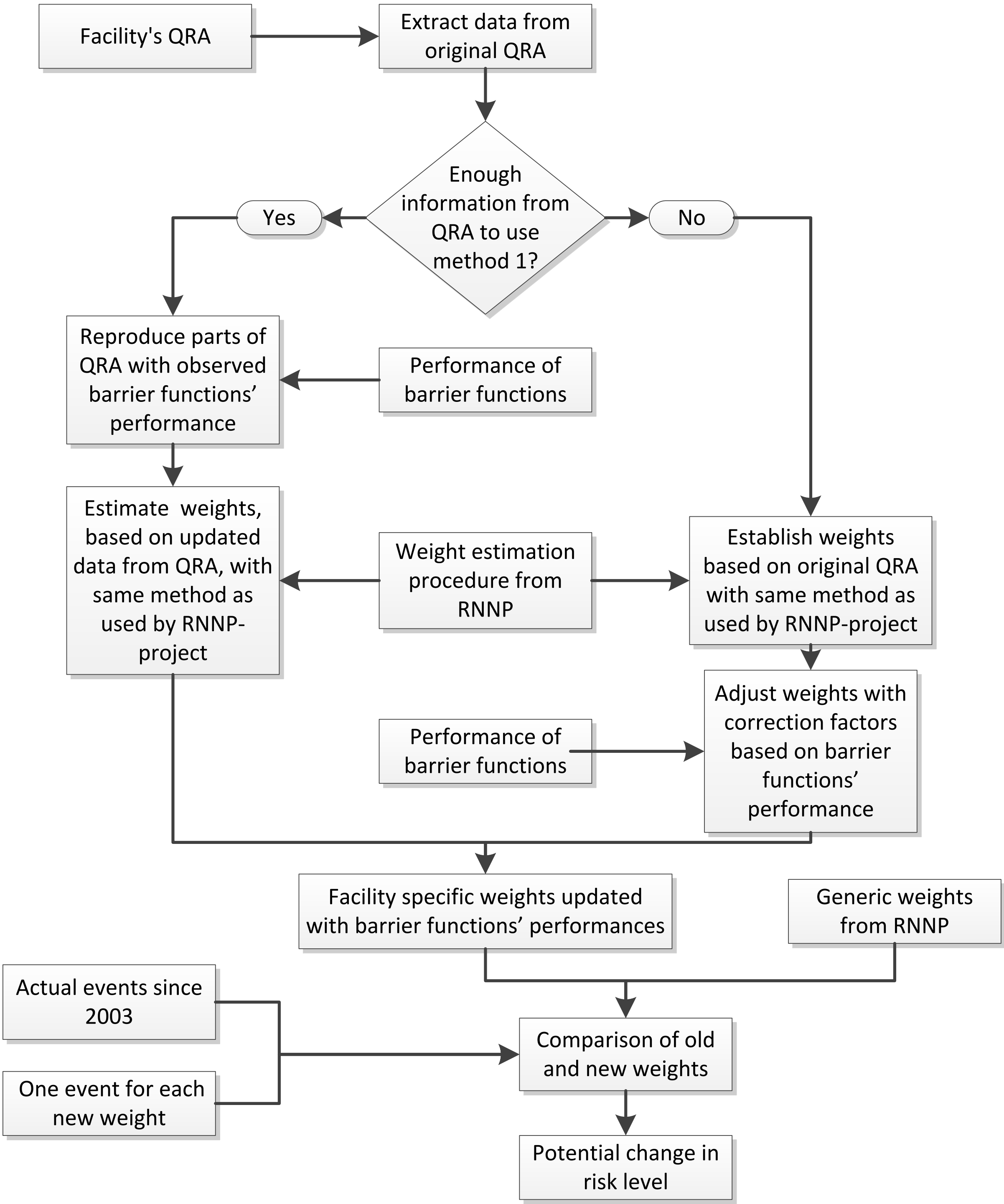


Figure 4 Flow of new model

4.3.1.2. Data from the Quantitative Risk Analysis

The first step when establishing weights is to extract information from the facility's QRA. The information extraction should be done from the latest available QRA revision to assure that the estimated weights are as up to date and correct as possible.

The extracted data should include the following:

- Assumptions the QRA is based on
- Which events the QRA is based on
- Barrier functions in place to stop/mitigate the events
- Risk expressed as PLL and/or FAR
- Persons OnBoard (POB), manning distribution and part of time they are exposed to risk
- Assumed frequencies of events

And if method 1 is to be used:

- Method used to estimate the risk
- Data necessary to reproduce parts of risk analysis

The extracted data will be used as basis for deciding which path in figure 4 to follow, and as input for both weight estimation methods.

4.3.1.3. Choice of Method

Which of the two methods that will be used to estimate the facility specific weights are chosen based on the available information from the QRA.

The two methods can be described as follows:

Method 1: This method is based on updating the assumed performance of barrier functions from the QRA with the reported barrier functions' performances. Then expected fatalities given a DFU incident is estimated with the same method as the generic weights from the "updated" QRA.

Method 2: When data from the QRA not is detailed enough to reproduce parts of risk analysis with updated barrier functions' performances. Method 2 estimates the risk level by an equation with a constant term and a variable term.

There are pros and cons with both methods. The first method demands a more detailed QRA and will be more time consuming, but it will give a more detailed and correct risk picture. The second method can give estimations based only on reported barrier functions' performances, risk from original QRA and benchmarks, but the estimated risk picture will not be as detailed and facility specific as from method 1. As a correct as possible risk picture is desired, method 1 is preferred when the necessary data is available. For further discussion, see chapter 6.2.1.

4.3.2. Method 1

4.3.2.1. Descriptions of Method 1

Method 1 is the left path in figure 4 and is described more detailed in figure 5. Method 1 uses observed barrier functions' performances to update and reproduce parts of the risk analysis. When the risk analysis has been "updated" with the observed performance, expected fatalities given an incident are calculated for each year barrier functions' performances are available. This calculation follows the same procedure as the generic weights, and is described in "Metoderapporten" (PSA, 2011b).

The QRA needs to contain enough information to be partly reproduced and there must be corresponding barrier functions used by the QRA and in the reported performances. To achieve a correct risk estimate as possible, barrier functions assigned to DFUs as described in appendix b will not be followed for method 1. Instead all barrier functions' performances that can be updated will be updated.

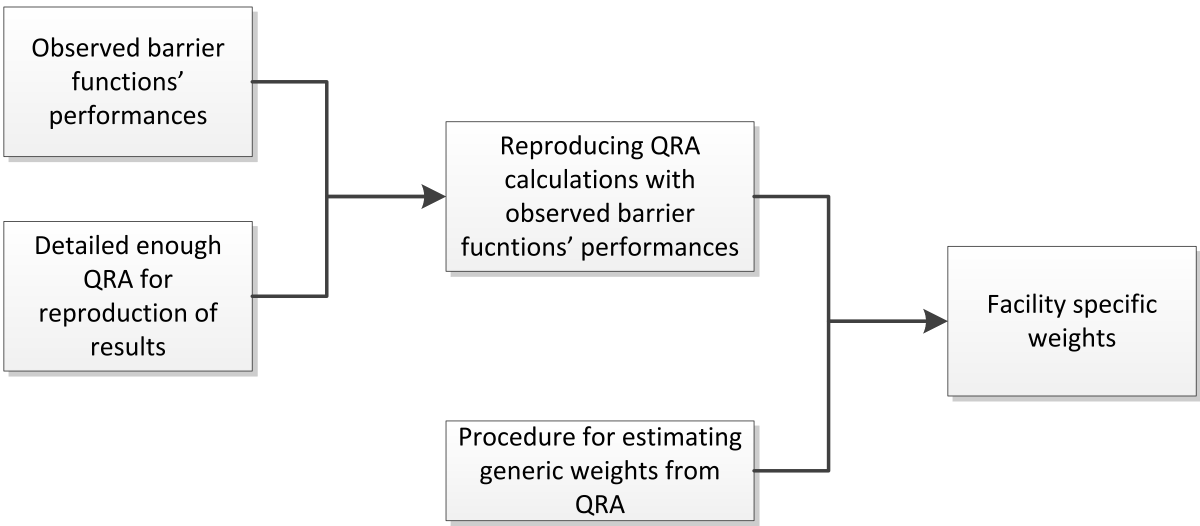


Figure 5 Estimation of weights, flow of method 1

4.3.2.2. Procedure for Method 1

As seen in figure 5, method 1 consists of two steps; reproducing parts of the QRA calculations and estimating facility specific weights.

Step one, reproducing QRA calculations, consist of updating the assumed performances of barrier functions in the QRA with the facility's observed performances. This can typically include estimating the PFD for the observed performances and implementing these into event trees. After the assumptions have been updated, the original QRA calculations have to be reproduced with the new performance data. How the reproduction is done and methods used depend on the specific QRA, but it will mainly be standard risk analysis methods. For an example, see test case 1 – CA, chapter 5.2.1.1.

The second step, estimating weights, consist of following the procedure described in "Metoderapporten"(PSA, 2011b) to estimate facility specific weights from the updated and reproduced QRA results.

4.3.3. Method 2

4.3.3.1. Description of Method 2

If the available QRA not is detailed enough to update and reproduce parts of the calculations, method 2 will still allow facility specific weights influenced by barrier functions' performances to be estimated.

This is achieved through the use of an equation consisting of a constant and a variable term as expressed by equation 18.

$$Weight_{Method\ 2,\ DFU\ i} = Weight_{From\ orignal\ QRA} * (1 + Correction\ factor_{DFU\ i})$$

Equation 18

The constant term is the facility specific weight estimated from the original QRA, and the variable term is an adjustment of it. The adjustment is done by the use of a weight correction factor. Further description is given in the next chapter.

Method 2 will like method 1 give facility specific weights, but the correction factors make the weights less specific and increases the chances of wrong risk estimates. See chapter 6.2.4 for further discussions.

4.3.3.2. Procedure for Method 2

The flow of method 2 is outlined in figure 6. Unlike method 1, method 2 consists of three steps.

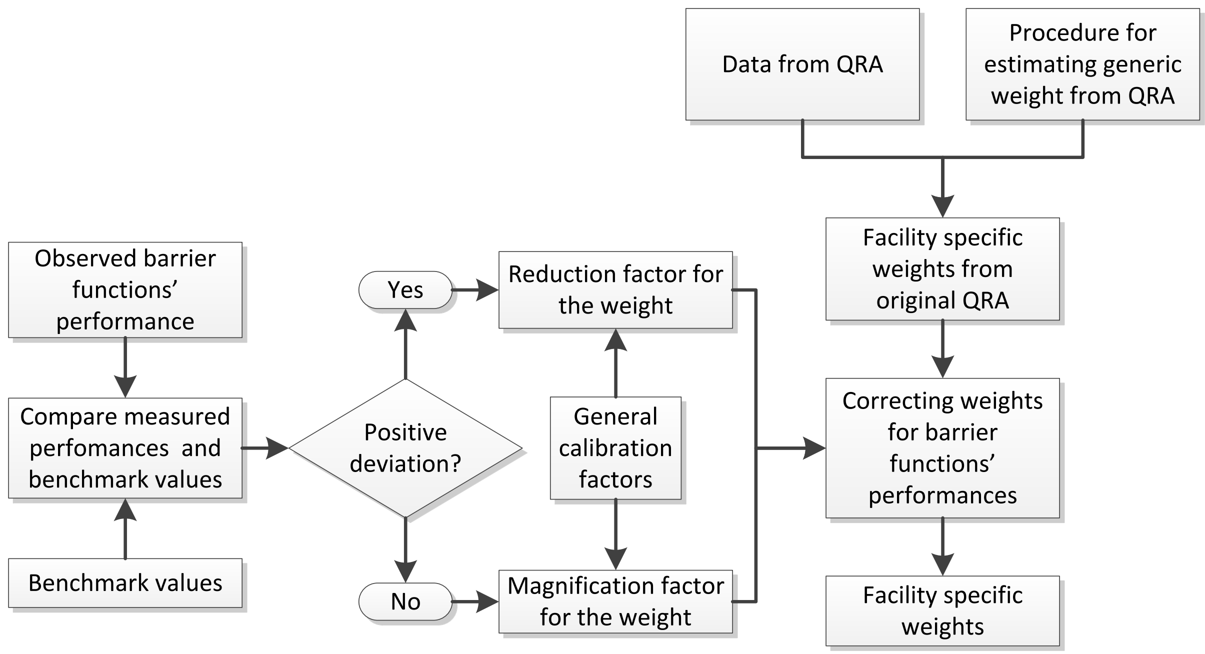


Figure 6 Estimation of weights, flow of method 2

The first step is to estimate facility specific weights based on the original QRA. The weights are estimated by the same procedure as the generic weights, given by "Metoderapporten" (PSA, 2011b). The estimated weight will be the constant term and basis for the variable term in equation 18.

The second step consists of finding the weight correction factors and the process can be described as follows:

1. Comparing barrier functions' performances with benchmarks. Benchmark failure rates can be found in appendix c and is discussed in chapter 6.2.4.1. The performance comparison is expressed as a "Deviation ratio", calculated by equation 19.

$$Deviation\ ratio_j = \frac{Barrier\ function\ failure\ rate_j - Benchmark\ failure\ rate_j}{Benchmark\ failure\ rate_j}$$

Equation 19

Equation 19 will give a negative deviation ratio if the observed barrier function's performances are better than the benchmark and a positive deviation if the observed performances are worse than the benchmarks.

2. Barrier function correction factors. This is the difficult part of method 2 and will be further addressed in chapter 4.3.3.3 and discussed in chapter 6.2.4. The barrier function correction factor will either be a reduction factor or a magnification factor, depending on whether the deviation ratio is negative or positive.

Each barrier function correction factor will be calculated by equation 20.

$$Correction\ factor_{Barrier\ function\ j} = (Deviaton\ ratio_j) * (Calibration\ factor_j)$$

Equation 20

Calibration factor will be described in chapter 4.3.3.3.

3. Weight correction factors. When barrier function correction factors have been established for all barrier functions, weight correction factors to adjust the weights established from the original QRA are calculated. This is done by an arithmetic average, and described by equation 21.

$$Correction\ factor_{DFU\ i} = \frac{\sum_{j=1}^{n_i} Correction\ factor_{Barrier\ function\ j}}{n_i}$$

Equation 21

Where n_i is the total number of barrier functions assigned to DFU i.

When both weights from the original QRA and weight correction factors have been estimated, the last step of method 2 is adjustment of weights. It consists of adjusting the weights established in step one with the weight correction factors established in step 2 by the use of equation 18.

Method 2 will be used for both of the test cases, see chapter 5 for further demonstration.

4.3.3.3. Calibrating Method 2

The use of correction factors introduces certain challenges. These challenges are interconnected and can be defined as follows.

- Acceptable range for the deviation ratio
- Minimum and maximum values of correction factors
- Finding "correct" calibration factor

A thorough discussion around these challenges can be found in chapter 6.2.4.5. Only a brief description and solution to the challenges will be given in this chapter.

Deviation Ratio

All calculated deviation ratios will be accepted. This is based on the assumptions made in chapter 4.2.2, stating that a three year rolling average will be used to avoid unrealistically high failure rates, and if no failures has occurred during testing a half failure is assumed to avoid a failure rate of zero. Acceptable values of deviation ratios are further discussed in chapter 6.2.4.2.

Correction Factor

For correction factors there are certain limitations to what the acceptable values can be. As a lower value, the correction factor can never be less than or equal to – 1 (negative 1). This would indicate no risk or a decrease in the risk level following a major accident precursor.

The absolute upper value of correction factors are the ones giving expected fatalities equal to people onboard. Acceptable values of correction factors are discussed in chapter 6.2.4.4.

Calibration Factors

Method 2 is calibrated to decide the barrier functions' performances effect on the risk level. This is done by first establishing facility specific weights by method 1, and then method 2 is used to find the calibration factors that give the same weights as estimated by method 1.

From demanding weights estimated by method 1 being equal to weights estimated by method 2 and using equation 18, equation 20 and equation 21, an expression for calibration factors can be derived. This expression is given in equation 22.

$$\sum_{n_i} C_j * D_j = \left(\frac{\text{Weight}_{\text{Method 1, DFU i}}}{\text{Weight}_{\text{QRA, DFU i}}} - 1\right) * n_i$$

Equation 22

The following notation is used in equation 22:

- n_i = number of barrier functions assigned to DFU i.
- C_j = Calibration factor for barrier function j.
- D_j = Estimated deviation ratio for barrier function j.
- $Weight_{Method\ 1,\ DFU\ i}$ = Weight estimated for DFU $_i$, by method 1.
- $Weight_{QRA,\ DFU\ i}$ = Weight estimated for DFU $_i$ from original QRA.

Equation 22 is only the basis for estimating calibration factors. The process will be based on iteration, i.e. trial and error until a good fit is achieved. This is first done for each year and then arithmetic average of the annual calibration factors is used to estimate the facility's calibration factors.

As given in appendix b there is more than one barrier function assigned to each of the DFUs. This will give certain challenges with assigning "correct" calibration factor to the "correct" barrier function. This issue will be critical when calibration factors only are based on a single facility. As more than one facility is to be used when general calibration factors are estimated, this issue should not be a problem.

As calibration factors determines the barrier function correction factors, certain guidelines for values calibration factors can take on have been established.
If the deviation ratio equals zero, the calibration factor equals zero.
The lower limit of calibration factors is rooted in the correction factors' lower limit. Calibration factors should never give barrier function correction factors equal to or less than -1 (negative 1).
Calibration factors are limited upward by the maximum allowed value of the estimated weights. I.e. calibration factors can never give correction factors resulting in weights with higher expected fatalities than there are people onboard. As all values of correction factors less than this are allowed, all values of calibration factors below this will be allowed as well. But if expected fatalities approaches total loss, the calibration factors used should be individually assessed and if necessary adjusted, as there almost always will be more barriers in place than the ones performances is reported for.
Method 2 is based on linearity, i.e. all equations are linear. This assumption is questionable, but it has been decided that benefits achieved from not using linearity not will justify the extra complications it would bring to the weight estimation method. This is further addressed in chapter 6.2.4.3.
There is not necessarily symmetry for the calibration factors, i.e. there can be different calibration factors for positive and negative deviation ratios for the same barrier function. Calibration factors for positive deviation ratios must therefore be estimated from facilities with positive deviation ratios, and correspondingly, calibration factors for negative deviation ratios must be calibrated from facilities with negative deviation ratios.
General calibration factors are estimated so method 2 can be used whenever not enough information is available to estimate weights by method 1. This will be further addressed in chapter 4.5.1. The description given in this section is only intended to describe how calibration factors can be estimated for a single facility.
For a more detailed walkthrough and demonstration of calibration, see chapter 5.2.1.3.

4.4. Weight Comparison

The last part in figure 4 is comparison between the estimated weights and the generic weights. Generic weights are taken from the report "Metoderapporten" (PSA, 2011b) and reproduce in appendix a.
Two cases will be compared; actual incidents for the facility and one incident for each new weight.

Actual Incidents

The actual incidents comparison looks at the facility's incidents from year 2003 to and including year 2012. This estimates the potential change in contribution to the risk level for the specific facility.

One Incident for Each New Weight

To show the new weights' total potential to change the estimated risk, one incident is assigned to each DFU new weights have been estimated for. For DFU 1 incidents, the comparison will not consist of a single incident, but plotting of the whole range of leak sizes.

4.5. Generalization of Model

This chapter will give a brief description of the generalization of the model developed previously in this chapter.

4.5.1. General Calibration Factors

When calibration factors are estimated for a single facility as described in chapter 4.3.3.3, arithmetic average of annual calibration factors are used to estimate the facility calibration factors. The same approach is proposed used when establishing general calibration factors. But instead of taking arithmetic average of annual calibration factors, arithmetic average is taken of several facility calibration factors to estimate a set of general calibration factors.
It is impossible to define a definite amount of facilities the general calibration factors should be based on. As a minimum there should be included so many facilities that calibration factors can be established both for positive and negative deviation ratios for all barrier functions. There is no clear upper limit to how many facilities that should be included. It should be kept in mind that as more facilities are included, each facility's calibration factors will have less impact on the general calibration factors, eliminating the issue of facility specific attributes disproportionately affecting the general calibration factors.
As only one set of general calibration factors are estimated, all weights estimated with them must be manually controlled. This is done to assure that the estimated weights are within possible limits. I.e. all correction factors are higher than -1 (negative 1) and no weights give excepted fatalities higher than people on board.
Only establishing one set of general calibration factors leads to certain challenges, these will be further addressed in chapter 6.2.4.5.

4.5.2. Generalization and Implementation into “Trends in Risk Level” Project

There has not been identified any issues with the new model causing need for adaptions to be made before it can be generalized for all facilities. Description of the general model will therefore be the same as given in chapter 4.3, except for the use of general calibration factors estimated according to chapter 4.5.1.

An implementation of the model into the RNNP-project will however include a lot of initial work, as weights must be estimated for each specific facility. This work does not necessarily have to be done all at once. A reduction in the initial workload could be achieved by only estimating weights as they are needed, e.g. estimate CA’s DFU 1 weights only after a DFU 1 incident has occurred at CA.

5. Testing and Verification of Model

Testing is included as part of the new weight estimation model development. Two test cases are used for this purpose.

The test cases were chosen from a selection of available QRAs. These two cases were chosen based on that they are manned during normal operation, they represent different facility categories in the RNNP-project and the available QRAs are fairly detailed. The two facilities are made anonymous by the same tags as used by the RNNP-project, CA and AK, and will be referred to by these tags throughout this master thesis.

Originally a third facility, BH, was also supposed to be included as a test case. It has however been left out. This will be further addressed in chapter 6.4.2.

The first of the two test cases, CA, allows both methods to be used for weight estimation. The second test case, AK, only allows the use of method 2.

This chapter demonstrates the weight estimation process, results of estimation and a comparison between generic and estimated weights. Estimation of calibration factors based on CA is also included.

5.1. Available Data

For testing of the new model, the following data has been available.

Occurrences of Major Accident Precursors

Incidents reported since 1996 has been available (PSA, 2013b). This is the same data the RNNP-project estimates the risk level from.

For the test cases this data is used to identify incidents the comparison between generic and estimated weight can use.

Observed Barrier Functions' Performances

Observed barrier functions' performances has been reported since 2002 as part of the RNNP-project (PSA, 2012a, 2013c). As the master thesis limitations states that only data from year 2003 and later is to be used, the first year of reported barrier functions' performances is ignored.

Observed barrier functions' performances are one of the key parameters when estimating facility specific weights. It is used both by method 1 to update and reproduce parts of the risk analysis and by method 2 to calculate the deviation ratio.

Quantitative Risk Analyses

QRA-reports have been available for both AK and CA (Det Norske Veritas, 2007, SAFETEC, 2009). For CA, parts of the risk analysis has also been available (SAFETEC, 2013).

The QRA-reports are used when estimating facility specific weights. Both necessary data and assumptions the analysis is based on are extracted. The part of the risk analysis available for CA has been used to estimated weights by method 1, as not all the required data was available from the QRA-report.

5.2. Test Case 1 – CA

This chapter describes how both methods are used to estimate CA's facility specific weights. A description of how method 2 is calibrated is also included.

By facility categorization as used in the RNNP-project, CA is a floating production unit.

Facility specific weights will only be estimated for DFU 1. No weights are estimated for DFU 3 or DFU 9. Further discussion of this can be found in chapter 6.3.1.

5.2.1. Model

5.2.1.1. Method 1

Performance of CA's barrier functions has been reported since 2008. The failure rate is expressed as a three year rolling average, this gives three failure rates for each barrier function; year 2010, year 2011 and year 2012. Assuming that failure rate can be described by an exponential distribution, corresponding PFDs have been calculated. These are reproduced in appendix d.

Risk has been analyzed by event trees. The event trees' node probabilities have been updated with PFD calculated from observed barrier functions' performances. Assumptions made regarding node probabilities can be found in chapter 6.3.1.1. All other data than observed barrier functions' performance, are taken from CA's risk model and QRA-report (SAFETEC, 2009, SAFETEC, 2013). Observed barrier functions' performances are taken from the RNNP project (PSA, 2013c).

DFU 1 – Non-ignited Hydrocarbon Leaks

For DFU 1, the barrier functions used both in the event tree analysis and reported in the RNNP project are BDV, fire detection, deluge valve and gas detection. Of these, only BDV and gas detection are assigned to DFU 1 by "Metoderapporten" (PSA, 2011b), but all four will be used to update the event trees.

CA's event tree analysis is performed for three representative leak rates and 57 different leak scenarios. PLL is estimated for each of the three leak rates. This is analyzed for 4 sets of barrier functions' performances; as assumed in the QRA and observed performances for 2010, 2011 and 2012. The resulting risk estimate for each of the four sets of barrier functions' performances are presented in column 2 to 5 in table 3.

Table 3 Risk (PLL) for DFU 1, Method 1, CA

Leak category	QRA	2010	2011	2012
Small Leak	6,20E-04	6,32E-04	6,17E-04	6,07E-04
Medium Leak	2,72E-04	2,80E-04	2,63E-04	2,51E-04
Large Leak	5,50E-04	5,54E-04	5,40E-04	5,29E-04
Total	1,44E-03	1,47E-03	1,42E-03	1,39E-03

To find expected fatalities given a DFU 1 incident, risk estimates from table 3 is combined with the QRA's assumed annual leak frequencies for each of the leak categories. The results are presented in table 4, where expected fatalities given an incident for each of the four sets of barrier functions' performances are given in column 2 to 5.

Table 4 Expected fatalities given DFU 1 occurrence, Method 1, CA

Leak category	QRA	2010	2011	2012
Small Leak	1,50E-03	1,52E-03	1,49E-03	1,46E-03
Medium Leak	2,12E-03	2,18E-03	2,05E-03	1,95E-03
Large Leak	1,00E-02	1,01E-02	9,82E-03	9,64E-03

Table 4 present three point values of expected fatalities given a leak for the four sets of barrier functions' performances. These are used as basis for linear regression in Microsoft Excel and weight equations are estimated. The estimated weight equations are given in table 5. For easy comparisons, the generic weights are also included. No weight equations are estimated for leaks larger than 10kg/s, this is the same as for generic weights.

Table 5 DFU 1, weight equations, Mehtod 1, CA

Leak category	Small leak	Medium leak	Large leak
Leak size	0,1 to 1,0 [kg/s]	1,0 to 10 [kg/s]	More than 10 [kg/s]
QRA	0,0002*X+0,0014	0,0005*X-0,0007	Individually assessed
2010	0,0002*X+0,0015	0,00055*X-0,0007	Individually assessed
2011	0,0001*X+0,0014	0,0005*X-0,0008	Individually assessed

2012	0,0001*X+0,0014	0,0005*X-0,0008	Individually assessed
Generic weights	0,0016*X+0,0003	0,0045*X-0,016	Individually assessed

The input values for method 1 calculations, can be found in appendix g.

5.2.1.2. Method 2

DFU 1 – Non-ignited Hydrocarbon Leaks

As part of method 1, the three leak categories' expected fatalities given an incident were estimated based on the QRA's assumed barrier functions performances. The resulting risk estimates are reproduced in table 6.

Table 6 DFU 1, expected fatalities per leak, QRA assumptions, CA

Leak category	Leak size [kg/s]	QRA
Small leak	0,5	0,0015
Medium leak	5	0,0021
Large leak	20	0,0100

Estimated expected fatalities per leak from the QRA are adjusted using correction factors according to equation 18. The correction factors are estimated from barrier function's performances and calibration factors. Calibration factors are described in chapter 5.2.1.3. After the point values of risk given in table 6 have been adjusted, linear regression in Microsoft Excel is used to estimate weight equations for year 2010, year 2011 and year 2012. The resulting correction factors and weight equations are presented in table 7.

Table 7 DFU 1, correction factors and weight equations, Method 2, CA

Year	**Correction factor**			**Weight equation**	
	0,5 Kg/s	5 Kg/s	20 Kg/s	Leak rate (X) 0,1 to 5,5 [kg/s]	Leak rate (X) 5,5 – 10 [kg/s]
2010	0,069	0,175	-0,061	0,0002*X+0,0015	0,0005*X+5E-5
2011	-0,037	-0,074	-0,090	0,0001*X+,0014	0,0005*X-0,0006
2012	-0,041	-0,127	-0,095	0,0001*X+0,0014	0,0005*X-0,0008

5.2.1.3. Calibrating Method 2

Ideally general calibration factors based on several facilities should be estimated. In this master thesis however, it will only be done for one test case to demonstrate the calibration process.

This chapter will present the estimation of calibration factors based on CA. It can be viewed as a walkthrough or a demonstration of how calibration factors can be estimated.

DFU 1 – Non-ignited Hydrocarbon Leaks

Most QRAs for the NCS use three leak sizes as DFU 1 incidents (PSA, 2011b). When estimating weights by method 2, expected fatalities for each of these three leak sizes are adjusted by correction factors. This means that three sets of calibration factors must be estimated, i.e. one for each leak size.

For DFU 1 barrier function correction factors are estimated for gas detection, BDV, Pressure Safety Valve (PSV), riser ESDV and wing and master valve. All these barrier function correction factors must be calibrated.

The calibration is based on demanding weights estimated by method 2 being equal to weights estimated by method 1. This is expressed in equation 23 and described in chapter 4.3.3.3.

$$Weight_{Bar.\ i,\ Met.\ 1} = Weight_{Bar.\ i,\ QRA} * (1 + Correction\ factor_{Bar.\ i})$$

Equation 23

For CA, necessary correction factors are calculated for leak sizes 0,5 kg/s, 5 kg/s and 20 kg/s. Iteration is used to estimate calibration factors. In the following a description of the iteration process is described.

1. Finding necessary weight correction factors.

   By equating weights for year 2010, year 2011 and year 2012 by method 1 and using weights estimated from the QRA, the necessary weight correction factors to have method 1 weights equal to method 2 weights can be estimated according to equation 24.

$$Necessary\ correction\ factor_{DFU\ i,\ Year\ j} = \frac{Weight_{DFU\ i,\ Met.\ 1,\ Year\ j}}{Weight_{DFU\ i,\ QRA}} - 1$$

Equation 24

2. Adjusting annual calibration factors.

   When the necessary weight correction factors have been established, the annual calibration factors are adjusted so that the estimated weight correction factors equal the necessary weight correction factors. This process is based on trial and error and is done for each year failure rates are available, i.e. year 2010, year 2011 and year 2012.

3. Estimating calibration factors.

   The facility's calibration factors are calculated as an arithmetic average of the annual calibration factors. These are used to calculate weight correction factors.

4. Controlling estimated weight correction factors.

   Do the estimated weight correction factors approach the necessary weight correction factors? If not, go back to step 2 and adjust the annual correction factors. When acceptable fit of weight correction factors has been achieved, the estimated calibration factors will be accepted.

5. Comparing all leak sizes.

   To control that the final calibration factors give sufficiently good weight correction factors, a comparison of method 1 weight equations and method 2 weight equations are graphed for all leak sizes. If these graphs not are similar, the annual calibration factors should be adjusted and step 2, 3 and 4 repeated. For CA, these graphs can be seen in chapter 5.2.1.4.

Different calibration factors are estimated for positive and negative deviation ratios. This means that 30 calibration factors should be estimated for CA. In table 8 CA's estimated calibration factors are presented; both the annual and the facility's calibration factors. As it can be seen, not all 30 calibration factors are estimated. This is a consequence of CA not having both positive and negative deviation ratios for all barrier functions. The calibration factors not estimated from CA are negative riser ESDV and positive PSV and wing and master valve. Another important issue that can be seen from table 8 is that calibration factors not necessarily are symmetric for positive and negative deviation ratios.

Table 8 DFU 1, calibration factors, based on CA

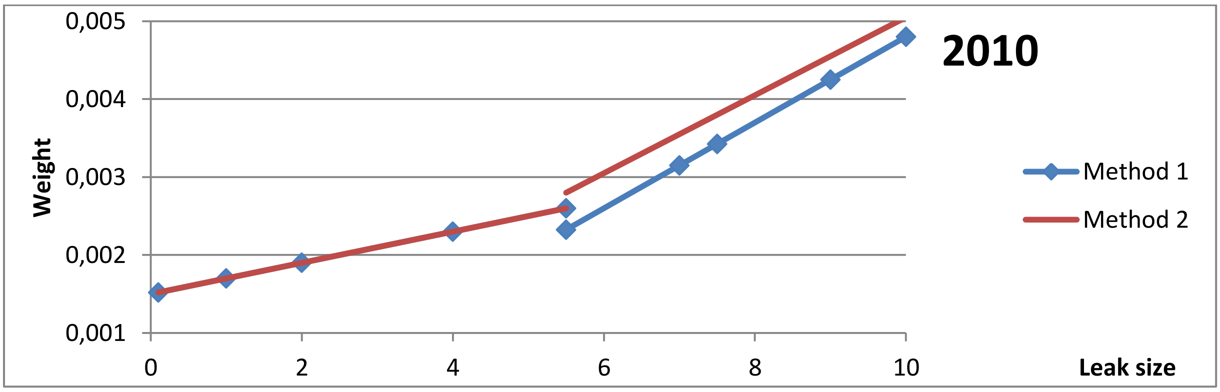
Leak size	**Barrier function**	**2010**		**2011**		**2012**		**Average calibration factor**	
Deviation ratio		**Negative**	**Positive**	**Negative**	**Positive**	**Negative**	**Positive**	**Negative**	**Positive**
Leak rate 0,5 [kg/s]	Gas detection		0,15		0,15	0,05		0,05	0,15
	BDV		0,40	0,30		0,12		0,21	0,40

	PSV	0,75		0,60		0,45		0,60	
	Riser ESDV		0,26		0,34		0,35		0,32
	Wing and master valve	0,70		0,55		0,45		0,57	
Leak rate 5 kg/s	Gas detection		0,45		0,30	0,30		0,30	0,38
	BDV		0,60	0,75		0,40		0,58	0,60
	PSV	0,75		0,75		0,20		0,57	
	Riser ESDV		0,30		0,30		0,30		0,30
	Wing and master valve	0,75		0,75		0,60		0,70	
Leak rate 20 kg/s	Gas detection		0,05		0,01	0,05		0,05	0,03
	BDV		0,10	0,05		0,05		0,05	0,10
	PSV	0,40		0,20		0,20		0,27	
	Riser ESDV		0,06		0,12		0,10		0,09
	Wing and master valve	0,50		0,50		0,50		0,50	

5.2.1.4. Results CA

DFU 1 – Non-ignited Hydrocarbon Leaks

The weight equations for DFU 1, estimated both by method 1 and by method 2, can be seen graphed in figure 7. The two methods give fairly similar results. Worth noticing is that both methods gives some discontinuities for the weight equations, this is addressed in chapter 6.7.



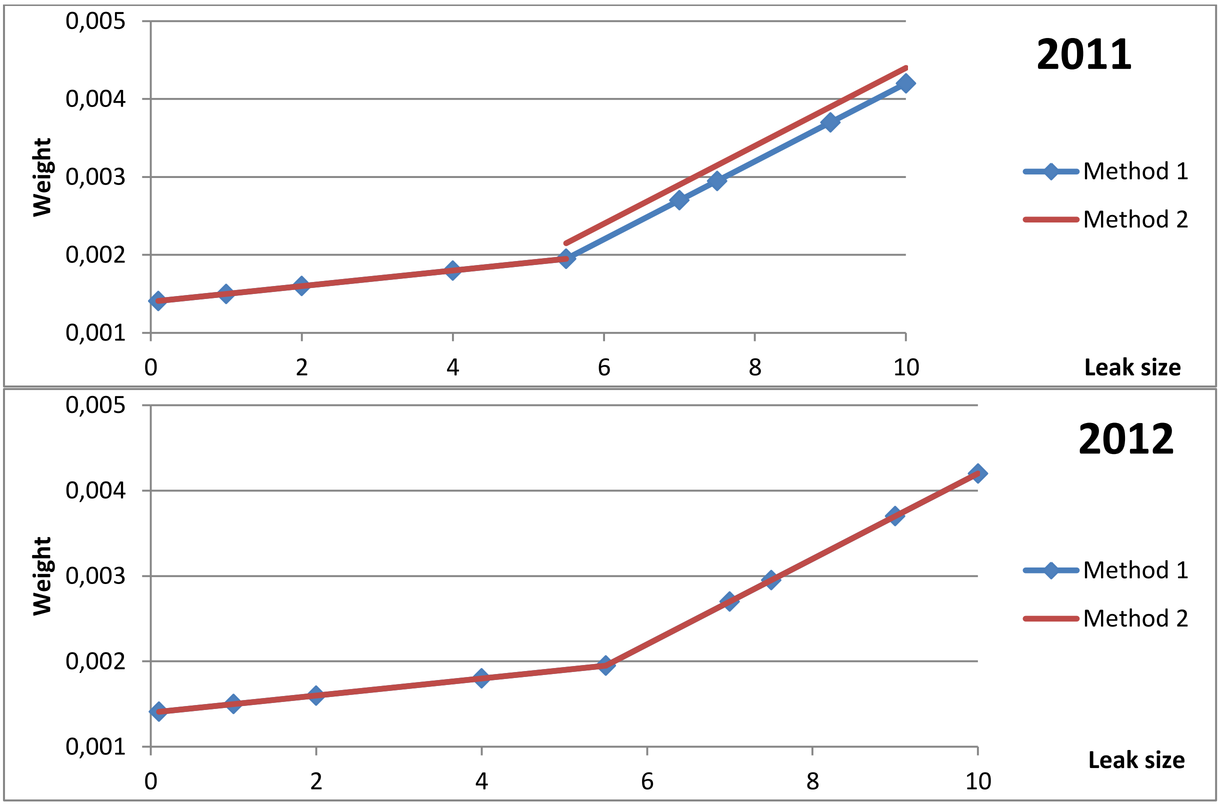
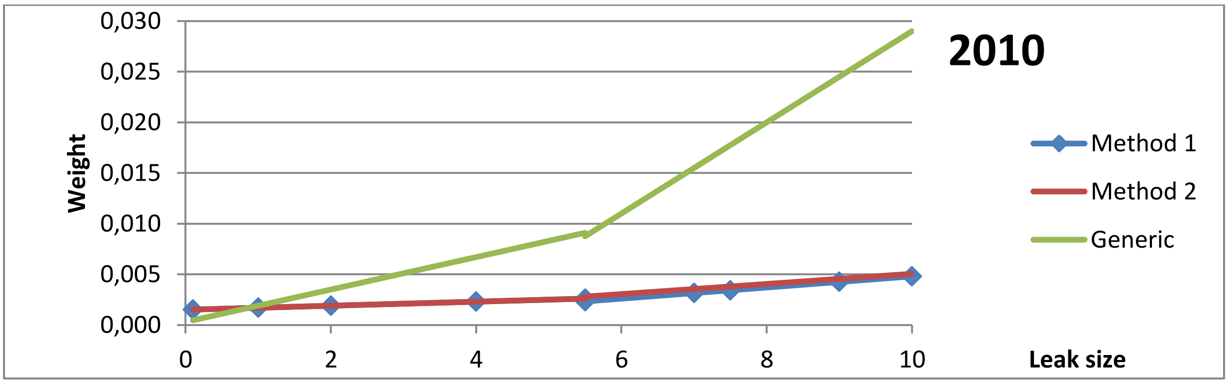


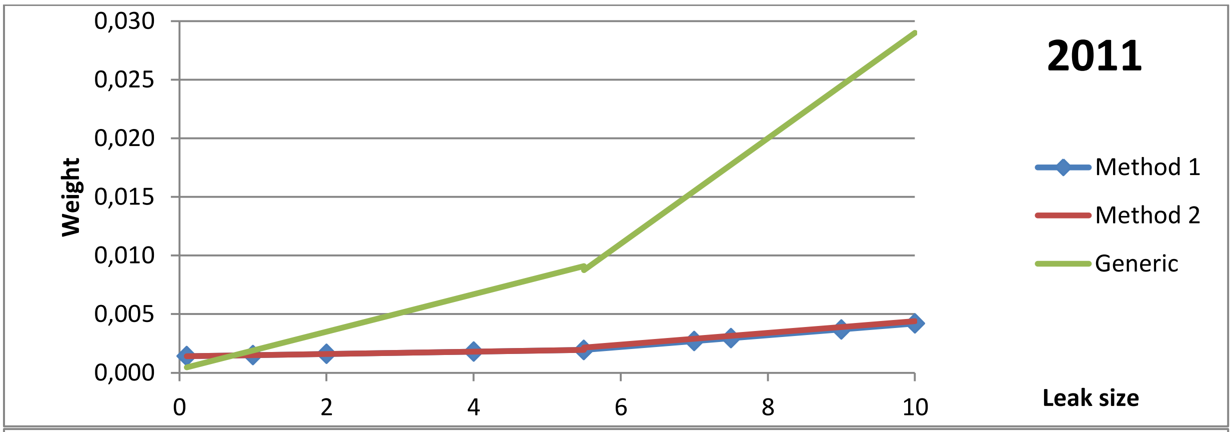
Figure 7 DFU 1, weights for method 1 and method 2, PLL, CA

5.2.2. Comparison of Results

DFU 1

In figure 8 a comparison of the estimated and the generic weight equations for DFU 1 is presented. In figure 7 it can be seen that there are some differences in the results for the two estimation methods, but as shown in figure 8, these differences are small compared to the difference between the generic and estimated weights.





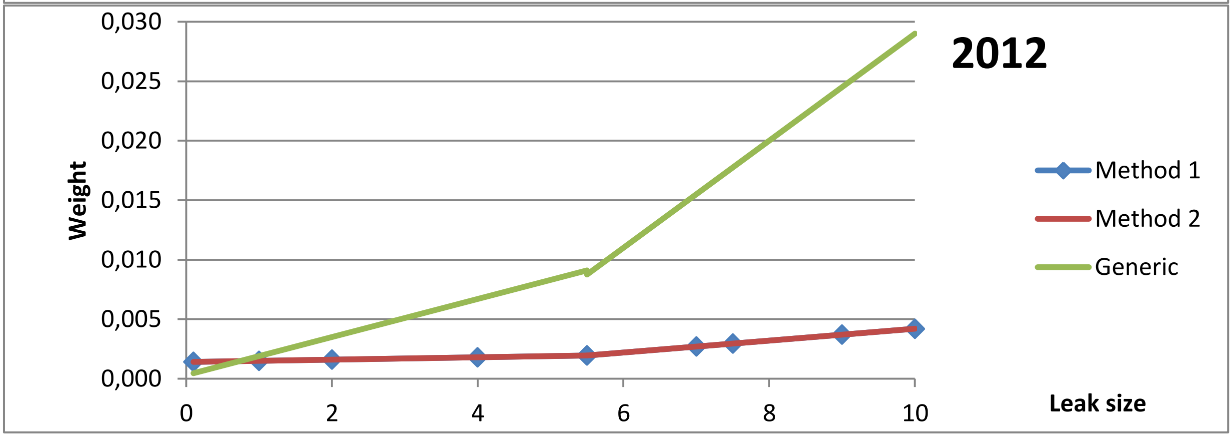


Figure 8 DFU 1, comparison of estimated weight equations, PLL, CA

Chapter 4.4 states that a comparison based on actual events shall be performed as well. For CA the only DFU 1 incident was in the first year of operation; year 2008. As a three year rolling average is used, it is not possible to estimate weights for 2008. In figure 9 however a comparison of the estimated risk contribution the incident would have had if it happened after 2009 is given for method 1, method 2 and the generic weights. From the figure it can be seen that the generic weight gives the highest estimated risk level contribution for all years. It can also be seen that that there is a reduction in CA's estimated contribution to the risk level, and that the 2011 and 2012 contribution is the same.

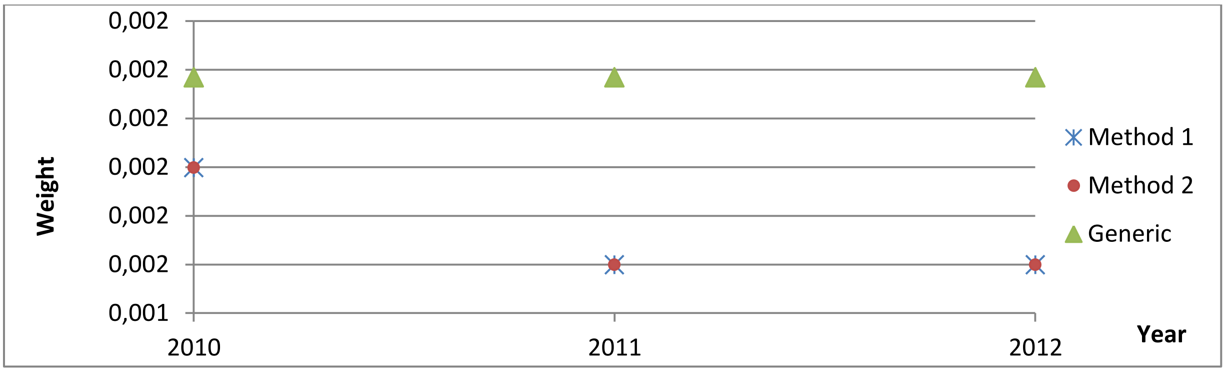


Figure 9 DFU 1, risk from 2008 incident, PLL, CA

5.3. Test Case 2 – AK

AK's QRA is not detailed enough for method 1 to be used; weights are therefore estimated by method 2.

For AK, facility specific weights are estimated for DFU 1. No weights are estimated for DFU 3, DFU 8 and DFU 9.

According to the RNNP-categorization, AK is a fixed facility. All data used in the weight estimation process has been taken from AK's QRA (Det Norske Veritas, 2007).

5.3.1. Model

5.3.1.1. Method 2

DFU 1 – Non-ignited Hydrocarbon Leaks

Like CA, AK has three leak categories. The largest of the leak categories however use a larger leak sizes as representative leak rate. Despite the difference in representative leak rates, the same calibration factors will be used.

In table 9 leak categories, representative leak rate and estimated fatalities given a representative leak incident are presented. Expected fatalities have been estimated from AK's QRA, combining PLL and annual leak frequencies for the representative leak rates. Data used for the calculations are reproduced in appendix h.

Table 9 DFU 1, leak rate and expected fatalities from QRA, AK

Leak category	Representative leak rate	Expected fatalities given representative leak
0,1 to 1,0 [kg/s]	0,5 [kg/s]	0,000218
1,0 to 10 [kg/s]	5 [kg/s]	0,00308
More than 10 [kg/s]	50 [kg/s]	0,0258

Based on AK's barrier functions' performances and method described in in chapter 4.3.3.2, weight correction factors for the expected fatalities are estimated. The calibration factors used are the ones estimated for CA in chapter 5.2.1.3. Calibration factors were not estimated for both negative and positive deviation ratios for all CA's barrier functions. Where this is the case, symmetry of calibration factors has been assumed.

Estimated weight correction factors and expected fatalities given leak adjusted by weight correction factors, are presented in table 10. Correction factors are also graphed in figure 10 for an easy general overview of barrier functions performances' development.

Table 10 DFU 1, correction factors and adjusted expected fatalities given leak, Method 2, AK

Year	Correction factor			Adjusted expected fatalities given representative leak		
	0,5 [kg/s]	5 [kg/s]	50 [kg/s]	0,5 [kg/s]	5 [kg/s]	50 [kg/s]
2005	1,44	1,85	0,382	0,000533	0,00878	0,0357
2006	1,20	1,56	0,298	0,000480	0,00789	0,0335
2007	1,05	1,37	0,259	0,000447	0,00730	0,0325
2008	0,422	0,386	0,0593	0,000310	0,00427	0,0274
2009	0,329	0,257	0,0303	0,000290	0,00387	0,0266
2010	1,72	2,37	0,384	0,000593	0,0104	0,0358
2011	3,12	4,52	0,737	0,000900	0,0170	0,0449
2012	2,90	4,19	0,669	0,000850	0,0160	0,0431

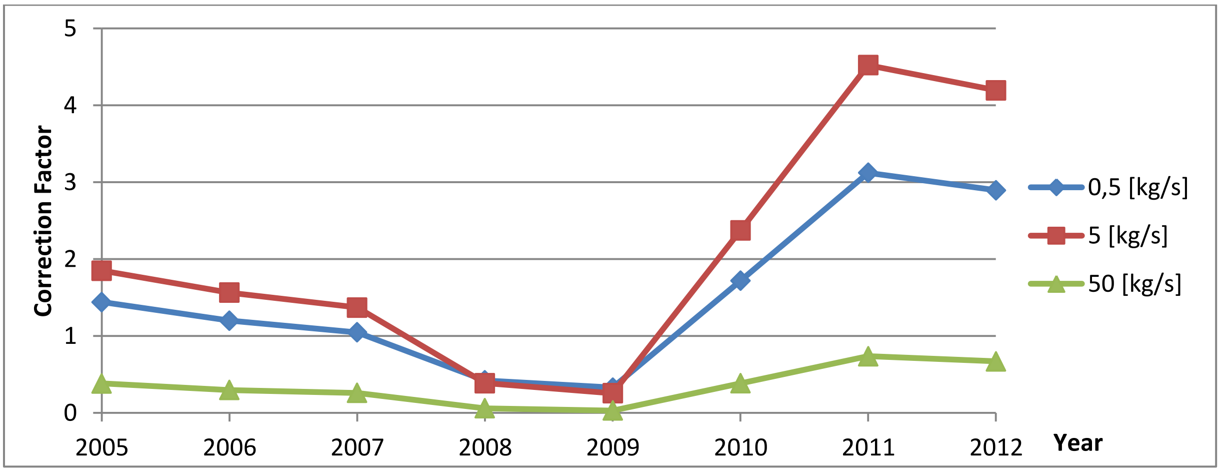


Figure 10 DFU 1, plot of correction factors, Method 2, AK

From figure 10 it is worth noticing that all correction factors for AK are magnification factors. This indicates that the barrier functions' performances for AK generally have been worse than the benchmark values.

Based on the adjusted expected fatalities, linear regression in Microsoft Excel is used to estimate the weight equations for year 2005 to and including year 2012. These equations are presented in table 11. For easy comparison, the generic weight equations are also included in the table.

Table 11 DFU 1, estimated weight equations, Method 2, AK

Year	Leak rate (X) [kg/s]	
	0,1 to 5,5 [kg/s]	5,5 to 10 [kg/s]
Generic weight	0,0016*X+0,0003	0,0045*X-0,016
2005	0,0016*X	2E-5*X+0,0087
2006	0,0015*X	2E-5*X+0,0078
2007	0,0014*X	2E-5*X+0,0072
2008	0,0008*X	2E-5*X+0,0042
2009	0,0007*X	2E-5*X+0,0038
2010	0,0019*X	2E-5*X+0,0103
2011	0,0032*X	2E-5*X+0,0169
2012	0,0030*X	2E-5*X+0,0159

Graphs of the estimated weight equations can be seen in chapter 5.3.1.2 and a comparison with the generic weights is given in chapter 5.3.2.

5.3.1.2. Results

DFU 1 – Non-ignited hydrocarbon leaks

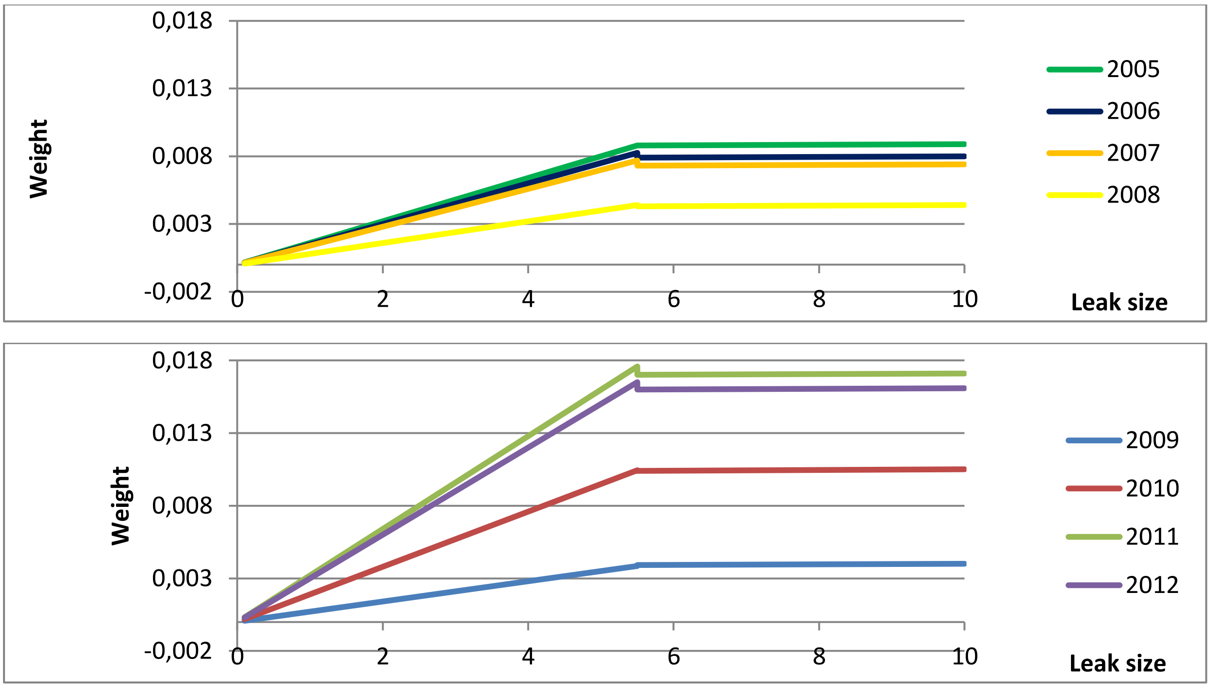


Figure 11 DFU 1, estimated weight equations, Method 2, AK

Figure 11 shows the estimated weight equations for AK. It can be seen that year 2010, year 2011 and year 2012 generally has higher estimated weights. The main reason for this increase in the risk level is more failures of wing and master valves after year 2010.

5.3.2. Comparison of Result

DFU 1 – Non-ignited Hydrocarbon Leaks

Figure 12 compares the generic weight equations and a selection of the estimated weight equations. Generic weights give the highest weight for very small leaks and leaks larger than 7,5 kg/s. For leak sizes between this, estimated weights are both smaller and larger, depending on annual variations in barrier functions' performances.

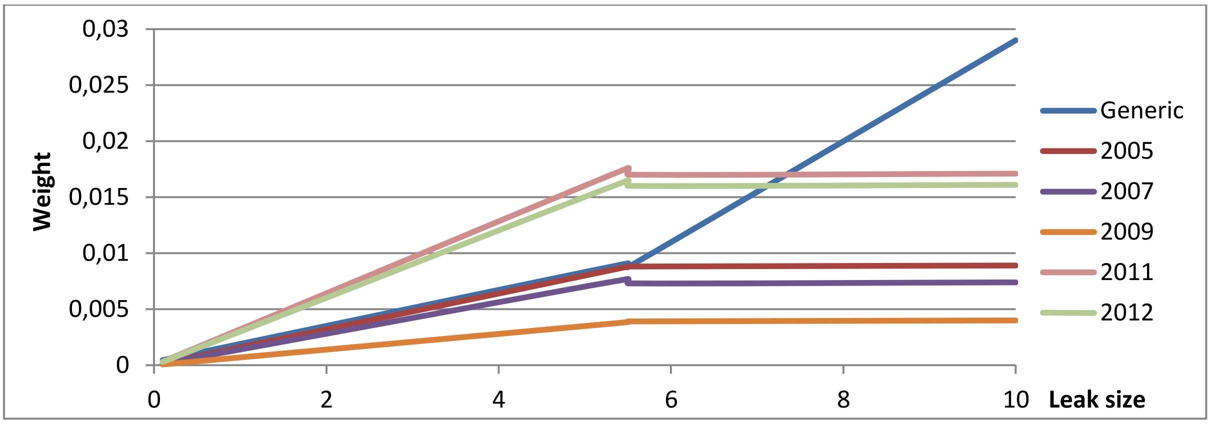


Figure 12 DFU 1, generic vs. selected estimated weight equations, Method 2, AK

The biggest difference between the estimated and the generic weight equations can clearly be seen in figure 12. For leaks larger than 5,5 kg/s, the generic weight equations increases much steeper than the estimated weight equations.

This probably has to do with a representative leak rate of 50 kg/s being used. For CA the largest representative leak rate is only 20 kg/s, which gives weight equations with similar trends as the generic weights. The increase in potential consequences is not necessarily too large from 20 to 50 kg/s, as a leak of 20 kg/s will have a large potential to cause damage. But as AK's linear interpolation is done from 50 kg/s and not 20 kg/s, the resulting weight equation estimates will have a more gentle increase.

For AK, like CA, there have not been any DFU 1 incidents during the time period new weight equations have been estimated for. There was however an incident in 2003. The estimated risk contribution this incident would have had if it happened after 2004 have been calculated and is presented in figure 13 together with the generic weight. As it was a small leak, 0,1 kg/s, the generic weight equations will give the highest weight. Year 2008 and year 2009 would give the lowest weights.

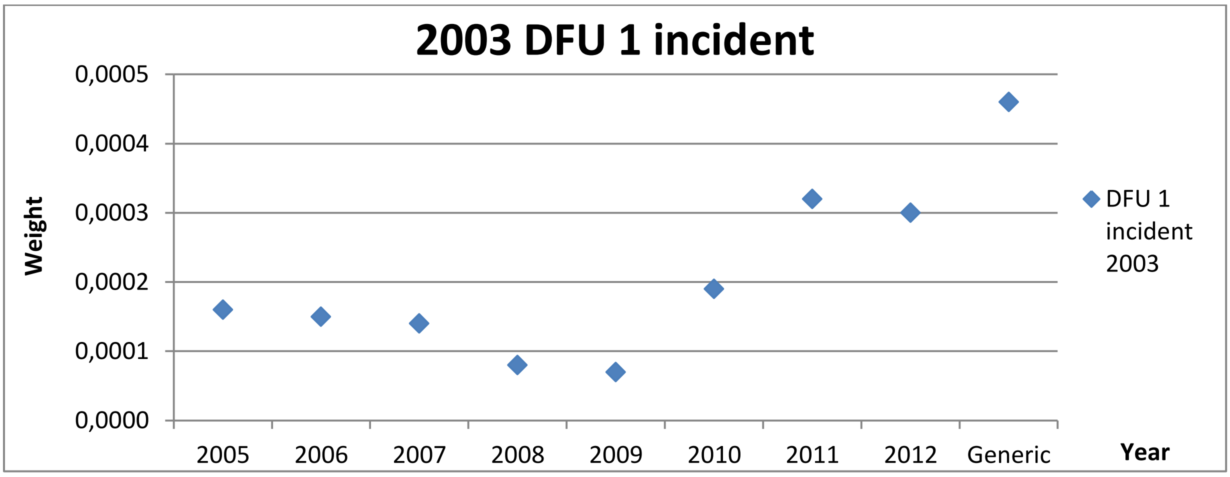


Figure 13 DFU 1, risk from 2003 incident, PLL, AK

5.4. Summing Up an Comparison of Cases

It can be seen that there are some differences in the estimated weight equations. Where CA's weight equation more or less have the same trends as the generic weights, AK's weight equation for leaks between 5,5 and 10 kg/s has a less steep increase. As discussed in chapter 5.3.2, the reason for this is most probably the representative leak rates used in the QRAs.

When comparing the two test cases, interesting properties that undermine the use of generic weights surfaces. For CA, all estimated weight equations give lower risk contribution than the generic weights, whilst for AK the estimated weights varies between better and worse than the generic weights. This shows the potential of having facility specific weights taking barrier functions' performances into consideration.

6. Discussion of Model Development, Assumptions and Limitations

6.1. Barrier Function Failure

6.1.1. Probability Distribution Describing Barrier Function Failure

The exponential distribution has been chosen to describe barrier function failure. The benefit of this distribution is that it is easy to use and have no memory of the past. The challenge with it, is that it might be too simple and not describe the real world sufficiently (Walpole et al., 2012).

Test data is reported for single barrier elements, i.e. tests of detectors are reported for each detector and not of a redundant detector system (PSA, 2010a). Estimation of failure rate and PFD as given in chapter 4.2.2 will give a failure rate and PFD for one single barrier element. The exponential distribution can then be used to describe single barrier elements, series and parallel structures describing the barrier functions.

Alternative distributions that have been considered are Weibull and gamma distributions. Both these distribution are described by two parameters (Walpole et al., 2012). The benefit of using a two parameter distribution is that it allows for more complicated courses of failure to be described. But fitting of a two parameter distribution will be more time-consuming than a one parameter distribution.

Due to limitations in the reported data and the extra time it would take to use a two parameter distributions compared to a one parameter distribution, it has been concluded that the weaknesses of the exponential distribution are acceptable and that the exponential distribution is the best alternative.

6.1.2. Failure Rate

Choice of Failure Rate

Given that failure of the barriers can be described by the exponential distribution, a failure rate has to be established. As presented in chapter 4.2.2, four different failure rate alternatives have been identified. They are as follows:

1. Annual failure rate. As given by equation 13.
2. Average failure rate. As given by equation 14.
3. Total failure rate. As given by equation 15.
4. Three year rolling average. As given by equation 16.

In figure 14 a comparison of the different alternatives can be seen. The plotted dataset can be found in appendix f.

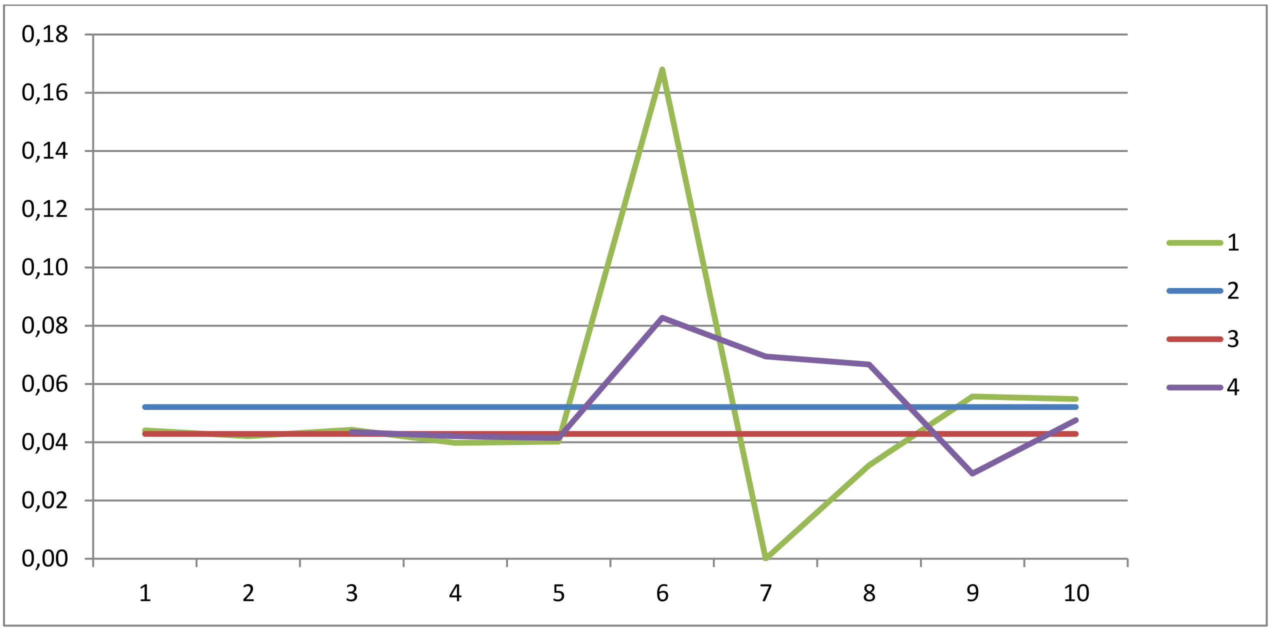


Figure 14 Comparison of failure rate estimates

Failure rate 1 (green) is the annual failure rate. As it can be seen from figure 14, for year 6 and year 7 this failure rate is very sensitive to annual variations. These variations make the annual failure rate unsuited for estimating and describing trends in the risk level. As a consequence of this, it has been chosen not to use annual failure rate to describe barrier element failure.

Both failure rate 2 (blue) and failure rate 3 (red) only give one value and not a time varying failure rate. Failure rate 2 is the arithmetic average of failure rate 1 for all the years data is available. This allows years with extreme values to have a bigger influence than they would if failure rate 3 is used. The latter estimates failure rate as the total amount of failures for all barrier function tests performed. Since failure rate 2 and 3 not are time dependent, it has been decided not to use them to estimate the failure rate describing barrier element failure.

Failure rate 3 can be seen as a version of failure rate 4, three year rolling average, but instead of three year averages it estimates the failure rate as an average of all years, finding an average value that is independent of time. Failure rate 4 on the other hand, will give a time depended failure rate estimate, as long as the observation period is longer than three years.

Compared to failure rate 1, the three year rolling average is less sensitive to fluctuations in the reported values. This is beneficial as one of the goals is to identify trends in the risk level. In figure 14 the difference in sensitivity can be seen. For year 6 and year 7, three year rolling average has less fluctuation than the annual failure rate.

Other arguments for using three year rolling average is that it already is used in the RNNP-project today (PSA, 2011b) and that it will be a measure for avoiding extreme values.

A limitation for failure rate 4 is that it needs more than three years of data reported. This will limit the risk level indications the first years of operation for new facilities. This issue has however not been considered important enough for an alternative method to be developed.

Based on the above discussion it has been chosen to use failure rate 4, three year rolling average, when describing barrier element failure.

Extreme Values of Failure Rate

Two scenarios give extreme failure rate values; no failures when testing and failure when a small number of tests are performed.

A failure rate of zero is unrealistic. Even though a large series of tests shows no failures, the actual real life failure rate for any barrier function relevant for use in offshore oil and gas industry will never be zero. As failure rates will be used in the estimation of the risk level, a failure rate of zero will give a too low picture of the risk.

To solve this issue two solutions have been considered; saying that there always will be at least a half failure or assuming that the next test will be a failure. The first of the two methods are chosen.

If one or more failures occur when only a limited amount of tests are performed, it can appear to be a higher failure rate than what actually is the case. The worst case example of this would be one test and one failure. This problem is minimized by the use of a three year rolling average for barrier functions' performances.

There is still a possibility of only a limited number of tests being performed within the three years, and the estimated failure rate being higher than what really is the case. This has however been ignored, and it is assumed that the use of three year rolling averages will solve the issue with too high failure rate estimates.

Limitations of Failure Rate Caused by Reported Data

There are identified several limitations and weaknesses with the reported data affecting the goodness of the estimated failure rate. These are as follows.

- Barrier functions' performances are reported for the facility as a unit. As a consequence of this, it is not possible to separate failure rates based on the barrier function's location.
- The reported data does not differentiate between types of devices used on the facility to fulfill the same barrier function. E.g. one or several types of gas detectors.

- From the reported data it is not possible to conclude if one or several devices fail more often than others. I.e. if failures are evenly distributed within the test population.
- Test frequency of items is not reported. The reported data does not indicate if some items are tested more often than others.
- Criticality of barrier elements failure is not reported. By criticality it is meant how much redundancy there is for the failed devices. E.g. is the gas detector 1oo1, 1oo3, 2oo3 or 2oo5?

The first three limitations are assumed to be small. The last limitation can partly be solved when weights are estimated; if the facility's QRA includes redundancy this will be partly included in the weights. If the QRA does not take redundancy into consideration, the estimated weights will not be affected by the criticality of barrier elements failure.

As the reported performances is the only data available, it has been concluded that the data must be trusted enough to be used for predicting the risk level. It is however also considered important to be aware of the limitations following from using them.

Limitations from the available data are further addressed in chapter 6.4.1.1.

6.1.3. Probability of Failure on Demand

The long run average PFD is used to estimate the failure rate. For this to be calculated some assumptions needs to be made.

The first assumption is about the test interval, τ. The performance data is reported twice a year, but the data available is only presented on an annual basis (PSA, 2010a). Of this reason the test interval has been sat equal to one year. A weakness with this assumption is that some barrier functions might have been tested several times while others might not have been tested at all. The way barrier functions are reported there is no way to establish the testing rate of the individual barrier elements. A test interval of one year is therefore seen as the only sensible choice.

When the barrier functions are made up of several barrier elements, i.e. 2oo2 or KooN, it is assumed that approximation formulas for PFD given in "System Reliability Theory", table 10.1 (Rausand and Høyland, 2004) can be used. This includes assuming that all PFD values are small, which is a reasonable assumption.

The calculated long run average PFD is as the name says an average. PFD can also be expressed as a time dependent property, with a lower PFD right after testing and a higher PFD at the end of the test interval.

Since establishing a time depended PFD would require more details from the reported data, most probably give an impression of having a higher accuracy than what is the case and the long run average PFD being a well establish and much used method in several industries (Rausand and Høyland, 2004). The long run average PFD is considered to give a sufficient description of the barrier functions' PFD.

Choice of test interval is further addressed in chapter 6.4.1.1.

6.1.4. Correlation between Barrier Functions

The critical value for correlation described in chapter 2.4.1.4 depends on two variables; degrees of freedom and the significance level.

The significance level is decided based on using the same level as already used in the RNNP-project, and is a 90% prediction interval. Degrees of freedom is given by (Ayyub and McCuen, 2003):

$$Degrees\ of\ freedom = N - 2$$

Equation 25

Where *N* equals years.

As the limitations made for this master thesis states that data from year 2003 to and including year 2012 is to be used, and it has been decided to use a three year rolling average to estimate failure rate, the maximum degrees of freedom will only be 6. For facilities newer than 2003, there will be less degrees of freedom, giving even higher critical values. For test case CA, data have only been recorded for a limited amount of years, and the degree of freedom will only be 1.

An important aspect when considering correlation between performances of barrier indicators is that there should be a physical and/or logical link between the barrier functions. I.e. if the risk level is to be affected by correlation, there needs to be some logical explanation of the link between the barrier functions and not just correlation above the critical value. This is supported by the fact that correlation not can prove causality (Wikipedia, 2013). It can only indicate that there can be a causal relationship between two variables.

If correlation is to be taken into consideration when establishing weights, careful assessment of any correlation found to be statistically significant should be performed. The assessment should be done on a facility specific basis to make sure that facility specific properties are accounted for.

Weights are estimated based on the facilities' QRAs which mostly ignore correlation. Allowing for effects of correlation could lead to higher uncertainties for the estimated risk level than if correlations are disregarded. These uncertainties would be further magnified by the fact that there is no good and obvious method for including correlation in the weight estimation process.

The demand for a careful assessment does not necessary mean that there has to be a physical one to one relationship between types of barrier functions to explain a causal relation. Organizational factors, safety culture, safety climate and economy may all explain a link between barrier functions that seems illogical (Øien, 2001a, Olive et al., 2006). These will however be difficult to uncover by only testing correlation between barrier functions and would call for an assessment of each facility's properties.

Vinnem et al. (2010) concludes that both safety climate and noise level are statistically significant correlated to major accident precursors. The same article also concludes that correlation between leak frequencies and barrier performance, serious injuries, falling objects, or age of facility not can be argued for being statistically significant.

The article also discusses randomness as an important factor. Because of the small amount of data, some of the correlation may just be coincidences. Meaning that even though the correlation are higher than the critical value, there is a good chance of this being just a coincidence, a random effect from the fact that there always will be some failures of barrier elements during testing.

Øien and Sklet (2001) has identified that several indicator aggregation methods take interaction between indicators into consideration to create a more detailed and realistic risk picture. Based on this work, not considering or allowing corrections for correlation to be made would imply accepting a less realistic risk picture.

It has been concluded that the uncertainties correlation would bring into the risk level estimate at this stage are so large, that including correlation not can be justified, even though this implies accepting a less detailed risk picture.

6.2. Model Development

When developing a new model for predicting the risk level, certain challenges need to be resolved.

The first issue is whether new weights should be estimated for all DFUs. Barrier functions are only allocated to some of the major accident precursors by ”Metoderapporten” (PSA, 2011b). It is decided to use this allocation as a guide for which DFUs new weights are to be estimated for. The argument for this decision is that the aim of this master thesis is to develop a model that include both observed barrier functions' performances and major accident precursors. It will then be of little relevance to estimate weights for major accident precursors that does not have relevant observed barrier functions' performances reported.

It is however decided to make one exemption from the "Metoderapporten" allocation; DFU 9 is assigned barrier functions as specified in chapter 6.2.2. This is done as several of the reported barrier functions are relevant for stopping/mitigating consequences from the major accident precursor leak from risers, pipelines and production systems.

It can be argued that relevant barrier functions are reported for DFU 4 as well, but DFU 4 will be omitted from the new weight estimation model. The estimation procedure will partly be based on the facilities' QRAs, and only a few QRAs includes this scenario (PSA, 2011b). Not being able to estimate new weights for DFU 4 by neither method 1 nor method 2 for the greater part of the NCS, most facilities would still have to use the generic weights. Estimating new weights for only some facilities can then cause a biased risk estimate. Conclusion is therefore made that no effort should be put into estimating new weights for DFU 4.

For some of the major accident precursor scenarios, generic weights are not considered to be sufficient to estimate the risk level. In today's model these scenarios will be individually assessed if and when they occur. The new weight estimation model will follow the same procedure. I.e. no new weights will be estimated for the scenarios generic weights are individually assessed for. In case of occurrence of any of these scenarios, an individual assessment will be performed for the new model as well.

The generic weights split some of the DFUs into ranges and categories. E.g. DFU 1 is calculated for 0,1 to 5,5 kg/s, 5,5 to 10 kg/s or larger than 10 kg/s and DFU 3 is divided into five categories. As no major arguments are found for using other distinction for weights estimated by the new model, today's distinction will be used as far as possible.

For DFU 3 there has however been identified some issues with using the same weight categories. As shallow gas not is relevant for all facilities, new weights will only be estimated for shallow gas when this is relevant. Furthermore some facilities' QRAs are found to only contain enough information for one general DFU 3 weight to be estimated. Where this is the case, categories used for the generic weights are put aside and only one general DFU 3 weight estimated.

For some facilities not all major accident precursors are relevant. As there are no reason to estimate weights that never will be used, no weights will be estimated when the corresponding major accident precursors not are relevant.

6.2.1. Choice of Method and Data from Quantitative Risk Analysis

In chapter 4.3 the two weight estimation methods are described. They estimate weights in two different ways, but both method estimates facility specific weights based on data from QRAs and observed perfomance of barrier functions.

Method 1, updating QRA and then estimate weights, requires QRAs with a high degree of detail as part of the QRAs' calculations has to be reproduced with the observed barrier functions' performances. This limits the range of application only to facilities with detailed enough QRAs available for part of the analysis to be reproduced.

Reproducing parts of the QRA as part of establishing weights means erroneous and outdated perfomance assumptions are "corrected" with the observed performances of barrier functions. This gives the most facility specific weight estimates, as no other QRAs are included and will give the best possibility of estimating correct weights and predicting the actual risk level.

Method 2, establishing weights and adjusting with correction factors, is a solution for facilities not having a detailed enough QRA available for method 1 to be used. All this method requires from the QRA, is that it is possible to estimate weights based on the presented results. No reproduction of the calculation and results are done. Influences from barrier functions' performances are included by correction factors, established based on differences between observed performances and benchmark values.

A weakness of method 2 is the correction of weights. The adjustment is based on calibration factors estimated using QRAs from other facilities, detailed enough for method 1 to be used. If the correction factor calibration was perfect, method 2 would be equally good as method 1. Achieving this is however not very likely, even if the calibration is based on similar facilities. The use of imperfect calibration based on other facilities' QRAs, will give weights less specific to the facility. This gives a higher probability of the estimated risk level being incorrect.

Another benefit of method 2 is that it will be less time-consuming. This is due to the fact that once the general calibration factors have been established; they can be used for all facilities.

Perhaps one of the most important limitations for method 1 is that it will be time consuming. The procedure for estimating weights from the QRA is the same for both methods, but method 1 includes a time demanding update of the risk analysis before estimating weights, whereas method 2 first estimates weights from the original QRA and then simply adjust these weights by correction factors.

From the preceding discussion it can clearly be seen that the most correct weight estimates are achieved from method 1. Method 2 will also give estimates of the risk level influenced by observed barrier functions' performances, but not as facility specific as method 1. As a most correct and facility specific risk level estimation is desired, method 1 will be the preferred method in this master thesis.

6.2.2. Assigning Barrier Functions to DFUs

"Metoderapporten" (PSA, 2011b) gives an overview of what performances that are measured for which barrier functions, and which DFU it prevents/mitigates. This overview has been used as a basis for the establishment of facility specific weight, and is reproduced in appendix b with some modifications.

When weights are estimated by method 1 and the QRA uses other barrier functions than those assigned to the DFUs by appendix b, the barrier function allocation from the QRA will be followed. E.g. if the QRA use the barrier function fire detection to estimate the risk for DFU 1, method 1 will update performance of fire detection for the DFU 1 weight estimation as well.

The allocation of barrier functions given by "Metoderapporten" has certain limitations. The perhaps most severe limitation is to only assign each barrier function to one DFU. E.g. gas detection is only assigned to DFU 1 and not DFU 2 and DFU 3 as well. Another limitation is that no performance indicators are assigned to DFU 9 and DFU 10.

This chapter will give a discussion on the fitness of the assigned barrier functions' performances to estimate weights for each DFU, and assign barrier functions to DFU 9.

DFU 1 and DFU 2

DFU 1 and DFU 2 are closely linked. DFU 2 can be viewed as a successor of DFU 1, an escalation of an already dangerous situation. There are not established any generic weights for DFU 2 as there never have been any need for it. As a consequence of this, the limitations of this master thesis states that no weights will be established for DFU 2. Due to the way barrier functions' performances are chosen to affect the risk level, not establishing weights for DFU 2 will lead to three of the reported barrier functions' performances not affecting the estimated risk level when strictly following the allocation of barrier functions done by "Metoderapporten".

An approach that could have been used to include the three barrier functions is to let them influence the weights for DFU 1. It has however been decided that the allocation given by "Metoderapporten" is to be followed as much as possible. This is based on the assumption that not following "Metoderapporten" would be similar to claiming that "Metoderapporten", which the RNNP-project is based on, is wrong and would require strong arguments.

The issue with the three performance indicators for DFU 2 not being allowed to affect the estimated risk level is partly solved by assigning performance indicators to DFU 9. For further discussion see paragraph about DFU 9.

For DFU 1 the following barrier functions' performances are used:

- *Availability of gas detection*.

  HC leaks are discovered either by personnel or by gas detectors. The time it takes for a leak to be discovered, will affect the risk level inflicted by the leak incident. Gas detectors performance indicates the detectors ability to signal gas detection when tested (PSA, 2010a). If the barrier function's performance is poor, it is an indication that there is an increased risk of detectors failing on demand. A detector failing on demand means that gas has to disperse further before detection occurs, increasing the risk of fire and explosion as more ignitions sources can be exposed.
- *Availability of blow down valve.*

  When an incident occurs, blow down of process pipes and equipment is an important measure to limit the potential consequences (NORSOK, 2008). The availability of BDV is a measure of the valves ability to open within the specified time limit during testing (PSA, 2010a). If the BDVs use too long time or fails to open on demand, pressurized HC will be available in the vicinity of the incident, increasing the potential consequences following a leak.
- *Availability of process safety valve.*

  PSVs are installed to avoid overpressurization of process equipment that could lead to a HC leak. The barrier function's performance is tested by checking that the PSV opens at a certain overpressure (PSA, 2010a) to bleed off HC to a safe area.

  Too high pressure in pipes and equipment could potentially lead to ruptures and subsequent HC leaks, increasing the risk of a DFU 1 incident occurring.
- *Closure and leak test of riser emergency shutdown valve.*

  Riser ESDVs are installed to give the possibility to stop the flow of HC before it reaches topside. Available HC topside is an important factor for the risk imposed by a DFU 1 incident. Performance of the barrier function riser ESDV measures the ability to close the valves and leakages when the valves are closed (PSA, 2010a). It will give an indication on whether the ESDV will be able to perform its task on demand and stop the flow of HC when activated.

  Failure of riser ESDV can potentially give more HC topside, increasing the risk level following a DFU 1 incident.

- *Closure and leak test of wing and master valve.*

  Wing and master valves are closed to stop HC flow from the Christmas tree. Failure of these valves can lead to increased amounts of HC topside and thereby increasing the risk level after an HC leak incident, much in the same manner as a riser ESDV failure.

Based on the preceding arguments, all barrier functions' performances allocated to DFU 1 by the "Metoderapporten", will be used as influencing factors when estimating new weights for the major accident precursor non-ignited HC leaks.

DFU 3

DFU 3 is not relevant for all facilities. Most floating production facilities will have all wells located subsea at large enough distance away for a well event not to cause major accident risk. This should be assessed for all facilities before the weight estimation process is started.

For DFU 3 the following performance indicators are allocated.

- *Availability of downhole safety valve.*

  In case of loss of well control the DHSV will allow the HC flow to be stopped before it reaches the BOP. The performance of DHSVs gives the valves observed ability to stop leakage of HC when closed (PSA, 2010a).

  Failure to seal off the well stream before the BOP will increase the probability of HC reaching the surface, creating a fire and explosion risk for the facility following a well event.

- *BOP barrier functions performances.*

  The other barrier functions allocated to DFU 3 depends on the location of the wellheads; i.e. subsea or topside. The categories are the same for both locations of the wellheads, they are as follows.

  - *Availability of drilling BOP*
  - *Availability of coiled tube BOP*
  - *Availability of pressure pipes BOP*
  - *Availability of wireline operation BOP*

  Failure of barrier functions are defined as leakage or malfunction of valves (PSA, 2010a, NORSOK, 2004); i.e. not being able to open or close on demand. All four performances for the BOP's barrier functions give indication on the risk, given a well event. Reports on BOP barrier functions are however of limited quality as most of the test data are kept by the drilling contractors (PSA, 2012a). And, there are no benchmarks to compare the performance of these indicators against.

It has been decided that only the barrier function DHSV will influence the estimated weights for DFU 3. BOP indicators could have been used when estimating weights by method 1, it has however been decided not to include BOP indicators as the quality of the reported data is varied and limited.

DFU 8

The barrier functions allocated to DFU 8 by "Metoderapporten" are only related to marine systems. Barrier functions are therefor only reported for floating production units and mobile installations. The three performance indicators are as follows.

- *Availability of watertight doors.*

  The watertight doors are considered to be available when they are able to close and lock in closed position within specified time (PSA, 2010a). The time to close the door is given by §39 in the stability regulation given by the Norwegian Maritime Authority (1991), but there are no benchmarks values to compare the observed availability with.

- *Availability of ballast system valves.*

  The failure mechanism of ballast system valves is defined as not being able to open or close within specified time (PSA, 2010a). In a meeting 3rd of April Professor J. E. Vinnem informed that criteria for successful tests are based on the operator's demands for the specific facility. These demands are based on the performance standards for barriers, required established by the management regulation.

  In a conversation on 2nd of April 2013, Principal Engineer M. R. Stensland, Det Norkse Veritas, informed that the time necessary for closing of ballast valves, the closing characteristic, depends on the valve's size and type and actuators used. When asked about if there are any regulations for time to open or close ballast system valves similar to the ones for watertight doors, he said that there were no regulations that he was aware of for single valves. Stensland however emphasizes that there are regulations related to performance of the entire ballast systems being able to get the facility back on an even keel after an incident within a given time limit. He also mentions another issue with testing of ballast valves; testing of their failsafe mechanisms, fail to closed position.

  Failsafe mechanisms are not mentioned in the demands for reporting of barrier functions (PSA, 2010a). The reported test data for ballast system valves does not contain any information on whether the performed tests are failsafe tests or closing and/or opening tests. Based on the nature of the reported data, it is not possible to establish separate performance indicators for testing of failsafe function and ability to perform on demand. It has therefore been concluded that no interpretations with regard to what kind of tests the reported data comes from should be done.

There are no benchmarks values for the availability of ballast system valves.

- *Metacentric height.*

  Metacentric heights (GM) are reported at the turn of the year for semi-submersible facilities and facilities shaped as ships (PSA, 2010a). The GM-values is not suited for comparison with a benchmark value, because the reported data will be good or bad based on the properties of the facility and not strictly on a numerical comparison. However, higher GM-values generally indicate better stability.

  The GM-value depends to a large degree on changing circumstances, such as ballast condition, payload and type of operation, and not just on the facility's fixed properties.

  Since the GM-values only are reported once a year, the impact of the reported value on the estimated risk level for the entire reporting interval is questionable. At best, it must be said that the GM-value gives a snapshot of the risk level influence at the turn of the year.

Based on the above arguments, it is concluded that GM-values are unsuited as an influencing factor when establishing weights for DFU 8, mostly because performance only is reported once a year.

The two other performance indicators on the other hand are found possible to include, but only when method 1 is used for estimating weights.

None of the three performance indicators allocated to DFU 8 by "Metoderapporten" are reported for fixed facilities, complexes or normally unmanned facilities. Weights for these three types of facilities will therefore not be influenced by the performance of barrier functions so no facility specific weights will be estimated.

DFU 9

One of the main limitations of the allocation from "Metoderapporten" is that it does not assign any performance indicators to DFU 9 and DFU 10. After consultation with Professor J. E. Vinnem, it has been concluded that some of the reported barrier functions should be used to estimate the risk imposed from a DFU 9 incidents as well. Because of the limitations made in this master thesis, no weights will be established for DFU 10, neither will any performance indicators be assigned.

For DFU 9, the barrier functions amongst other include the barrier functions assigned to DFU 2 by “Metoderapporten”; fire detection, fire pumps and deluge valves. As no weights are established for DFU 2, assigning these performance indicators to DFU 9 will give the benefit of allowing them to still influence the estimated risk level. This eliminates the limitations encountered when not establishing weights for DFU 2 incidents and not assigning DFU 2’s barrier functions to DFU 1.

The two other barrier functions included, gas detection and riser ESDV, are assigned to DFU 1 as well. This means that they are allowed to influence the risk level twice, i.e. both for DFU 1 incidents and for DFU 9 incidents. As they in real life also influences risk imposed from both types of incidents they will be allowed to affect the estimated risk level for both types of incidents, without any kinds of weighing or reduction in influence on the risk.

Arguments for letting performance of these barrier functions affect the estimated risk from a DFU 9 incident are as follows.

- *Availability of fire detection.*

  One of the main consequences with potential to cause a major accident is fire. Early detection of fires will allow the deluge system to activate, process blow down to start and activation of general alarm (NORSOK, 2008). The performance indicator for fire detectors are based on whether the detectors signals detection or not during testing (PSA, 2010a). Poor performance of fire detectors during testing would indicate that it would take longer time for an actual fire to be detected, giving an increased risk level.

- *Availability of deluge valves.*

  If fires occur, the deluge system will try to limit the damages by stopping or slowing down the fire escalation (NORSOK, 2008). As the capacity of the water supply will limit the size of the fire that can be combated (Buchanan, 2001), early activation and maximum flow capacity are important factors for the efficiency of the deluge system.

  The performance of deluge valves is their ability to open on demand (PSA, 2010a). If the deluge valves fail to open, it will significantly increase the potential consequences of a fire following a riser or pipeline leakage.

- *Availability of fire pumps.*

  Fire pumps are essential to combat fires with deluge. The fire pumps’ performance is their ability to deliver fire water following the first attempt to start the pump (PSA, 2010a). High unavailability of fire pumps will result in lower probability of enough fire water being delivered to the deluge system. As mentioned earlier, the fire size the deluge systems potentially can handle depends on the maximum water capacity. Poor performance of fire pumps during testing will indicate that the system might not be able to deliver sufficient fire water to the deluge system on demand. This can result in the deluge system not being able to combat the fires it is designed for, increasing the risk level imposed on the facility following a DFU 9 incident.

- *Availability of gas detection.*

  To avoid fires following a leakage, early blow down, isolation of HC flow and isolation of ignition sources are important measures. In order for these measures to be initiated, the HC gas leak needs to be detected (NORSOK, 2008).

  Gas detectors' performance measures the detectors ability to signal detection during testing (PSA, 2010a). If the performance indicator shows low ability to signal during testing, it can be an indication that gas detectors will fail upon an actual demand. Failure on demand can increase time before detection and allow the gas plume to diffuse further into the facility giving an increased risk from explosion and/or fire following a DFU 9 incident.

- *Availability of riser ESDV.*

  When riser ESDV is activated it will close down the flow of HC into the riser. Stopping the flow of HC will limit the potential gas plume size and reduce the risk significantly.

  There are two types of performance indicators for riser ESDV; availability of closing and leakage when closed (PSA, 2010a). Both types of failure modes will give an increased potential of HC gas volume, leading to a higher risk from a DFU 9 incident.

Based on the above arguments, it has been decided that these barrier functions' performances are to be used as influencing factors when estimating weights for DFU 9.

6.2.3. Method 1 Methodology

Method 1 is based on reproducing part of the QRA and estimate weights by the same procedure as used for the generic weights.

The methodology used in method 1 is mostly standard risk analysis methods, and will depend on the QRA being reproduced. For test case CA this means that event tree analysis is used. Event tree analysis is a well-established method and no further presentation of the analysis method will be given.

As no new methodology is developed for method 1, no further discussion of methodology is given.

6.2.4. Method 2 Methodology

Where no new methodology are developed for method 1, weight estimation by method 2 are based on a method developed in this master thesis and some discussion and clarification is therefore needed. This chapter will discuss the elements of the weight estimation process.

6.2.4.1. Benchmarks

As described in chapter 4.3.3, barrier functions' performances are compared with benchmarks values. These benchmarks values are taken from the RNNP-reports (PSA, 2012a) and are reproduced in appendix c.

In the RNNP-reports these benchmarks are referred to both as industry standards and Statoil's internal guidelines for barrier functions performance. After a conversation with Professor J. E. Vinnem 8th of Mars 2013, it became clear that these benchmarks are taken from Statoil's internal guidelines, but that they are more or less the same for all operators on NCS.

It could be argued that using only one set of benchmarks can result in erroneous correction factors for facilities not having the same internal guidelines as Statoil. The time it takes to establish separate benchmarks for all operators and/or facilities not being operated by Statoil, and the added complexity this would introduce to method 2, are not considered to make up for the benefits of removing the potential errors following from only using one set of benchmarks. Therefore it has been decided that Statoil's internal guidelines will be used as benchmarks for all facilities.

An alternative approach to benchmarks could be to use trend analysis of barrier functions' performances to estimate changes in the risk level. In the 2011 RNNP main report (PSA, 2012a) this is suggested done for the barrier functions for BOP. The argument for using trend analysis instead of benchmarks is that availability demands not are found to be suited for use on BOP barrier functions.

A combination of using trends and benchmarks could have been to use benchmarks for all barrier functions this is defined for and trend analysis for the rest. The combination would allow performance of the BOP barrier functions, watertight doors and ballast valves to be used as influencing factors on the estimated risk when using method 2.

One of the main benefits of method 2 is that it is easy to use. Partly or fully implementation of trend analysis would increase the complexity of the method. Combined with poor quality of the reported data for many of the facilities' BOP barrier functions' performances (PSA, 2012a), it has been decided that only benchmarks are to be used at this stage.

6.2.4.2. Deviation Ratio

Deviation ratio is a measurement established to express how well a facility's barrier functions performs relative to the benchmark values defined in appendix c.

The measurement assumes that the facilities' availability demands are the same as the benchmark values. If however the facility's safety philosophy uses other availability demands for their barrier functions, the calculated deviation ratio will lead to wrong weight estimates.

As stated in the previous chapter, benchmark values are originally based on Statoil's internal guidelines, but these are more or less the same for the entire industry. Method 2 is used when not too much information is available, and it therefore is a good possibility that the barrier functions' assumed availability is unknown. Using the same availability for all facilities, i.e. the benchmark values, is therefore considered to be the best solution.

The acceptable range of the deviation ratio has to do with what barrier function performances that will be accepted as reasonable. As stated in chapter 4.2.2, a half failure will be assumed if there are no observed failures in the reported data. As this avoids a failure rate of zero, all deviation ratios lower than zero will be accepted.

Like the lower limit, all upper values of the deviation ratio will also be accepted. This is based on the following arguments.

The upper limit of the deviation ratio will be achieved if all reported tests are failures. A three year rolling average is used to avoid annual fluctuations in the failure rate following from failures on only a few performed tests. A high failure rate is therefore assumed to indicate that there actually is poor barrier function performance.

And calibration of the correction factors, as described further down, is intended to avoid correction factors of unreasonably high values, allowing all values of deviations ratios to be accepted.

6.2.4.3. Linearity

All calculations in method 2 are linear. This gives a simple and easy to use calculation process. There are however certain issues with this assumption.

If the facility is designed according to NORSOK S-001 (NORSOK, 2008), there will be implemented redundancy in some of the barrier functions' configurations. This means it typically will be a KooN configuration of the barrier elements. Ideally the failure distribution used to estimate the risk level should then include both serial and parallel structure. The deviation ratio will however only consider a single barrier element's performance when estimating the deviation, taking neither parallel nor serial structures into consideration.

Not all barrier functions are relevant for every terminal event. E.g. if there is an immediate ignition after a HC leak starts, gas detection will not have an impact on expected fatalities. As the correction factors adjust the sum of the terminal events, and not each terminal event separately, they will not take this into account. I.e. the deviation ratio ignores that some barrier functions might not have an impact on all possible outcomes.

If any of the two preceding issues were to be taken into consideration, more data than what method 2 is meant to be based on would have to be available. As this not is aim of method 2, these issues are assumed not to be too critical, and linearity will be used.

6.2.4.4. Correction Factors

Correction factors are used to adjust the weights estimated from the QRA based on the barrier functions' annual performances. There are two types of correction factors; one for barrier functions and one for weights. The weight correction factors are calculated as an arithmetic average of the barrier function correction factors.

Since weight correction factors are based on more than one barrier function correction factor, poor performance of one of the barrier functions will not necessarily result in a correspondingly larger weight correction factor and vice versa.

This is considered both to be a positive and negative property of method 2. It is negative in the sense that critical barrier functions do not have higher influence on the weight correction factor than less critical barrier functions. The negative consequences will be largest if the most critical barrier functions perform poorly whilst the rest of the barrier functions perform very well.

On the other hand, the expected consequences of a major accident precursor may be reduced if most of the barrier functions perform well, whilst one performs poorly. Not allowing a single barrier function correction factor to disproportionately influence the estimated weights will express this.

The positive and negative consequences are considered to make up for each other, so the weight correction factor estimation procedure suggested is considered to be the best solution.

Estimating weights by the use of correction factors based on general calibration factors will never be equivalent to updating parts of the risk analysis, like done for method 1. This will be a consequence of barrier functions' performances not having the same influence on the causal chain as they would in for example an event tree. Another effect of not influencing the casual chain is that redundancy not can be properly accounted for by correction factors.

Weights estimated by the use of correction factors will however be an improvement from generic weights, as they take into account performance of barrier functions.

A correction factor lower than – 1 (negative 1), would indicate that the occurrence of a major accident precursor would decrease the risk level. This is unreasonable and can never be the case. The accepted lower range for correction factors will therefore be no less than or equal to -1 (negative 1).

For the upper range of the correction factor, there will be limitations coming from the maximum possible expected fatalities after the occurrence of a major accident. I.e. there can never be more fatalities than there are people onboard.

When correction factors are calculated, the calibration factors should ideally account for the upper acceptable limit. As there are no guaranties that this is the case, quality assurance should be implemented in the form of controlling the estimated weights, to make sure they are within acceptable limits.

All correction factors that give weights with expected fatalities less than people onboard will be accepted. This is based on that the calibration process should give a link between deviation ratios and correction factors that is assumed to be reasonable and not approaching an unrealistic value. And that a very high failure rate, when a three year rolling average is used to avoid unrealistically high failure rates, will indicate poor barrier function performance and a following severe increase in the risk level.

6.2.4.5. Calibration Factors

To express the link between deviation ratios and barrier function correction factors, calibration factors are used. The calibration factors can be viewed as a measure of how important changes in the observed barrier functions' performances are for the estimated risk level.

Estimation of Calibration Factors

As not all values of correction factors are possible, certain limitations are given for the calibration factors.

The first limitation is that observed barrier function performance equal to the benchmark value, indicates that the perfomance is as expected. A deviation ratio equal to zero will therefore give a calibration factor of zero, as no adjustments of risk level estimated from the QRA should be made.

The second limitation is the lower accepted value of correction factors. The reduction factor can never be lower than or equal to -1 (negative 1). As mentioned earlier, a lower value than

this would give reduction of the risk level following an incident. The estimated calibration factors can therefore not give correction factors equal to or lower than -1 (negative 1).
The last limitation is the upper limit. Absolute upper limit the weights can take on is the potential loss of all personnel onboard the facility. The calibration factors will therefore be absolutely restricted upwards by the number of personnel onboard.
Even if barrier functions' observed failure rate approaches 1, there will most often be other barrier functions in place that performance not is reported for. This means that total loss following poor observed barrier functions' performances and an incident is highly unlikely.
As total loss might be a consequence given lousy performance of critical barrier function, it will be accepted. But if the weights approach total loss, individual focus should be put on each barrier function's correction factor to see if individual adjustments should be made.
Given that linearity is used when estimating weights by method 2, the question of symmetry of calibration factors must be asked. I.e. if calibration factors are the same for both positive and negative deviation ratios for the same barrier function. Based on that the range of the negative deviation ratios is -1 (negative 1) to zero and positive deviation ratio has a range from zero to 199. There are no reasons to assume that symmetry will be the case. Calibration done in test case 1 also supports the argument of no symmetry. Calibration factors should therefore be estimated both for positive and negative deviation ratios.

Test Case Calibration

Calibration factors have to be estimated for positive and negative deviation ratios for each barrier function. In order for this to be done for test case 1 – CA, the three years failure rate is estimated for would have to give both positive and negative deviation ratios for all reported barrier functions. As this not is the case, it has been impossible to estimate calibration factors for both magnification factors and reduction factors for all barrier functions.
The calibration is based on trial and error. It is first done for each year failure rate has been estimated for, and then arithmetic average is taken of the annual calibration factors to find the facility calibration factors. This process is continued until the best fit of correction factors is achieved.
As the calibration factors are established by adjusting one and one calibration factor until the necessary correction factor is achieved, there are no guaranties that the calibration factors established are the "correct" ones. The fact that only one test case is calibrated magnifies this potential source of error. As discussed in the next section, this issue will be smaller as more facilities are calibrated and general calibration factors are estimated.

The calibration factors established from test case 1 are used to estimate weights by method 2. For test case 2 –AK this brings two new challenges; not all calibration factors have been estimated and using a FPU to calibrate a fixed facility.

The lack of calibration factors are solved by assuming symmetry. As addressed in the previous section, symmetry cannot necessarily be assumed. But the only alternative approach identified, is to use arithmetic average of the other positive calibration factors to estimate calibration factors for the barrier functions PSV and wing and master valve, and arithmetic average of the negative calibration factors to estimate negative calibration factor for riser ESDV. As this approach includes the use of other barrier functions' performances, it has been concluded that the best approach is to use symmetry, which only depend on the actual barrier function.

Calibration of a fixed facility by a FPU does not necessarily give correct results. But as the next section about general calibration factors concludes that only one set of calibration factors is to be used for all facilities, it is assumed that using a FPU to calibrate a fixed facility will give a good enough risk level estimate.

General Calibration Factors

The calibration factors estimated for a single facility will only indicate the impact on the estimated risk from changes in barrier functions' performances for the specific facility. Since there are differences between facilities on the NCS, calibration factors estimated for one facility might not describe all other facilities sufficiently well.

Because of this, general calibration factors are established based on calibration factors from several facilities. This is done to avoid facility specific properties having to big impact on the estimated risk level.

The use of general calibration factors is what makes weight estimated by method 2 less facility specific. Instead of only looking at the impact on the risk level from the specific facility's barrier functions' performances, the impact on other facilities risk level from their barrier functions' performances are used to estimate changes in the risk level.

An alternative approach could have been to establish general calibration factors for each of the five facility types used in the RNNP-project. This would give a lot of extra work and it would demand a sufficient amount of detailed QRAs to be available for weights to be estimate by method 1 for five categories. This can be an issue and it is decided that only one set of general correction factors will be used.

This brings the discussion to the next issue. How many facilities should be used to estimate the general calibration factors? As an absolute minimum there needs to be included so many facilities that both positive and negative calibration factors can be estimated for all barrier functions. This minimum has its basis in that assuming symmetry not is considered a sufficient solution.

The upper amount of facilities the calibration factors should be based on will depend on when convergence is achieved. I.e. when change in arithmetic average no longer is noteworthy as more facilities are added.

The use of general correction factors makes the estimated weight less facility specific. A consequence of this is that the estimated correction factors might give unrealistic weights. This issue should not cause too severe troubles, but all estimated weights should be controlled to ensure they are realistic.

6.2.5. Comparing Results

The intention of the comparison is to assist in the development of a new weight estimation model and to monitor annual changes. The comparison will indicate to which degree the weights are changed and if the potential consequences are worse or better than estimated by the generic weights. The comparison will also reveal changes in risk level from year to year.

The comparison will be an alternative approach to today's presentation of barrier functions' performances. It will not only give the performances, but as weights are estimated it will be a measurement of the consequences changes in barrier functions' performances could have on the risk level.

As major accident precursors are rare, a general comparison should be made. One incident is therefore assigned to each new weight, and the resulting risk level is compared. This comparison does not represent the risk level actually caused by occurrences of major accident precursors, but it is a way to estimate the potential risk level.

6.3. Test Case Assumptions

Weights will only be estimated for DFU 1. This is a consequence of the data available from CA's QRA not being detailed enough to use method 1 for DFU 9. This will be further addressed in chapter 6.4.2.1.

6.3.1. Assumptions for CA

Barrier functions as listed in table 12 are used to estimate weights by method 1 for CA.

Table 12 Method 1, performance indicators for facility CA

Major accident precursor	Barrier functions
DFU 1: Non-ignited hydrocarbon leaks	Availability of gas detection
	Availability of BDV
	Availability of deluge valves
	Availability of fire detection

And, barrier functions as listed in table 13 are used to estimate weights by method 2 for CA

Table 13 Method 2, performance indicators for facility CA

Major accident precursor	Barrier functions
DFU 1: Non-ignited hydrocarbon leaks	Availability of gas detection
	Availability of BDV
	Availability of PSV
	Availability of riser ESDV
	Availability of wing and master valve

6.3.1.1. Method 1

When reproducing CA's QRA calculations with the observed barrier functions performances, certain assumptions has to be made. This chapter will give a description of these assumptions.

DFU 1

The following assumptions are made when updating the event tree's node probabilities.

- *Fire detection.*

  The event tree analysis for CA uses two configuration of the fire detectors; 1oo1 or 2oo2. When there is a 2oo2 configuration, the QRA express PFD as two times the PFD for one detector. This is similar to the conclusion on PFD for series systems in "System Reliability Theory" (Rausand and Høyland, 2004, page 431). When reproducing the QRA's calculations, the same configurations and method for expressing the PFD for 2oo2 systems have been used.

  In CA's event tree analysis, fire detector redundancy is not accounted for, i.e. there are no places where there are 1oo2 or similar configurations. As the reproduction of the calculations only are supposed to update the QRA results, the same procedure is followed.

Not including redundancy is a questionable assumption, but as this is done in the original QRA, it will be assumed to be good enough for the reproduction. Not including redundancy will be discussed in chapter 6.4. This chapter will only conclude that the same procedure as in the QRA will be followed.

- *Gas detection.*

  CA's QRA states that successful gas detection demands two detectors signaling detection. In the event tree analysis this is accounted for by a gas detection configuration of 2oo2 for all scenarios except "Stern discharge system on poop deck", which will be discussed separately. Like fire detectors, the same procedure as used in the QRA will be used when estimating the probability of successful gas detection. I.e. PFD for two detectors are expressed as two times the PFD for one detector.

  The original event tree analysis does not include redundancy for gas detectors. The reproduction of the QRA with observed barrier functions' performances will therefore not include redundancy either. For further discussion, see chapter 6.4.

  For the scenario "Stern discharge system on poop deck", the event tree analysis uses a different probability of successful gas detection than for the other scenarios. Linear interpolation, as used for deluge valves, could have been used to estimate the PFD, but as no assumptions are given in the QRA for the choice of PFD value, the PFD from the original QRA will be used in the event tree reproduction.

- *Blow down valves.*

  For all scenarios except "Stern discharge system on poop deck" and "VOC on vessel upper deck", the configuration of BDV is 1oo1. The same assumption has been used for the reconstruction of the event tree analysis. This includes that no redundancy is included, which is the same as for the preceding barrier functions.

  For the scenarios "Stern discharge system on poop deck" and "VOC on vessel upper deck" the original event tree analysis uses a PFD equal to 1, without giving any arguments for this in the QRA. As the assumptions not are given in the QRA, the reproduction of the event trees with observed barrier functions' performances will use the same PFD as the original analysis.

- *Deluge valves.*

  For all scenarios other than "Methanol tank on top of TP30", "Stern discharge system on poop deck" and "VOC on vessel upper deck", CA's QRA states that to include that more than one deluge valve has to work, and account for possible blockages in the deluge system, PFD of the deluge system is set equal to 0,1. It also states that when the fire escalates to other areas, the PFD of the deluge system is equal to 1. When the latter is the case, PFD of deluge system will be set equal to 1 for the reconstruction of the event tree analysis as well.

When the deluge system has a PFD equal to 0,1, linear interpolation is used to adjust the assumed PFD from the original QRA with the observed barrier functions' performances.

To use linear interpolation a second interpolation point beside the point assumed from the original QRA has to be established. System PFD will be equal to 1 when PFD of single detectors equals 1. This will be used as the second interpolation point.

The interpolation formula for deluge system PFD is expressed in equation 26.

$$PFD_{System,Observed} = PFD_{System,QRA} + \left[\frac{\left(PFD_{System,Assumed} - PFD_{System,QRA}\right)}{\left(PFD_{1\ valve,Assumed} - PFD_{1\ valve,QRA}\right)} * \left(PFD_{1\ valve,Observed} - PFD_{1\ valve,QRA}\right)\right]$$

Equation 26

Where $PFD_{System,\ Assumed}$ and $PFD_{1\ valve,\ Assumed}$ is the assumed interpolation point, $PFD_{System,\ QRA}$ and $PFD_{1\ valve,\ QRA}$ is the assumed PFD values from the QRA and $PFD_{1\ valve,\ observed}$ is the observed PFD value for a barrier element.

Ideally equation 26 would have taken into account blockage of deluge system, redundancy and nonlinear effects caused by redundancy. The nonlinear effects will never be accounted for when using a linear interpolation approach. For observed barrier function's performances close to assumed performance in QRA this should not be a problem. But if there is large deviation between observed and assumed perfomance, nonlinear effects will have larger impacts on the result. This will be a weakness of this approach.

The QRA does not mention anything about redundancy. If the system PFD used by the QRA includes redundancy, the interpolation formula will also partly include redundancy. As the aim only is to reproduce the QRA's analysis, the uncertainty about redundancy being included is ignored.

The QRA states that part of the reason why the system PFD value is as low as it is, is because of possible deluge system blockage. From this it follows that the interpolation formula not should give system PFD equal to 0 when the observed PFD equals zero. This criterion is fulfilled. It should however be noticed that the failure rate never will be zero, see chapter 4.2.2.

As a reproduction of system PFD not is possible from information in the QRA. It has been concluded, even with the limitation as mentioned above, that the linear interpolation formula is the best approach for updating the node probability with observed barrier function's performances.

For the scenarios "Methanol tank on top of TP30", "Stern discharge system on poop deck" and "VOC on vessel upper deck", there are used other PFDs than for the rest of the 54 scenarios. For the reproduction this gives two cases; PFD equal to 0,01 and PFD equal to 0,05. The first case is assumed to be the same as only one deluge valve PFD, and the observed PFD for one single valve is used. For the latter case, the PFD used in the original event tree analysis is used, as no assumptions are given in the QRA.

CA's QRA has different leak categories for gas and for liquid HC. When weights are estimated, representative leak sizes are assumed to be the same as for gas leaks. This is done since most of the scenarios are gas leaks.

6.3.1.2. Method 2

DFU 1

Like method 1, method 2 assumes the same leak categories for DFU 1, i.e. only gas leaks. No other assumptions are made for DFU 1.

6.3.2. Assumptions for AK

Barrier functions as listed in table 14 are used to estimate AK's weights by method 2.

Table 14 Method 2, performance indicators for facility AK

Major accident precursor	Barrier functions
DFU 1: Non-ignited hydrocarbon leaks	Availability of gas detection
	Availability of BDV
	Availability of PSV
	Availability of riser ESDV
	Availability of wing and master valve

6.3.2.1. Method 2

For AK weights are only estimated by method 2. The following assumption is made for the weight estimation.

DFU 1

As mentioned in chapter 5.3.1.1, symmetry is assumed where calibration factors not have been estimated both for positive and negative deviation ratios from CA. This applies to negative deviation for riser ESDV and positive deviations for PSV and wing and master valve.

This assumption leads to limitation in the estimated correction factors and thereby the estimated weights. Assuming symmetry for calibration factors are further addressed in chapter 6.2.4.

6.4. Quality and Limitations of Available Data

6.4.1. "Trends in Risk Level" Data

This chapter will address some data quality issues and limitations from the available data.

6.4.1.1. Performance of Barrier Functions

Quality of Reported Data

Barrier functions' performances have been reported since 2002. The first year was a pilot study, with some limitations on the data quality (PSA, 2004). The methods developed in this master thesis depend heavily on barrier functions' performances. Poor quality of the reported data would severely influence the estimated weights. However as the master thesis limitations states that only data from 2003 and after shall be used, quality issues with the data from the 2002 pilot study are not considered a critical problem.

Some issues with barrier function testing are mentioned in the 2011 RNNP-report (PSA, 2012a). It states that there are indications that some operators lubricate valves before testing and failed tests might not be reported as failed if they are succeeded by successful test. If this is true, it would severely deteriorate the quality of the reported data.

If actual barrier function performance is worse than reported, the weights estimated by the new model will be lower than they should be, indicating a safer facility. No measures are included in the new model to account for erroneous reported performances. This could be viewed as a weakness, but as there is no way of being certain about which operators or facilities wrong performances might be reported for, not compensating for possible erroneous reports are considered the best solution.

Another issue mentioned is the lack of reported data for the BOP barrier functions. The reason for this is that the drilling contractors usually have the barrier functions test records. The reported performances must therefore be viewed as partial and limited.

Limitation of Reported Data

Certain limitations on the use of barrier functions in the weight estimation process are caused by the nature of the reported data.

Performance is only reported for barrier functions as listed in appendix b. As a consequence of this, not all barrier function in place can be accounted for when estimating weights.

According to “Krav til rapportering av barrierer – 2008” (PSA, 2010a) only the most critical barrier functions have been selected for the RNNP-project. If barrier functions’ performances were reported for all barrier functions, a more correct risk picture could have been estimated. Including all barrier functions’ performances would significantly increase the time-consumption. Both when estimating facility specific weights and for the operators reporting the observed performances. Based on that the most common and critical barrier functions’ performances are reported and the extra time it would take to include all barrier functions, it has been concluded that only applying the reported observed performances will give sufficiently good weight estimates.

The reported data does not differentiate between components of tested barrier functions, e.g. smoke or heat detectors, or if there are different manufacturers of the equipment, i.e. if there are more than one supplier of gas detectors. Furthermore, the reported data does not include any information on the age of the detectors or their location, e.g. process area or living quarter.

Tests are performed and reported mainly based on barrier elements’ performances and not for barrier systems. As a consequence of this, redundancy is not covered by the reported data. What type of test being performed is not specified either, e.g. testing performance on demand or failsafe functions.

Weights estimated based on the data available today, gives a risk picture for the facility as a unit. No efforts are made to predict risk level caused by an incident based on where on the facility it occurs. A more detailed description of the tests and failures could have been to include location of incident, type and age of barrier elements and redundancy. However, if impressions of too precise risk level estimates are made, it could increase the possibility of misinterpretations of the result, yielding too high confidence in the results.

As the reporting of data is fairly well incorporated as it is, and the possibility of obtaining risk estimates yielding to high confidence in the results. Conclusion is made that performance as reported today gives sufficient basis for weight estimation.

Equation 17 is used to estimate the PFD. An assumption made in this equation is to assume a one year test interval. As the reported perfomance do not contain information on how often tests are performed, this assumption is assumed to be the best option in chapter 6.1.3. As a constant failure rate is assumed when choosing failure distribution, actual test intervals other than one year will give a wrong PFD estimate. E.g. if the actual test interval is twice a year, a reduction in PFD by a factor of two will be obtained (PSA, 2004).

Not reporting test intervals is viewed as the main weakness of the reported performances. The other weaknesses identified in this chapter, except for redundancy, only limits what potentially can be estimated. But if wrong test interval is assumed, it will automatically give wrong PFD. But as stated in chapter 6.1.3, test intervals of one year will be used to estimate PFD, as this is the only sensible assumption that can be made.

6.4.1.2. Reported Major Accident Precursor Occurrences

Occurrences of major accident precursors on the NCS are reported on an annual basis. For the model development and testing of the developed methods, reported occurrences since 1996 have been available. There are certain quality issues with the reported data for the first years of this period (Oljedirektoratet, 2001), but as only incidents occurred the last ten years will be used in this master thesis, this should not be an issue.

The only application of the reported incidents is to make a comparison between the estimated and the generic weights. As this only is a comparison and not affecting the weight estimation process, poor quality of the reported major accident precursor will not be a critical issue. Not being a critical issue, no further discussion of the data quality will be included.

6.4.2. Quantitative Risk Analysis Data

Risk analysis is not an exact science and two independent risk analyses of the same object will not necessarily give the same results. The outcome of a risk analysis depends on assumptions made and available data. Unfortunately the precise risk estimate often presented after a quantitative analysis will give the impression of high accuracy (Holmgren and Thedéen, 2010) and a risk model more similar to the real world than it actually is (Knudsen, 2010). This is both a benefit and a pitfall of the quantitative risk analysis approach, and it must be taken into consideration when weights are estimated.

The weight estimation process uses QRAs to estimate facility specific weights. The high dependency on QRAs can yield erroneous weights if assumptions or available data the QRA is based on not are sufficiently good and up to date. If the work done preparing and making the QRA not is sufficiently good, the estimated weights will be of limited value. This is further addressed in chapter 6.6.

Article16 in the management regulation demands establishment of criteria for when the risk analysis should be updated. The criteria should be based on conditions that can change the risk of the facility (PSA, 2011a). To achieve the most correct weights, estimation should always be based on the latest revision of QRAs, as this is supposed to always reflect the current and updated risk level of the facility.

According to Knudsen (2010), there are however no guaranties that the QRAs are up-to-date and represents the present design. Knudsen further states that some of the QRAs are manipulated or adjust to achieve the desired outcome, i.e. risk level, to prove that the facility is safe enough. If any of these statements are true for QRAs used to estimate facility specific weights, the weights will be more or less wrong, and of limited value for estimating the risk level on the NCS.

6.4.2.1. Test Cases

All the available QRAs are some years old and new and/or updated QRAs have probably been made. Independently of the QRAs' age, the data quality is assumed to be good enough and sufficient for testing of the new model. If however the new model is to be implemented into the RNNP-project, the estimated weights should be controlled to assure they represent the present state of the facilities.

CA

CA's QRA-report states that event trees are used for estimating risk for hazard scenarios corresponding to DFU 1 and DFU 9. Based on this and the good possibility of getting access to parts of the risk analysis for CA, it was decided that CA would be used to test weight estimation by both methods and to demonstrate the estimation procedure for calibration factors.

When access to parts of the risk analysis was gained, it however became apparent that no event trees were available for DFU 9. In an email received 15th of April 2013, Principal Engineer Terje Dammen, Safetec, confirms that there are no event trees made for DFU 9 in

CA's QRA. Based on the lack of event trees, no weight estimates could be made for DFU 9 by method 1.

For DFU 1 on the other hand, the available data was sufficient for both weight estimation methods to be used.

No redundancy is accounted for in CA's QRA calculations. The assumption gives a conservative result, so it can be argued for being a reasonable assumption for a QRA that is meant to show that there is an acceptable risk level. However when weights are estimated, it can lead to larger consequences following an incident, giving higher expected numbers of fatalities and thereby higher weights.

As the aim of weight estimation method 1 not is to establish a new QRA, but simply reproduce the calculation with varying barrier functions' performances, the weight estimation process will use the same assumptions as the QRA, and thereby accepting the possibility of higher risk estimates.

No weights are estimated for DFU 8. The performance indicators assigned to DFU 8 are no doubt relevant for the risk level of CA, but neither the QRA report nor the available part of the risk analysis contains the necessary information for DFU 8 weights to be estimated.

AK

AK's QRA does not contain detailed enough data for weights to be estimated by method 1, there is however sufficient data available for method 2 to be used. There are no reasons to suspect the quality of the QRA to be questionable so all extracted data will be treated like it is of good quality.

If weights were to be estimated for DFU 3, the shallow gas categories would not be relevant for AK. This decision is based on information given by Professor J. E. Vinnem during a meeting 8th of March 2013, where he stated that most of the time before a fixed facility is installed; test drilling is performed to avoid shallow gas under the facility.

During the same meeting, it was also agreed that if DFU 3 weights were to be estimated, they should only include one category and not three like the generic weights. This was based on that the QRA does not contain enough details to distinguish between degrees of severity of a blowout.

BH

Originally three facilities were supposed to be included as test cases. They were partly chosen based on that their QRA-reports were available. After receiving the QRAs and a meeting with

Professor J. E. Vinnem, 28th of February 2013, it was decided that only two facilities was to be used, and thereby omitting BH as a test case.

BH was omitted as a test case because the available part of the QRA-report is too coarse for weight estimation based solely on it to be possible. The report only contains FAR-values for the facility's different areas from the defined hazards, but if any weights are to be estimated assumed frequencies and manning distribution used in the risk analysis is also needed.

DFU 1 weights are estimated based on the same procedure as used for the generic weights. If this was to be done for BH, risk from process leaks would have to be expressed for different leak sizes. This is not available from the QRA-report. Neither is leak sizes used as representative leaks given (Det Norske Veritas, 2005).

Due to lack of information from the QRA-report, no further effort has been put into assessing the quality of the available data for BH.

6.5. Dual Assurance

The RNNP-project uses two general types of indicators; major accident precursors and barrier functions' performances. By the distinction between types of indicators as defined in chapter 2.2.2, the RNNP-project's two types of indicators would be classified as follows:

- Major accident precursor – Lagging indicator
- Barrier functions' performances – Leading indicators

As described in chapter 2.2.3, dual assurance is HSE's concept for combining leading and lagging indicators. The aim of this concept is to have continuous monitoring of the indicators' goodness. Even though the RNNP-project does not mention dual assurance as an aim, at least by the name dual assurance, the model used today can be viewed as a version of dual assurance.

The model developed in this master thesis will combine the leading and lagging indicators into one total risk indicator. This indicator will not allow comparison between lagging and leading indicators in the same way as it can be done today. A combination of the new total risk indicator and a separate presentation of major accident precursors and barrier functions' performances can then be an alternative approach, if dual assurance is desired.

6.6. Uncertainty

A model like the one developed in this master thesis will most certainly be exposed to significant uncertainties. To express this, the word "estimate" is used throughout the entire master thesis.

As discussed in chapter 6.4, quality of the available data will have great impact on the results. This and other issues influencing uncertainty of the estimated weights will be addressed in this chapter.

Correlation

Correlation between barrier functions' performances is not included in the new model. This can lead to both reduced and increased uncertainties of the estimated weights.

Higher uncertainties would be a consequence of ignoring interactions between barrier functions, maintenance and safety culture. But if interactions were to be included, but not all interactions are identified or some are left out on purpose, this could increase the uncertainties as well.

Successful implementation of correlation would mean that all interaction are identified and accounted for. This would no doubt yield lower uncertainties, but it is found highly unlikely that all interactions can be identified. And including them correctly would be an enormous job as it would have to be done separately for all facilities.

Based on the above it has been concluded that implementation of correlation in the new model most probably would give higher uncertainties and it is therefore left out.

Quantitative Risk Analyses

As discussed in chapter 6.4, QRAs will have limitations and does not necessarily represent the real world in a sufficient manner. As the new model bases itself on risk estimated by the risk analysis and thereby the same assumptions, uncertainties from the QRAs will be include in the estimated weights.

The famous saying "shit in, shit out" goes for weights estimated by the new model as well. If the QRAs have high uncertainties, the estimated weights will have high uncertainties. But since the new model is based on already existing QRAs, there is little that can be done to reduce these uncertainties.

Performance of Barrier Functions

The new model uses observed barrier functions' performances to estimate weights. This means that assumed performances is replaced with the facility's observed performances.

A question is then if there actually are changes in performance or if they are random variations. In appendix d a control of significance is included for CA's barrier functions' performances. The method used to control if changes are significant, is to assume a sufficient amount of tests for normal distribution to be used and estimating a 90 % prediction interval

based on the previous year's failure rate. If observed performance is within this interval, the change is not significant.
As it can be seen from the table in appendix d, only five out of 30 observed barrier functions' performances change significantly. When estimating new weights, observed performances are used independently of whether the changes are significant. It can be argued that as most of the changes not are significant, an average of all years should have been used instead of a three year rolling average. A three year rolling average is however used for all failure rates.
Quality of reported performance data is discussed in chapter 6.4. Some challenges and limitations were identified. These can give larger uncertainties for the estimated weights, but it is assumed that the quality of the reported barrier functions' performances is good and that it not will give a significant contribution to the models uncertainties.

Estimation by Method 2

As discussed in chapter 6.2.4, method 2 will be less detailed than method 1. As a consequence of this higher uncertainties must be expected from method 2 than method 1. This is a result of the method's nature and no actions to minimize uncertainties caused by choice of method have been identified.

6.7. Weaknesses and Limitations of the New Model

The new model has several weaknesses and limitations. This chapter will give a discussion of the ones identified.

Time-consuming

If the new model is to be implemented, it will demand significantly more work than the model used today. Initial work would include estimating specific weights for all facilities. But it is not just the initial work that will be time-consuming, new annual weights would have to be estimated every year, as the new model's weights change on an annual basis.
When QRAs are updated, new initial weights would have to be estimated. Implementation of the new model would therefore have to include establishment of a reporting or monitoring system that identifies when the initial weights no longer are valid and needs to be updated.
An alternative approach can be to estimate weights only when they are needed. This would reduce the necessary initial work, but more work would have to be performed every year, as not only new annual weights would have to be estimated.

Weight Estimation Method

As mentioned several times earlier in this master thesis, method 1 requires a fairly detailed QRA. This is the main weakness of method 1. But when all necessary information is available, method 1 will give more accurate weights than method 2.

Method 2 will be dependent on the goodness of the general calibration factors. This means that there is a higher possibility of method 2 estimating erroneous weights than method 1. The weakness is as described in chapter 6.2.4 caused by the lack of possibility to update the causal chain with observed performances.

Another weakness of method 2 is the need for benchmarks. There are not benchmarks values for all barrier functions. Where this is the case, barrier functions' performances cannot influence the estimated weights, as no deviation ratios can be established.

When using linear regression to estimate weight equations for DFU 1, there are some discontinuities between the leak ranges. The test cases show that there can be as much as 10% difference between the two leak size ranges for leak size of 5,5 kg/s. It has however been concluded that this issue only can be solved by using more valid digits to express expected fatalities. This would give the impression of having higher accuracy than what actually will be the case and the discontinuities will be accepted as they are.

Phases of Operation

Weights estimated by the new model will be deepened on the scenarios analyzed in the QRA. If the facility's QRA do not contain scenarios for all phases of the facilities life, the weights estimated will only be valid for major accident precursors occurring during the analyzed phases. E.g. if QRA only evaluates risk during normal operation and a major accident precursors occurs during maintenance shutdown, the estimated weights will give an inaccurate risk estimate.

The same goes for POB, if POB changes significantly and no new weights are estimated, the risk level caused by a major accident precursor will be incorrect.

This weakness does not give poorer estimates of the risk than the use of generic weights would. Weights estimated by the new model can easily be adjusted to include changes to actual operational status. Whilst on the other hand, generic weights not are suited for adjustment to be made based on the operational phase at all.

Available Data

An important limitation follows from what barrier functions performance is reported for. As only some performances are reported, the weight estimation process can only use updated performances for some of the barrier functions in place. Thereby ignoring parts of the barrier system's status when estimating the risk level.

Estimated barrier function failure rate is assumed to be correct and no considerations are given to whether changes in failure rates are significant. If the estimated failure rates turn out to be wrong, the estimated weights will also be wrong.

Which method that is used to estimate weights are based on how much information that is available from the QRA. But if no QRA is available or it is impossible to estimate weights from the available QRA, neither of the two methods can be used. No alternative approach for estimating the influence from barrier functions' performance is developed, and if neither method 1 nor method 2 can be used, a risk indicator including both barrier functions' performances and major accident precursors will not be estimated.

Not Estimating New Weights for all Types of Major Accident Precursors

New weights are only estimated for the major accident precursors with barrier functions assigned, as given in appendix b. This means that generic weights still are used for several DFUs.

As shown by the test cases, facility specific weights can deviate significantly from the generic weights and vary from year to year. This is also supported by table 15 in "Metoderapporten" (PSA, 2011b, page 42) where it can be seen that expected fatalities given an incident varies with several powers of ten.

Not estimating new weights for all DFUs, and thereby continuing to use generic weights is a significant weakness. But as the aim of this master thesis is to combine observed performance of barrier functions with major accident precursors, this limitation is made.

7. Discussion and Further Work

7.1. Discussion of Model

In “Lessons from Longford – The Esso Gas Plant Explosion”, Andrew Hopkins (2000) points at how Lost Time Injury (LTI) are used to express risk. He argues that this indicator only is suited for expressing the safety from occupational accidents and not as a major accidents risk indicator. I.e. LTI does not indicate if a facility will be safe from major accidents in the future. The Longford accident happened in 1998, twelve years later the Macondo blowout occurs. The Deep Water Horizon rig had a perfect LTI record and was renowned for being a safe rig (United States National Commission on the BP Deepwater Horizon Oil Spill and Offshore Drilling, 2011).

This was not the first time an accident report claims that BP draws incorrect assumptions based on LTI reports. Both the “Baker report” (Baker, 2007) and the investigation of the BP Grangemouth incident (Health and Safety Executive, 2003) concluded that BP used LTI to draw wrong conclusions on risk from major accidents. This shows that the petroleum industry is a slow learner and if focus is to be placed on the potentially largest losses, i.e. major accidents, it needs a push.

The RNNP-project focuses on major accidents, occupational hazards and acute spills. This master thesis however has only focused on major accidents. Major accidents are one of the main focus areas of PSA. Another focus area is barriers. Combining major accident precursors with barrier functions’ performances into one total risk indicator, will give a risk level indication based on both these areas of priority.

This chapter discusses the developed methods, results and the new models potential.

Can we know if Weights are Correct

The developed model will estimate the risk level contribution from a major accident precursor through the use of two new weight estimation methods. This introduces the question of whether the new weights are more “correct” and if the generic weights are “wrong”.

Whether weights are correct or wrong depends on what is meant by correct. As the major accident precursors used to estimate the risk level mostly do not lead to major accidents, it can be argued that the “correct” weight should be 0,0 fatalities and all other “estimates” are “wrong” unless a major accident actually occurs.

Even though the total risk indicator is a lagging indicator, it is not intended to express actual consequences, but the potential consequences the incident could have had as identified in the

QRA. I.e. even though a facility operates its entire life span with several major accident precursor and no fatalities, the “correct” weight as defined in this section should never be 0,0. The “correct” weight will be rooted in the risk estimated in the QRA form that type of major accident precursor.

A risk analysis is never “correct”. There are no guaranties that the same result will be achieved for the same object if the analysis is performed by different analysts. The reason for this is that assumptions and limitations have to be made. A model will therefore always be a model, describing the real world with a varying degree of correctness, but never perfectly.

“Solum certum nihil esse certi” (Montaigne, 1580) or “It is only certain that there is nothing certain”. A risk analysis expresses the statistical expected fatalities. In a way it can therefore be said that all weights are “wrong”. They do no express actual people killed, but the expected fatalities for all the terminal events used in the risk analysis that the major accident precursor could have led to if the course of events had been different.

In this master thesis, “correct” and “wrong” weights refer to which degree the estimated weights reflect the risk estimated by the risk analysis. I.e. the generic weights will only be “correct” for CA if weights estimated based on CA’s risk analysis are the same.

This does not mean weights estimated by the RNNP-project are “wrong”, leaving the projects’ risk level indication worthless. It only means that the generic weights not are considered to give good risk estimates compared to risk estimated based on the specific facility’s risk analyses. This is a consequence of the generic weights only estimating an average of what the risk potentially could be.

In this discussion chapter correct means that the weights reflects risk estimated by the specific facility’s risk analysis, either based on original assumptions or updated with observed barrier functions’ performances. Correspondingly, wrong means that weights not are considered to reflect the risk estimated from the risk analysis sufficiently well.

Quantitative Risk Analysis

The preceding section defined correct and wrong according to whether the risk level estimate reflects risk estimated by the QRA. This leads to the question of whether a risk analysis can be “correct”.

Chapter 6.4.2 addresses the issue of QRAs being manipulated and outdated. Manipulation and risk analyses reflecting what the government want to see is further addressed by Aven (2011). This article uses the phrase “mercenary scientists” to express how analysts can adjust the risk analyses to demonstrate that the risk is acceptable so that a company can avoid expensive measures to reduce the risk. The article further emphasizes that the scientific basis and uncertainties in the risk analysis process must be addressed in order to improve the results' quality.

Aven's article addresses risk assessment in general and not QRAs for the NCS in particular. Also Knudsen (2010) claims that risk analyses performed for facilities on the NCS are manipulated to achieve an acceptable risk level estimate. Based on this, it is no reason to believe issues addressed in Aven's article not are relevant for the new model.

The result of a risk analysis will to a large degree depend on the analyst (Aven and Heide, 2009). It is difficult, if not impossible, to say that one analysis is more correct than the other. The risk estimation model developed in this master thesis does not only depend on the results from risk analyses, it is based on them. If the QRAs are manipulated or the analyses are outdated, the risk level estimated by the new model will give inaccurate indications on the risk level.

This is an issue when saying that the new weights are more "correct" than the generic weights. If the risk analyses not describe the "true" risk level sufficiently well, the estimated weights will not necessarily yield a better estimate of a major accident precursor's risk level contribution.

In this master thesis it has been assumed that the available risk analyses describe the "true" world sufficiently well, so that the risk level can be estimated based on them. This is also the basis for this discussion chapter.

The New Model

There are several weaknesses with today's RNNP model. The perhaps most significant weakness is shown in figure 2. The yellow arrow demonstrates the generic weights' inability to account for the specific facility's barriers. Not assessing observed barrier functions' performances in the weighting process, amplifies this weakness further.

The model developed in this master thesis attempts to improve these weaknesses. This is achieved by both assessing the barrier functions actually in place and their observed performances. Figure 3 demonstrate how the new weights will be different from the generic weights and vary on an annual basis.

It is important to remember that a model like this includes both lagging and leading indicators. But the new models' intention is not to predict the future risk level. The leading indicators are only used to find more "correct" estimates of the potential consequences following from the lagging indicators. I.e. the observed barrier functions' performances influence on the potential consequences from occurrences of major accident precursors. The total risk indicator must therefore be viewed as a lagging indicator, expressing the NCS offshore oil and gas industry's ability to manage risk.

Methodology

The new model estimates facility specific weights by one of two methods; method 1 or method 2. Their estimation procedures are described in chapter 4.3.2 and chapter 4.3.3.

As discussed in chapter 6.2.1 both methods have their benefits. Method 1 is the method considered to give the most "correct" weight estimates as it consists of updating and reproducing parts of the risk analysis. On the negative side it demands more detailed information from the risk analysis than method 2, and it will be the most time-consuming approach as the risk analysis partly has to be reproduced.

Method 2 is the alternative approach whenever the available information from the QRA not is sufficiently detailed for method 1 to be used. It benefits from being easy to use and it estimates weights quickly when the general calibration factors are established. A weakness of method 2 is that it is less specific to the facility because the causal chain not is directly updated with the barrier functions' performances, but instead includes them through the application of correction factors.

No weight estimation method for facilities without available QRAs has been developed. It can be argued that a third method should have been developed as barrier functions' performances are available.

An alternative approach could have been to adjust weights by the same procedure as method 2, but using generic weights instead of weights estimated from the QRA. This alternative is however not considered to give good weight estimates. This is based on table 15 in "Metoderapporten" (PSA, 2011b, page 42), where it can be seen that expected fatalities given an non-ignited HC leak, varies with several powers of ten for the same type of facility.

Estimating new weights based on the generic weights and observed barrier functions' performances would then give the impression of having a much higher accuracy than what probably will be the case, undermining the new models' credibility as a more "correct" weight estimation model.

Method 1's weight estimation procedure is more or less based on standard risk analyses methods, and as mentioned in chapter 6.2.3 it serves little purpose to discuss method 1's methodology. Method 2's methodology on the other hand are mostly developed in this master thesis and certain elements needs to be addressed.

Benchmark values are only based on Statoil's internal guidelines. If other operators base their safety philosophy on different barrier function performances, deviation from the benchmark values will give "wrong" estimations of the risk level. This is however considered to be a limited problem, as Professor J. E. Vinnem in a meeting 8th of Mars 2013 confirms that most of the operators on the NCS mainly use the same demands for barrier functions' performances.

Weights estimated by method 2 are adjusted by correction factors. This means that the causal chain not is updated with observed barrier functions' performances like done for method 1. Instead correction factors adjust the terminal events' risk sum. This makes the estimated weights less facility specific and includes a source of uncertainty.

Neither of the preceding issues is considered to be the critical issues with method 2. The correctness of weights estimated by method 2 mainly depends on the calibration factors.

Calibration factors are the relationship between deviation ratios and correction factors. As shown for test case 1 – CA, it is fully possible to find facility calibration factors giving weights estimated by method 2 equal to weights estimated by method 1.

Facility calibration factors are estimated for only one of the test cases. These will include properties specific only to CA. The estimation procedure of facility calibration factors is like described in chapter 4.3.3.3 based on trial and error. As discussed in chapter 6.2.4.5 there are no guaranties that the estimated calibration factors will be the correct ones. This is a severe weakness and is part of the reason why general calibration factors are established.

General calibration factors is a measure used to avoid individual facilities' specific properties disproportionately influencing the estimated weights and to minimize the issue of wrong facility calibration factors being estimated because of the trial an error process.

Calibration is not based on other facilities' risk, but the effect on other facilities' risk level from their barrier functions' performances. Method 2 estimates risk from the specific facility's risk analysis and then adjust it by correction factors. This will give more facility specific weight estimates than the generic weights, allowing a more "correct" picture of the risk to be estimated. Less facility specific weights will however always be a consequence of using method 2 instead of method 1.

Presentation

The new model is rooted in a desire for a more "correct" risk level estimate, but it will have another benefit as well. When all aspects of the risk are included in the same indicator, it will be easier to express the general trend of all risk influencing factors. One single total risk indicator gives a better overview of the current risk level and makes it easier to present the risk picture. This can make it easier for the involved parties to understand the message PSA tries to send through the RNNP-project, and thereby be a facilitation tool for risk level improvement.

The all including total risk indicator has its weakness as well. As barrier functions' performances only are used to estimate weights for the major accident precursors and not presented as separate indicators, it will be harder to see which barrier functions most effort should be put into approving.

The performances of riser ESDV and DHSV were discussed in the presentation of the 2012 RNNP-results on 25th of April 2013. The RNNP-model used today shows that the performances of these barrier functions not are as good as the industry standard indicates that they should be. If only the new model is used, this would not be possible to see. Poor barrier function performances would only yield higher expected fatalities from an incident, and not include indication of which barrier functions that causes this increase.

The best alternative would therefore be a combination of the model used today and the model developed in this master thesis. The combination should include a presentation of the risk level, numbers of major accident precursors and observed performances of barrier functions.

Another possible application of the new model is to assign one event to each of the updated weights every year, like done in the comparison, and see how the potential risk changes on an annual basis. The risk trend could be presented for each single facility, each operator and for the NCS as a whole. This eliminates the need for a major accident precursor occurrence in order for the risk level estimate to change, and yields a leading risk indicator.

Uncertainties

Several factors induce uncertainties into the estimated weights. The weight estimation is based on QRAs and observed performance of barrier functions. Any uncertainties in these data sources would automatically give uncertainties in the estimated weights. Due to the importance of these data sources for the weight estimation and these being the best data sources available, any uncertainties induced from them must be accepted.

Other sources that might induce uncertainties into the estimated weights are correlation and the previously discussed general calibration factors. Uncertainties following from the latter will be a consequence of the methodology used to estimate weights. The only identified measure to avoid these uncertainties is weight estimation by method 1.

Uncertainties following from correlation are difficult to avoid. Both if it is included or if it is left out, correlation will have the potential to cause uncertainties in the estimated weights. Correlation is discussed thoroughly in chapter 6.1.4, and it is concluded that correlation not will be a part of the weight estimation process.

Assigning Barrier Functions to Types of Major Accident Precursors

The new model use the same DFUs as today's model. A revision of the DFUs could have been an option. There have not been any occurrences of DFU 2 incidents and DFU 9 can be viewed as a successor of DFU 10. An alternative could therefore be to include DFU 2 and its barrier functions into DFU 1 and DFU 10 into DFU 9. But as the RNNP-project has been running for some years, it has been decided that there are no reasons to rock the boat and make changes to the fundamentals of today's model.

Allocation of barrier functions to the DFUs are mainly done based on "Metoderapporten" (PSA, 2011b). A thorough discussion of this is given in chapter 6.2.2 and will not be reproduced here.

All the barrier functions included in the model can be classified as technical barrier functions. Maintenance and mustering time are also reported, but as neither of them are assigned to any of DFU 1 to DFU 10 by "Metoderapporten", they are not included in the new model. This means that the only reported nontechnical barrier functions are left out of the risk level estimation.

An improvement of the new model could be to include performance of nontechnical barrier functions. As this would increase the complexity of the model and nontechnical barrier functions largely depend on the facility at question (Øien and Sklet, 2001), the more "correct" risk estimate achieved from implementing them most probably would not justify the effort and time the implementation would take.

Testing of the Developed Model

The new model is tested for DFU 1 on two test cases; the two facilities CA and AK. Originally the model was supposed to be tested on DFU 9 as well, but as the available data did not contain the necessary information to allow weight estimation by method 1, the model could not be tested on DFU 9. As weight estimation not is the main goal of the test cases, only testing the developed methods on DFU 1 is considered sufficient as it demonstrates that the new model works.

Both methods are tested on and facility calibration factors are estimated for test case 1 – CA. Figure 15 shows weights estimated by both methods compared with the generic weights. The new weights are more or less the same from year to year. This is as expected since change in failure rate compared to the previous year only is significant for 5 out 30 annual failure rates. For leaks larger than approximately 1kg/s on CA, the generic weight equations indicate a higher contribution to the NCS's risk level following a DFU 1 incident. For the largest leaks the generic weights indicate a risk approximately six times larger than the new model.

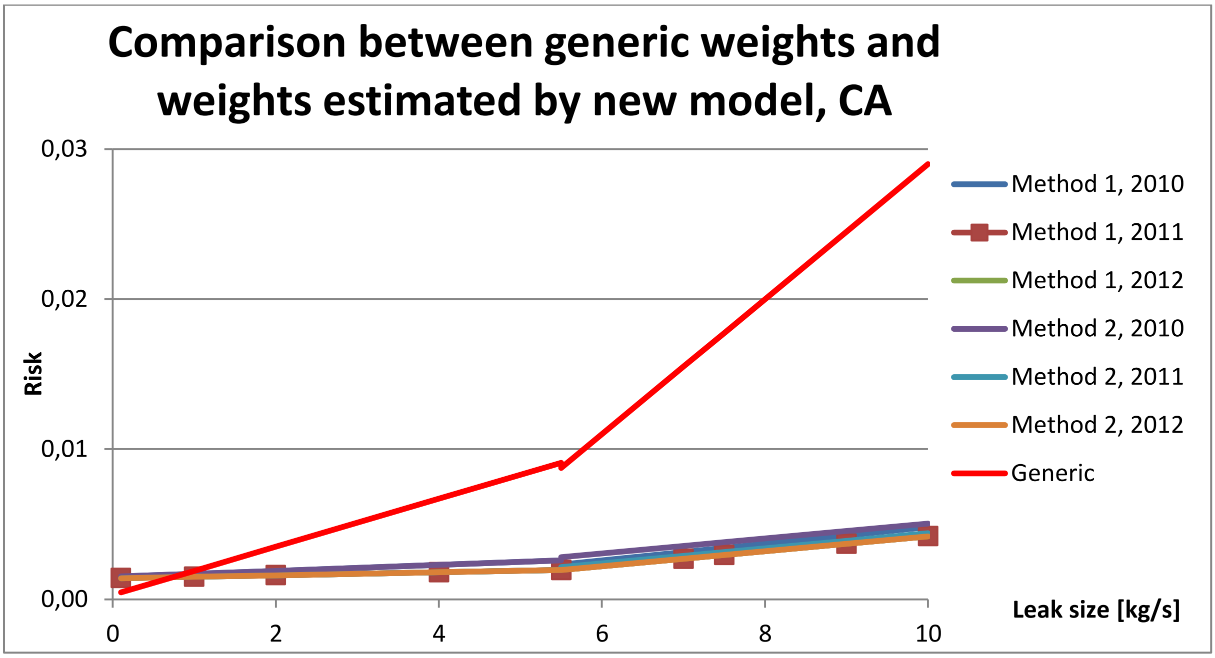


Figure 15 DFU 1, risk (PLL), comparison between estimated and generic weights, CA

Weights estimated for AK by method 2 shows considerably annual variations. The estimated weight equations are both larger and smaller than the generic weights depending on the year in question.

For AK weights are only estimated by method 2. It is therefore not possible to control if the estimated weights are "correct" or not. If weights could be estimated by method 1 as well, a comparison between weights estimated based on calibration factors from CA and method 1 could have been made. From this it could have been concluded whether facility calibration factors could be used to represent other facilities, or if they only are applicable to the facility they are estimated for.

The test cases show the potential of the new model. The risk level based on barrier functions is changing on an annual basis. The use of generic weights in today's model does not allow this to be taken into consideration. How much more facility specific the weights potentially can be is demonstrated by test case 1 – CA were the estimated weights are significantly lower than the generic weights.

As only one set of facility calibration factors are estimated from the test cases, it is not really possible to decide whether method 2 works. In order for this to be determined, general calibration factors must be established and tested on facilities which can use method 1 to see if the estimated weights are similar for both methods.

If it turns out that method 2 gives sufficiently "correct" results, the best option for generalization and implementation of the new model might actually be to use method 2 for all facilities. This will decrease the work load and still give time and barrier function deepened weight estimates.

Model Relative to Master Thesis Objective

The objective of this master thesis was to develop a new risk level estimation model, test to see if it works and determine if the necessary data is available.

As discussed previously, a model consisting of two methods have been developed. Test case 1 – CA demonstrated that method 1 can be used to estimate weights. The estimation process however demands a lot of the available data. For test case 1 – CA it turned out that only having the QRA-report was insufficient. Parts of the risk analysis were also needed if method 1 was to be used.

Method 2 has been tested on both test cases, but only with CA's facility calibration factors. Since method 1 not could be used on test case 2 – AK, it is not possible to decide whether the used facility calibration factors will give "correct" weights. Further testing of method 2 and establishment of general calibration factors are therefore necessary in order to decide whether method 2 can estimate "correct" enough weights.

For method 2 all the necessary data have been extracted from the QRA-reports. As long as this is the case, it will be relatively easy to get hold of the necessary data. For method 1 on the other hand, this can be more difficult, as it turns out that the necessary data might include parts of the risk analysis.

Whether it is possible to get hold of all risk analyses is questionable, but the operators should at least have the QRA-reports available. So when answering the question of necessary data being available, the answer will depend on how available is defined. The data exist, but it might be some issues with gaining access to it.

Both methods demonstrate the potential of including barrier functions' performances when estimating weights. The estimates can change from year to year, and weights can be significantly different from the generic weights.

Added Value of New Model

The big question when developing a model like the one in this master thesis is whether it gives added value compared to the old model. And if added value is identified, it must be decided if the added value justifies the added complexity and effort required by the new model.

Added value from the new model is clearly demonstrated by the test cases. From test case 1 – CA it can be seen that the risk level the generic weights would indicate following a DFU 1 incident is too high. And test case 2 – AK demonstrates how the annual variation of barrier functions' performances changes the estimated weights, and thereby changes the major accident precursors' contribution to the risk level.

Generic weights are based on expected fatalities given an event. Earlier in this master thesis it has been commented that these estimates varies by several powers of ten for the same type of facility. By the definition of correct and wrong used in this chapter, risk level estimated by the generic weights will then be wrong. It is simply an average of what it could be. Implementation of the new model would remove this source of error from the estimated total risk indicator.

But there is no such thing as free lunch. Implementation of the new model requires a lot of initial work to be done. It would include everything from collecting QRAs and parts of risk analyses to estimation of barrier functions' PFDs and reproduction of analysis to finally estimating new time and barrier function dependent weights. To assure that weights always are based on the latest risk analysis revision, a system monitoring when new risk analysis are made should also be created.

Compared with today's model, it would no longer be enough to multiply the number of incidents with the generic weights and then summing up for all facilities. The same work would still have to be performed, but it would also include the estimation of annual weights based on the last year's observed barrier functions' performances.

Whether the new model should be implemented would have to be based on a cost benefit analysis. This is not within the scope of this master thesis, and will not be addressed any further.

7.2. Further Work

During the master thesis work, areas needing further work have been identified. One of these areas is necessary if the model shall be implemented while others are ideas to how the model can be expanded even further.

- The next step of the model development is to establish general calibration factors and use these to test method 2 on facilities which can use both weight estimation methods. This will answer whether method 2 can give “correct” weight estimates.
- Correlation has been left out of the new model. Finding a good method for identifying and including correlation between barrier functions’ performances could give more “correct” weight estimates and reduce the uncertainties.
- No benchmark values are identified for some of the barrier functions’ performances. Either including trend analyses or establishing benchmark values for these barrier functions would mean that they could be included in the weight estimation by method 2.
- The new model only considers major accidents leading to personnel risk. Establishing a method that use barrier functions’ performances to predict risk to the environment from acute spills and risk to personnel from occupational accidents could be an interesting further development of the new model.
- Develop a method for deciding whether the QRAs describes the real world well enough for weights to be estimated based on them.

8. Conclusion

The aim of this master thesis was to develop a model combing major accident precursors and observed barrier functions' performances into one total risk indicator, test the new model and decide it the necessary data is available.

A model consisting of two methods has been developed; method 1 and method 2.

Two test cases are used to test the new model. Test case 1 demonstrates that method 1 estimates weights that reflect the risk analysis, but it proves to be time-consuming and requiring detailed QRAs and/or parts of the risk analysis.

Method 2 has been tested on both test cases. Calibration factors are however only estimated for one of the facilities. As a consequence of only being able to use method 1 for one of the two test cases, it is not possible to conclude on the correctness of weights estimated by method 2. None the less, the test cases demonstrate that annually varying facility specific weights including barrier functions' performances can be estimated by method 2.

Quality of the available data is concluded to be sufficiently good for weight estimation. Whether the necessary data is available or not depend on how available is defined. Risk analyses are performed for all facilities on the NCS, so they exist. As long as the operators are willing to share their QRA-reports, new weights can be estimated. This does not necessarily mean that weights can be estimated by method 1 for all facilities, as this demands fairly detailed QRAs.

The best way of implementing the new model will perhaps be to establish general calibration factors by use of method 1, and then estimate weights for all facilities included in the RNNP-project by method 2. This gives a total risk indicator including both barrier functions' performances and major accident precursors and reduces the work load compared to using method 1.

9. References

AVEN, T. 2011. Selective critique of risk assessments with recommendations for improving methodology and practise. *Reliability Engineering and System Safety,* 96**,** 509-514.

AVEN, T. & HEIDE, B. 2009. Reliability and validity of risk analysis. *Reliability Engineering and System Safety,* 94**,** 1862-1868.

AVEN, T. & VINNEM, J. E. 2007. *Risk Management with Applications from the Offshore Industry,* London, Springer London.

AYYUB, B. M. & MCCUEN, R. H. 2003. *Probability, statistics, and reliability for engineers and scientists,* Boca Raton, Fla., Chapman & Hall/CRC.

BAKER, J. A., BAKER, G. E., PRIEST, S., TEBO, P. V., I., ROSENTHAL, I., BOWMAN, F. L., HENDERSHOT, D., LEVESON, L., WILSON, D., GORTON, S., & WIEGMANN, D. 2007. The report og the BP U.S. refineries indepednent safety review panel.

BUCHANAN, A. H. 2001. *Structural design for fire safety,* Chichester, Wiley.

DET NORSKE VERITAS 2005. "BH" PLATTFORM – TOTALRISIKOANALYSE (TRA). REVISJON NR. 02 ed.

DET NORSKE VERITAS 2007. "AK", Risikoanalyse, As-built 2007. Rev. 04 ed.

HEALTH AND SAFETY EXECUTIVE 2003. Major Incident Investigation Report, BP Grangemouth Scotland (29th May – 10th June 2000).

HEALTH AND SAFETY EXECUTIVE 2006. *Developing process safety indicators,* London, HSE books.

HOLLNAGEL, E. 2004. *Barriers and accident prevention,* Aldershot, Ashgate.

HOLMGREN, Å. J. & THEDÉEN, T. 2010. Risk Analysis. *In:* GRIMVALL, G., HOLMGREN, Å. J., JACOBSSON, P. & THEDÉEN, T. (eds.) *Risks in Technological Systems.* Springer London.

HOPKINS, A. 2000. *Lessons from Longford: the Esso gas plant explosion,* Sydney, CCH Australia Ltd.

KHAN, F., ABUNADA, H., JOHN, D. & BENMOSBAH, T. 2010. Development of risk-based process safety indicators. *Process Safety Progress,* 29**,** 133-143.

KNUDSEN, R. H. 2010. *Myter og sannheter om risikoanalytisk tilnærming.* Master thesis, Norwegian University of Science and Technology.

MONTAIGNE, M. D. 1580. *The Essays of Montaigne*.

NORSOK 2004. Well integrity in drilling and well operations, NORSOK Standard D-010, Rev. 3. Lysaker, Norway: Standards Norway.

NORSOK 2008. Technical safety, NORSOK Standard S-001, Edition 4. Lysaker, Norway: Standards Norway.

NORSOK 2010. Risk and emergency preparedness assessment. NORSOK Standard Z-013, Edition 3. Lysaker, Norway: Standards Norway.

NORWEGIAN MARITIME AUTHORITY 1991. Regulation concerning stability, watertight subdivision and watertight/weathertight closing means on mobile offshore units. *No. 878.* Norwegian Maritime Authority.

OLIVE, C., O'CONNOR, T. M. & MANNAN, M. S. 2006. Relationship of safety culture and process safety. *Journal of Hazardous Materials,* 130**,** 133-140.

OLJEDIREKTORATET 2001. Utvikling i risikonivå - norsk sokkel. Pilotprosjektrapport 2000. Stavanger: OD.

OLJEDIREKTORATET 2002. Utvikling i risikonivå - norsk sokkel - Fase 2 rapport 2001. Oljedirektoratet.

OLJEDIREKTORATET 2003. Utvikling i risikonivå - norsk sokkel - Fase 3 hovedrapport 2002. Oljedirektoratet.

PSA 2004. Utvikling i risikonivå - norsk sokkel - Fase 4 hovedrapport 2003. Petroleum Safety Authority Norway.

PSA 2005. Utvikling i risikonivå - norsk sokkel - Fase 5 hovedrapport 2004. Petroleum Safety Authority Norway.

PSA 2006. Utvikling i risikonivå - norsk sokkel - Fase 6 hovedrapport 2005. Petroleum Safety Authority Norway.

PSA 2007. Utvikling i risikonivå - norsk sokkel - Fase 7 hovedrapport 2006. Petroleum Safety Authority Norway.

PSA 2008. Risikonivå i petroleumsvirksomheten - Norsk sokkel 2007. Petroleum Safety Authority Norway.

PSA 2009. Risikonivå i petroleumsvirksomheten - Hovedrapport, utviklingstrekk 2008, norsk sokkel. Petroleum Safety Authority Norway.

PSA 2010a. Krav til selskapenes rapportering av ytelse av barrierer RNNP. Petroleum Safety Authority Norway.

PSA 2010b. Risikonivå i petroleumsvirksomheten - Metoderapport - Akutte utslipp. Petroleum Safety Authority Norway.

PSA 2010c. Risikonivå i petroleumsvirksomheten -Hovedrapport, utviklingstrekk 2009, norsk sokkel Petroleum Safety Authority Norway.

PSA 2011a. Regulations Relating to Managemetn and the Duty to Provide Information in the Petroleum Activities and at Certain Onshore Facilities (The Management Regulations). Petroleum Safety Authority Norway.

PSA 2011b. Utvikling i risikonivå - norsk sokkel, Metoderapport for 2010,. Petroleum Safety Authority Norway.

PSA 2012a. Hovedrapport - Utviklingstrekk 2011 - Norsk Sokkel. Petroleum Safety Authority Norway.

PSA 2012b. Trends in Risk Level in the Petroleum Activity - Summary Report 2011. Petroleum Safety Authority Norway.

PSA 2013a. Prinsipper for barrierestyring i petroleumsvirksomheten. Petroleum Safety Authority Norway.

PSA 2013b. Reported DFUs (Definerte Fare- og Ulykkessituasjoner). Petroleum Safety Authority Norway.

PSA 2013c. Reported performance of barrier functions. Petroleum Safety Authority Norway.

RAUSAND, M. 2011. *Risk assessment,* Hoboken, N.J., Wiley.

RAUSAND, M. & HØYLAND, A. 2004. *System reliability theory : models, statistical methods, and applications,* Hoboken, N.J., Wiley-Interscience.

RAUSAND, M. & UTNE, I. B. 2009. *Risikoanalyse - teori og metoder,* Trondheim, Tapir akademisk forl.

SAFETEC 2009. "CA" – Quantitative Risk Analysis. Rev. 05 ed.

SAFETEC 2013. 04543 "CA" FPSO risk results.

SKLET, S. 2006. Safety barriers: Definition, classification, and performance. *Journal of Loss Prevention in the Process Industries,* 19**,** 494-506.

UNITED STATES NATIONAL COMMISSION ON THE BP DEEPWATER HORIZON OIL SPILL AND OFFSHORE DRILLING 2011. *Deep water: the Gulf oil disaster and the future of offshore drilling: report to the President,* [Washington, D.C.], National Commission on the BP Deepwater Horizon Oil Spill and Offshore Drilling.

VINNEM, J. E., HESTAD, J. A., KVALØY, J. T. & SKOGDALEN, J. E. 2010. Analysis of root causes of major hazard precursors (hydrocarbon leaks) in the Norwegian offshore petroleum industry. *Reliability Engineering & System Safety,* 95**,** 1142-1153.

WALPOLE, R. E., MYERS, R. H., MYERS, S. L. & YE, K. 2012. *Probability & statistics for engineers & scientists,* Boston, Mass., Pearson.

WIKIPEDIA. 2013. *Correlation and dependence* [Online]. Available: http://en.wikipedia.org/wiki/Correlation_and_dependence [Accessed 24/02/2013.

ØIEN, K. 2001a. A framework for the establishement of organizatioanl risk indicators. *Reliability Engineering and Syatem Safety*, 147-167.

ØIEN, K. 2001b. Risk indicators as a tool for risk control. *Reliability Engineering and System Safety,* 74, 129-145.

ØIEN, K. & SKLET, S. 2001. Metodikk for utarbeidelse av organisatoriske risikoindikatorer. Trondheim: SINTEF Teknologiledelse, Sikkerhet og pålitelighet.

10.Appendices

Appendix A Generic Weights Used in “Trends in Risk Level” Project

All weights presented in this appendix is reproduced from “Utvikling I risikonivå – norsk sokkel – Metoderapport 2010” (PSA, 2011b). Some of the DFUs will be assessed and assigned a weight individually if they occur. These indicators are not assigned a weight but instead says “Individually assessed” were the assigned weight would have been given.

DFU 1:

Table 15 Generic weights for DFU 1 (PSA, 2011b, 2012a)

Facility type:	Leak rate (x) 0,1 to 5,5 [kg/s]	Leak rate (x) 5,5 – 10 [kg/s]	Leak rate (x) > 10 [kg/s]
Fixed facilities	0,0008*x + 0,0002	0,023*x - 0,008	Individually assessed
Normally unmanned installations	0,05*(0,0008*x + 0,0002)	0,05*(0,023*x - 0,008)	Individually assessed
Floating production unit	0,0008*x + 0,0002	0,023*x - 0,008	Individually assessed
Complexes	0,0008*x + 0,0002	0,023*x - 0,008	Individually assessed
Mobile installations	0,0008*x + 0,0002	0,023*x - 0,008	Individually assessed

DFU 2:

No weights have created for DFU 2, as no ignited hydrocarbon leaks have occurred during the reporting period. Weights will be estimated if needed.

DFU 3:

Table 16 Generic weights for DFU 3 (PSA, 2011b)

Facility type:	Severity and type of event				
	Regular	**Severe**	**High risk**	**Shallow gas**	**Severe shallow gas**
Fixed facilities	0,0035	0,017	0,087	0,087	0,87
Normally unmanned installations	0,0035	0,017	0,087	0,087	0,87
Floating production unit	0,0035	0,017	0,087	0,087	0,87
Complexes	0,0017	0,0087	0,043	0,043	0,43
Mobile installations	0,0005	0,0027	0,013	0,013	0,13

DFU 4

Table 17 Generic weights for DFU 4 (PSA, 2011b)

Facility type:	Weight
Fixed facilities	0,021
Normally unmanned installations	0,0021
Floating production unit	0,021
Complexes	0,021
Mobile installations	0,021

DFU 5

Table 18 Generic weights for DFU 5 (PSA, 2011b)

Facility type:	**Weight**
Fixed facilities	0,0081
Normally unmanned installations	0,0013
Floating production unit	0,0009
Complexes	0,0073
Mobile installations	0,0023

DFU 6

Table 19 Generic weights for DFU 6 (PSA, 2011b)

Facility type:	**Weight**
Fixed facilities	0,0009
Normally unmanned installations	0,0002
Floating production unit	0,0001
Complexes	0,0002
Mobile installations	0,0003

DFU 7

Table 20 Generic weights for DFU 7 (PSA, 2011b)

Facility type:	Weight
Fixed facilities	0,00211
Normally unmanned installations	0,00129
Floating production unit	0,00134
Complexes	4,0*10^-6
Mobile installations	0,0042

DFU 8

Table 21 Generic weights for DFU 8 (PSA, 2011b)

Facility type:	Major event	Super major event
Fixed facilities	0,006	Individually assessed
Normally unmanned installations	0,003	Individually assessed
Floating production unit	0,011	Individually assessed
Complexes	0,001	Individually assessed
Mobile installations	0,004	0,080

DFU 9 and DFU 10

Table 22 Generic weights for DFU 9 and DFU 10 (PSA, 2011b)

Facility type:	**Per leak**	**Major damage**
Fixed facilities	0,48	0,096
Normally unmanned installations	0,00017	8,6*10^-6
Floating production unit	0,097	0,0097
Complexes	0,091	0,006
Mobile installations	-	-

The given weights for DFU 9 and DFU 10 are to be used for events that occur within 200 meters of the facility. From 200 meters to 500 meters they should be reduced with 75%. Events occurring further away from the facility than 500 meters will be assigned a weight of zero. No weights are assigned for mobile installations.

Appendix B What Barrier Functions are assigned to which DFUs

Table 23 Allocation of barrier functions to DFUs (PSA, 2011b, page 22, table 9)

DFU nr	**Description of DFU**	**Barrier indicators**
1	Non-ignited hydrocarbon leak	• Gas detection, availability test • Blow down valve, availability test • Process safety valve, availability test • Riser emergency shutdown valve, closure and leak test • Wing and master valve, closure and leak test
2	Ignited hydrocarbon leak	• Fire detection, availability • Deluge valves, availability test • Fire pumps, availability test
3	Well events/loss of well control	• Downhole safety valve, availability test • Surface BOP, availability test ▪ Drilling BOP ▪ Coiled tube BOP ▪ Pressure pipes BOP ▪ Wireline operation BOP • Subsea BOP, availability test ▪ Drilling BOP ▪ Coiled tube BOP ▪ Pressure pipes BOP ▪ Wireline operation BOP
8	Structural damage	• Watertight doors closing, availability test • Function test of ballast system valves, availability test • GM values
9	Subsea leak	• Gas detection, availability test • Fire detection, availability • Deluge valves, availability test • Fire pumps, availability test • Riser emergency shutdown valve, closure and leak test

Table 23, except for DFU 9, describes which DFUs barrier functions performance are reported for according to "RNNP – Metoderapport 2010" (PSA, 2011b). DFU 9 and corresponding barrier function performance is added according to consultation with Professor J. E. Vinnem.

Appendix C Benchmark Values for Barrier Functions' Performances

Table 24 Benchmark values for barrier functions' performances (PSA, 2012a, page 123, table 24)

DFU nr	Description of DFU	Barrier indicators		Benchmark value
1	Non-ignited hydrocarbon leak	Gas detection, availability test		0,01
		Blow down valve, availability test		0,005
		Process safety valve, availability test		0,04
		Riser emergency shutdown valve		0,01
		Riser emergency shutdown valve	Closure test	0,01
			Leak test	0,01
		Wing and master valve		0,02
		Wing and master valve	Closure test	0,02
			Leak test	0,02
2	Ignited hydrocarbon leak	Fire detection, availability		0,01
		Deluge valves, availability test		0,01
		Fire pumps, availability test		0,005
3	Well events/loss of well control	Downhole safety valve, availability test		0,02
		Test BOP		*
		Surface BOP, availability test	Drilling BOP	*
			Coiled tube BOP	*
			Pressure pipes BOP	*
			Wireline operation BOP	*
		Subsea BOP, availability test	Drilling BOP	*
			Coiled tube BOP	*
			Pressure pipes BOP	*
			Wireline operation BOP	*
8	Structural damage	Watertight doors closing, availability test		**
		Function test of ballast system valves, availability test		***
		GM values		****
9	Subsea leak	Gas detection, availability test		0,01
		Fire detection, availability		0,01
		Deluge valves, availability test		0,01
		Fire pumps, availability test		0,005
		Riser emergency shutdown valve		0,01
		Riser emergency shutdown valve	Closure test	0,01
			Leak test	0,01

* Availability requirements not considered suitable, use trend analysis instead.

** No benchmark values assigned, assessed by demands from stability regulations

*** No benchmark values assigned, depend on performance demands of barrier elements from QRA/risk philosophy

**** No benchmarks

In the RNNP-report for year 2011 (PSA, 2012a), the benchmark values are referred to both as industry standard and as Statoil's internal guidelines. Professor J. E. Vinnem confirms in a conversation 8th of Mars 2013 that the benchmarks actually are Statoil's internal guidelines, but that they are more or less the same for the entire NCS.

Appendix D CA’s Barrier Functions’ Performances

Table 25 CA's barrier functions' performances

Barrier element	Year	Three year rolling average failure rate	PFD	Failed tests or half failure	Significant difference
Tests fire detection	2010	3,17E-02	1,57E-02	Failed test	-
	2011	1,29E-02	6,41E-03	Failed test	No
	2012	3,12E-03	1,56E-03	Failed test	Yes
Tests gas detection	2010	3,81E-02	1,88E-02	Failed test	-
	2011	1,81E-02	8,99E-03	Failed test	Yes
	2012	3,04E-03	1,52E-03	Failed test	Yes
Tests riser ESDV	2010	3,92E-02	1,94E-02	Failed test	-
	2011	3,96E-02	1,95E-02	Failed test	No
	2012	4,17E-02	2,05E-02	Failed test	No
Tests closing, riser ESDV	2010	6,13E-02	3,01E-02	Failed test	-
	2011	6,29E-02	3,08E-02	Failed test	No
	2012	6,21E-02	3,04E-02	Failed test	No
Tests leak, riser ESDV	2010	5,43E-03	2,71E-03	Half failure	-
	2011	4,85E-03	2,42E-03	Half failure	No
	2012	7,04E-03	3,51E-03	Half failure	No
Test Christmas tree	2010	3,42E-03	1,71E-03	Failed test	-
	2011	1,62E-03	8,11E-04	Half failure	No
	2012	1,54E-03	7,71E-04	Half failure	No
Tests closing, Christmas tree	2010	6,85E-03	3,42E-03	Failed test	-
	2011	3,25E-03	1,62E-03	Half failure	No
	2012	3,09E-03	1,54E-03	Half failure	No
Tests leak, Christmas tree	2010	3,42E-03	1,71E-03	Half failure	-
	2011	3,25E-03	1,62E-03	Half failure	No
	2012	3,09E-03	1,54E-03	Half failure	No
Tests DHSV	2010	7,81E-03	3,90E-03	Half failure	-
	2011	4,07E-03	2,03E-03	Half failure	No
	2012	1,23E-02	6,15E-03	Failed test	No
Tests BDV	2010	5,62E-03	2,80E-03	Half failure	-

	2011	1,88E-03	9,38E-04	Failed test	No
	2012	3,43E-03	1,71E-03	Failed test	No
Tests PSV	2010	1,19E-03	5,96E-04	Half failure	-
	2011	7,28E-04	3,64E-04	Half failure	No
	2012	1,16E-03	5,78E-04	Half failure	No
Tests deluge valves	2010	2,87E-03	1,44E-03	Half failure	-
	2011	2,65E-03	1,32E-03	Half failure	No
	2012	5,88E-03	2,94E-03	Failed test	No
Tests firepump startup	2010	1,09E-03	5,44E-04	Half failure	-
	2011	8,01E-04	4,01E-04	Half failure	No
	2012	8,01E-04	4,63E-04	Half failure	No
Tests closing, water proof doors	2010	2,98E-03	1,49E-03	Half failure	-
	2011	2,96E-03	1,48E-03	Half failure	No
	2012	1,16E-02	5,79E-03	Half failure	Yes

Appendix E AK's Barrier Functions' Performances

Table 26 AK's barrier functions' performances

Barrier element	Year	Three year rolling average failure rate	PFD	Failed tests or half failure	Significant?
Tests fire detection	2005	1,43E-02	7,10E-03	Failed test	-
	2006	8,73E-03	4,35E-03	Failed test	No
	2007	2,49E-03	1,24E-03	Failed test	Yes
	2008	2,88E-03	1,44E-03	Failed test	No
	2009	3,19E-03	1,59E-03	Failed test	No
	2010	3,88E-03	1,94E-03	Failed test	No
	2011	3,18E-03	1,59E-03	Failed test	No
	2012	1,60E-03	8,01E-04	Failed test	No
Tests gas detection	2005	2,07E-02	1,03E-02	Failed test	-
	2006	1,70E-02	8,46E-03	Failed test	No
	2007	1,26E-02	6,26E-03	Failed test	No
	2008	2,52E-03	1,26E-03	Failed test	Yes
	2009	1,58E-03	7,92E-04	Failed test	No
	2010	1,67E-03	8,36E-04	Failed test	No
	2011	8,24E-04	4,12E-04	Failed test	No
	2012	9,98E-04	4,99E-04	Failed test	No
Tests riser ESDV	2005	1,00E-01	4,84E-02	Half failure	-
	2006	8,33E-02	4,05E-02	Failed test	No
	2007	6,25E-02	3,06E-02	Failed test	No
	2008	7,14E-02	3,49E-02	Failed test	No
	2009	7,14E-02	3,49E-02	Failed test	No
	2010	7,14E-02	3,49E-02	Failed test	No
	2011	6,25E-02	3,06E-02	Failed test	No
	2012	6,25E-02	3,06E-02	Failed test	No
Tests closing, riser ESDV	2005	-	-	Failed test	-
	2006	-	-	Failed test	-
	2007	1,67E-01	7,89E-02	Half failure	-
	2008	1,00E-01	4,84E-02	Half failure	No

	2009	8,33E-02	4,05E-02	Failed test	No
	2010	8,33E-02	4,05E-02	Failed test	No
	2011	8,33E-02	4,05E-02	Failed test	No
	2012	7,14E-02	3,49E-02	Failed test	No
Tests leak, riser ESDV	2005	-	-	Failed test	-
	2006	-	-	Failed test	-
	2007	-	-	Failed test	-
	2008	-	-	Failed test	-
	2009	5,00E-01	2,13E-01	Half failure	-
	2010	5,00E-01	2,13E-01	Failed test	No
	2011	2,50E-01	1,15E-01	Failed test	No
	2012	5,00E-01	2,13E-01	Failed test	No
Tests Christmas tree	2005	1,61E-02	8,02E-03	Half failure	-
	2006	1,25E-02	6,22E-03	Half failure	No
	2007	1,18E-02	5,86E-03	Half failure	No
	2008	1,29E-03	6,44E-04	Half failure	No
	2009	1,14E-03	5,68E-04	Half failure	No
	2010	7,03E-03	3,51E-03	Half failure	Yes
	2011	1,05E-02	5,23E-03	Half failure	No
	2012	8,24E-03	4,11E-03	Half failure	No
Tests closing, Christmas tree	2005	-	-	Failed test	-
	2006	-	-	Failed test	-
	2007	5,56E-03	2,77E-03	Failed test	-
	2008	3,31E-03	1,65E-03	Failed test	No
	2009	2,55E-03	1,27E-03	Failed test	No
	2010	2,12E-03	1,06E-03	Failed test	No
	2011	1,58E-03	7,91E-04	Failed test	No
	2012	1,18E-03	5,88E-04	Failed test	No
Tests leak, Christmas tree	2005	-	-	Half failure	-
	2006	-	-	Failed test	-
	2007	1,22E-02	6,07E-03	Failed test	-
	2008	3,05E-03	1,52E-03	Failed test	No
	2009	2,05E-03	1,02E-03	Half failure	No

	2010	1,20E-02	5,98E-03	Half failure	Yes
	2011	1,99E-02	9,91E-03	Half failure	No
	2012	1,65E-02	8,19E-03	Half failure	No
Tests DHSV	2005	1,06E-02	5,30E-03	Half failure	-
	2006	1,05E-02	5,24E-03	Half failure	No
	2007	4,17E-03	2,08E-03	Half failure	No
	2008	4,20E-03	2,10E-03	Half failure	No
	2009	4,03E-03	2,01E-03	Failed test	No
	2010	4,50E-03	2,25E-03	Failed test	No
	2011	1,08E-02	5,36E-03	Failed test	No
	2012	1,06E-02	5,30E-03	Failed test	No
Tests BDV	2005	5,50E-02	2,70E-02	Failed test	-
	2006	5,08E-02	2,50E-02	Failed test	No
	2007	4,80E-02	2,36E-02	Failed test	No
	2008	1,12E-02	5,60E-03	Failed test	Yes
	2009	7,04E-03	3,51E-03	Failed test	No
	2010	9,64E-02	4,67E-02	Half failure	Yes
	2011	1,89E-01	8,88E-02	Failed test	Yes
	2012	1,76E-01	8,33E-02	Failed test	No
Tester PSV	2005	6,00E-02	2,94E-02	Failed test	-
	2006	4,79E-02	2,36E-02	Half failure	No
	2007	6,17E-02	3,02E-02	Half failure	No
	2008	5,56E-02	2,73E-02	Half failure	No
	2009	4,76E-02	2,34E-02	Half failure	No
	2010	2,30E-02	1,14E-02	Half failure	Yes
	2011	1,04E-02	5,17E-03	Half failure	No
	2012	4,93E-03	2,46E-03	Half failure	No
Tests BOP	2005	1,00E-01	4,84E-02	Half failure	-
	2006	-	-	Failed test	-
	2007	4,90E-03	2,45E-03	Failed test	-
	2008	1,91E-03	9,54E-04	Failed test	No
	2009	1,16E-03	5,78E-04	Failed test	No
	2010	1,03E-03	5,15E-04	Failed test	No

	2011	1,09E-03	5,46E-04	Half failure	No
	2012	1,45E-03	7,24E-04	Failed test	No
Surface BOP: Tests drilling BOP	2005	-	-	Failed test	-
	2006	-	-	Failed test	-
	2007	-	-	Failed test	-
	2008	-	-	Failed test	-
	2009	-	-	Failed test	-
	2010	3,33E-03	1,66E-03	Half failure	-
	2011	1,77E-03	8,83E-04	Half failure	No
	2012	1,20E-03	6,01E-04	Failed test	No
Surface BOP: Tests wireline	2005	-		Failed test	-
	2006	-	-	Failed test	-
	2007	-	-	Failed test	-
	2008	-	-	Failed test	-
	2009	-	-	Failed test	-
	2010	1,00E-01	4,84E-02	Failed test	
	2011	1,00E-01	4,84E-02	Failed test	No
	2012	1,00E-01	4,84E-02	Half failure	No
Tests deluge	2005	6,98E-02	3,41E-02	Failed test	-
	2006	1,47E-02	7,32E-03	Failed test	No
	2007	1,92E-02	9,55E-03	Failed test	No
	2008	2,38E-02	1,18E-02	Half failure	No
	2009	1,85E-02	9,20E-03	Half failure	No
	2010	8,77E-03	4,37E-03	Half failure	No
	2011	4,67E-03	2,33E-03	Half failure	No
	2012	4,67E-03	2,33E-03	Half failure	No
Tests starting of fire pumps	2005	5,49E-04	2,74E-04	Half failure	-
	2006	1,39E-03	6,92E-04	Half failure	No
	2007	1,75E-03	8,77E-04	Half failure	No
	2008	4,63E-03	2,31E-03	Failed test	No
	2009	2,21E-03	1,11E-03	Failed test	No
	2010	2,02E-03	1,01E-03	Failed test	No
	2011	1,97E-03	9,86E-04	Failed test	No

	2012	1,61E-03	8,06E-04	Failed test	No

Appendix F Dataset for Comparison of Failure Rates

The following dataset is used to create figure 14. The numbers are generated randomly, and are only meant to show that three year rolling average is the most suited description of the failure rate for exponential distribution.

Table 27 Dataset for comparison of failure rates

Year	Total tests	Failed tests	Annual failure rate	Failure rate method			
				1	2	3	4
1	999	44	0,044	0,044	0,052	0,043	
2	950	40	0,042	0,042	0,052	0,043	
3	950	42	0,044	0,044	0,052	0,043	0,043
4	1030	41	0,040	0,040	0,052	0,043	0,042
5	946	38	0,040	0,040	0,052	0,043	0,041
6	250	42	0,168	0,168	0,052	0,043	0,083
7	802	0	0,000	0,000	0,052	0,043	0,069
8	998	32	0,032	0,032	0,052	0,043	0,067
9	683	38	0,056	0,056	0,052	0,043	0,029
10	711	39	0,055	0,055	0,052	0,043	0,048

Appendix G Data Used to Estimate DFU 1 Weights for CA

Data presented in this appendix are taken from CA's QRA (SAFETEC, 2009) report and part of CA's risk analysis (SAFETEC, 2013).

PLL from part of risk analysis				
Leak category	**QRA**	**2010**	**2011**	**2012**
Small Leak	6,20E-04	6,32E-04	6,17E-04	6,07E-04
Medium Leak	2,72E-04	2,80E-04	2,63E-04	2,51E-04
Large Leak	5,50E-04	5,54E-04	5,40E-04	5,29E-04
Total	1,44E-03	1,47E-03	1,42E-03	1,39E-03

Leak categories [kg/s]					
Leak category	**Gas/2-phase**		**Oil**		**Average/used in calculation [kg/s]**
	Range [kg/s]	**Typical [kg/s]**	**Range [kg/s]**	**Typical [kg/s]**	
Small Leak	0,1-1	0,50	0,1- 1	0,50	0,5
Medium Leak	1,0-10	5,00	1,0-20	15,00	5
Large Leak	>10	20,00	>20	50,00	20,00

Leak category	Annual leak frequency from QRA
Small Leak	4,14E-01
Medium	1,28E-01
Large Leak	5,49E-02
Total	0,5974565

Leak category	Expected people killed per leak			
	QRA	**2010**	**2011**	**2012**
Small Leak	1,50E-03	1,50E-03	1,52E-03	1,49E-03
Medium Leak	2,12E-03	2,12E-03	2,18E-03	2,05E-03
Large Leak	1,00E-02	1,00E-02	1,01E-02	9,82E-03

Appendix H Data Used to Estimate DFU 1 Weights for AK

Data presented in this appendix are taken from AK's QRA (Det Norske Veritas, 2007).

PLL from QRA	
Leak category	**QRA**
Small Leak	7,90E-5
Medium Leak	3,08E-4
Large Leak	2,61E-3
Total	3,00E-03

Leak Category	**Range [kg/s]**	**Representative leak rate [kg/s]**
0,1 to 1,0 [kg/s]	0,1	0,5
1,0 to 10 [kg/s]	1,0 -10	5
More than 10 [kg/s]	>10,0	50

Leak category	**Annual leak frequency from QRA**
Small Leak	0,362
Medium	0,100
Large Leak	0,101
Total	0,563

Year	**Correction factor**			**Adjusted expected fatalities from representative leak**		
	0,5 [kg/s]	5 [kg/s]	50 [kg/s]	0,5 [kg/s]	5 [kg/s]	50 [kg/s]
QRA	-	-	-	0,000218	0,00308	0,0258
2005	1,44	1,85	0,382	0,000533	0,00878	0,0357
2006	1,20	1,56	0,298	0,000480	0,00789	0,0335
2007	1,05	1,37	0,259	0,000447	0,00730	0,0325
2008	0,422	0,386	0,0593	0,000310	0,00427	0,0274
2009	0,329	0,257	0,0303	0,000290	0,00387	0,0266
2010	1,72	2,37	0,384	0,000593	0,0104	0,0358
2011	3,12	4,52	0,737	0,000900	0,0170	0,0449

2012	2,90	4,19	0,669	0,000850	0,0160	0,0431