Universitetet i Stavanger

Faculty of Science and Technology

MASTER'S THESIS

Study program/Specialization: Industrial economics/Technological innovation and entrepreneurship	Spring semester, 2021 ~~Open~~ / Restricted access
Writer: Marius A.S. Aasen	Marius Aasen ………………………………………… (Writer's signature)
Faculty supervisor: Håkon B. Abrahamsen External supervisor(s): Sigbjørn Daasvatn	
Thesis title: System Reliability of a Wax Control System: Application of Markov Chains in an Initial Reliability Analysis of New Subsea Technology	
Credits (ECTS): 30	
Key words: Subsea technology Reliability Probability Python Markov chain	Pages: 94 + enclosure: 117 Stavanger, 15th June/2021 Date/year

Abstract

One challenge related to subsea oil operations is wax deposition on the oil pipeline. The problem arises when hot crude oil comes in contact with a cold pipeline. The paraffin wax, present in hot crude oil, crystallizes and builds up from the pipeline wall and inwards. Subsea 7 is now developing a new mechanical solution to deal with wax deposition; the Wax Control System (WCS). In a WCS, a cleaning pig flows with the oil inside the pipeline. While the oil is cooled down to ambient temperature before further exportation, the pig removes any wax present on the pipeline wall.

This thesis uses the WCS as a case to study the usefulness of the Markov chain as a tool in the reliability analysis of new subsea equipment. Two objectives are made to answer this. The first objective is to analyse two different possible configurations of the WCS in light of costs, availability, and maintainability using Markov chains. The second objective is to discuss if and how Markov chains fit as an analysis tool for components under conditions not generally discussed in reliability literature. More specifically, a case where one component in a system is non-stationary, and a case where a component's availability might depend on another component's lifetime. For these components, existing tools might fall short. The research question is then: *How do Markov chains fit as a tool in the initial reliability analysis of complex components in a new subsea system?*

In short, the results suggests that the use of Markov chains in reliability analysis can add a degree of overview of a subsystem, mostly because of the transition graphs. Further, compared to an existing method used to calculate availability, the use of Markov chain theory can open up to more detailed and flexible analyses.

Even though the use of the Markov chain can open up for flexibility and more detailed analyses, there are some drawbacks. When using Markov chains, it is assumed that, for instance, failure, repair, and wear are exponentially distributed. This assumption might, however, not hold in practice. Many components' failure rate is not necessarily constant. Also, finding these rates can be quite challenging. Existing tools within risk and reliability also have some drawbacks when analysing a non-stationary component in a system.

Preface

For the last two years, I have studied for a Master's degree in industrial economics at the University of Stavanger, and this thesis marks the end of my time as a student. With a background in mechanical engineering, I wanted to write a master's thesis related to lifetimes of mechanical components, system reliability and how the final system is affected by various cost factors.

I want to thank my supervisor Håkon B. Abrahamsen for his tips, advices and for helping me in the right direction. I would also like to thank my lecturer in probability and statistics, Tore Selland Kleppe, for taking the time to answer my emails regarding topics from the curriculum.

Also, a big thanks to my external supervisor at Subsea 7, Sigbjørn Daasvatn. He provided me with all the possible information I could need only a short time after requesting it. He has also always been available for meetings. Working with Subsea 7 has given me insight into the industry and shown me how the topics I have learned at the University apply in the industry.

Marius A.S. Aasen
15th June 2021

Contents

List of Figures

List of Tables

Acronyms

AMPL Automated Multiple Pig Launcher.

CDF Cumulative Distribution Function.

ETA Event Tree Analysis.

FDV Flow Directional Valve.

FMEA Failure Mode and Effects Analysis.

FMECA Failure Mode, Effects and Criticality Analysis.

FTA Fault Tree Analysis.

HAZOP Hazard and Operability Study.

MLE Maximum Likelihood Estimation.

MTTF Mean Time to Failure.

MTTR Mean Time to Repair.

OREDA Offshore and Onshore Reliability Data.

P&ID Piping and Instrumentation Diagram.

PDF Probability Density Function.

PMF Probability Mass Function.

PSA Petroleum Safety Authority Norway.

RBD Reliability Block Diagram.

SPU Submerged Production Unit.

SWIFT Structured What-If Technique.

TRL Technology Readiness Level.

WCS Wax Control System.

Chapter 1

Introduction

1.1 Background

When paraffin wax present in hot crude oil comes in contact with a cold oil pipeline, it crystallises and becomes a solid because of the temperature difference. This solid builds up from the pipeline wall and in, reducing or even blocking the oil flow from a well to a processing facility.

One way of dealing with this problem is to heat the pipeline to a temperature at which the wax melts and is mixed with the crude oil. However, heating kilometres of pipeline from wells to processing facilities requires a substantial amount of energy [1]. The oil and gas industry, like many other industries, is in a position where the energy consumption and the CO_2 footprint must be reduced. Doing so is necessary to meet the goals in the Paris agreement.

A solution to this problem is to cool the oil down to ambient before exporting it to a processing facility. To deal with the solid wax on the pipeline wall, Subsea 7 is developing a Wax Control System (WCS).

At the time being, Subsea 7 is in a technology qualification process, following a procedure developed by DNVGL. This procedure is meant as a suggestion on how to qualify new technology. Qualification of new technology is vital, as it ensures the end-user that the technology works under specified conditions and limitations [2].

As a measure of how far technology has come in the qualification process, industries such as the oil and gas industry uses the Technology Readiness Level (TRL). The number of levels can vary, as some TRL scales are industry-specific. However, the one recommended for subsea technology ranges for TRL 0-7, suggesting that the technology is an "unproven concept" and "field-proven", respectively. In June 2020, the WCS reached TRL 4, suggesting that the technology is "environment tested" [1], [2, B.4].

In DNVGL's recommended practice for technology qualification, it is mentioned three tools for the calculation of the system's reliability. These tools are Fault Tree Analysis (FTA), Monte Carlo simulation, and Reliability Block Diagram (RBD). However, to use these tools in calculations, one needs to estimate various parameters. These estimations can, for instance, be based on data obtained by running the system over time. Unfortunately, the data necessary to do so do not exist at the time being.

1.2 Research Question and Thesis Objectives

This thesis aims to aid Subsea 7 in the technology qualification process, in addition to examining the fit of another analysis tool for reliability calculations, in light of the WCS. To do so, two objectives related to reliability of the WCS will be completed:

1. Analyse, compare, and evaluate two different configurations in light of availability, maintainability, and costs using Markov chains.

2. Use the WCS to construct case studies and discuss how and if Markov chains fits to describe components under conditions not commonly discussed in reliability theory. More specifically:

   (a) One case where the availability of one component might depend on the lifetime of another component, and

   (b) A system with a non-stationary component, where commonly used tools and methodologies within risk and reliability in some areas might fall short.

Based on the objectives mentioned above, the following research question can be formulated:

How do Markov chains fit as a tool in the initial reliability analysis of complex components in a new subsea system?

As a final note, it should be stressed that the interpretation of the word "initial" is here that
(1) analyses of the components have not been carried out yet, and
(2) little data is currently available.

1.3 Limitations

As the WCS is in an ongoing qualification process, little data currently exist other than some results from the accelerated testing (done as a part for qualification of TRL 4). Data unique for each component in the system (such as failure rates) is unavailable, as the system has not yet failed during testing. Estimating these rates is the topic of another thesis written in parallel and will therefore not be an objective in this thesis. However, in some situations, some estimations must be done to get something to work with.

Further, concrete costs and prices associated with the system and its components are, for the time being, unknown. The discussion related to costs will therefore be based on assumptions and some coarse estimates from Subsea 7.

In a design or qualification process, multiple standards and recommended practices exist. This thesis will not discuss any standards or recommended practices that might be relevant, other than some relevant topics in DNVGL-RP-A203 (DNVGL's recommended practice for technology qualification).

1.4 Structure

The structure of this thesis is presented below.

Chapter 1: Introduction
This chapter presents background information related to the wax control system and this thesis' main objectives, the research question, and the overall structure.

Chapter 2: Description of the WCS
In this chapter, the system's components, their functionality and the overall functionality of the system is presented and described.

Chapter 3: Theory
Chapter 3 is the first part of the literature review. Here, relevant theory within the fields of probability and reliability are presented. Some of the theory might be considered as methodologies but is thought to fit best under theory. The theory will later be used in the analyses.

Chapter 4: Methodologies
Chapter 4 is the second part of the literature review. This chapter presents tools and

methodologies often used within the field of risk and reliability. Some of these methodologies will be used in a discussion in Chapter 6.

Chapter 5: RAM Analysis: AMPL
In Chapter 5, an analysis is performed on two configurations involving the AMPL. Arguments supporting why the AMPL is chosen over the other components in this analysis are also given. This chapter is related to thesis objective 1 and 2(a).

Chapter 6: Discussion About and Analysis of a Non-Stationary Component
Chapter 6 is separated into two sections, with the common subject being the cleaning pig. The first section discusses how existing tools and methods used within the fields of risk and reliability in some areas might fall short when analysing a system with a non-stationary component. The second section is more analytical. Here, Markov chain theory is used to study the wear of the cleaning pig, assuming that the testing in the qualification testing is continued. This chapter is related to thesis objective 2(b).

Chapter 7: Results
This chapter presents the results from the analysis in Chapter 5 and the second section of Chapter 6.

Chapter 8: Discussion
In Chapter 8, the results presented in Chapter 7 are discussed. The aspect of costs is brought in to view the results from a different perspective. A discussion concerning the assumptions made and the validity of these is also given. Also, the discussion concerning the usefulness of the Markov chain as a tool is addressed in this chapter. This is done by comparing the Markov chain to an existing method, in addition to reviewing the experience using the Markov chain and the results from Chapter 5 and 6.

Chapter 9: Conclusion
Based on the discussion in Chapter 8, a conclusion is made. The research question, goals and objectives presented in Chapter 1 are included to help draw a conclusion.

Chapter 2

Description of the WCS

In the first section, some of the key components used in the WCS are explained. The theory from Section 2.1 is used to explain the whole system in Section 2.2. Unless otherwise noted, each paragraph in Section 2.2 is referring to the process shown in Figure 2.7. The information and figures used in this chapter are obtained from internal presentations from Subsea 7 and from [1], [3].

2.1 Key Components

2.1.1 Cleaning Pig

The cleaning pig is the tool used to remove the wax on the inside wall. It is designed to just fit inside a pipeline and flows with the oil. Initially, it is housed inside an Automated Multiple Pig Launcher (AMPL) before being launched in hot crude oil inside a bundled pipe. The cleaning pig is shown in Figure 2.1.

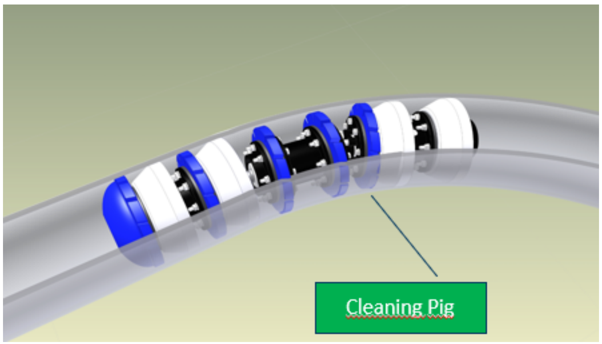


Figure 2.1: A pig flowing in a pipeline. Obtained from [4].

Figure 2.2 shows the dimensions of the pig and its composition of parts (such as seal rings) necessary to clean the pipeline walls. As one can see, at the front, the pig is fitted with bypass holes. This allows some oil to flow through the pig internally. This will reduce the speed of the pig relative to the oil flow, so that wax particles will be pushed forward by the pig. Figure 2.3 shows wax leftovers collected by the pig during testing for qualification to TRL 4.

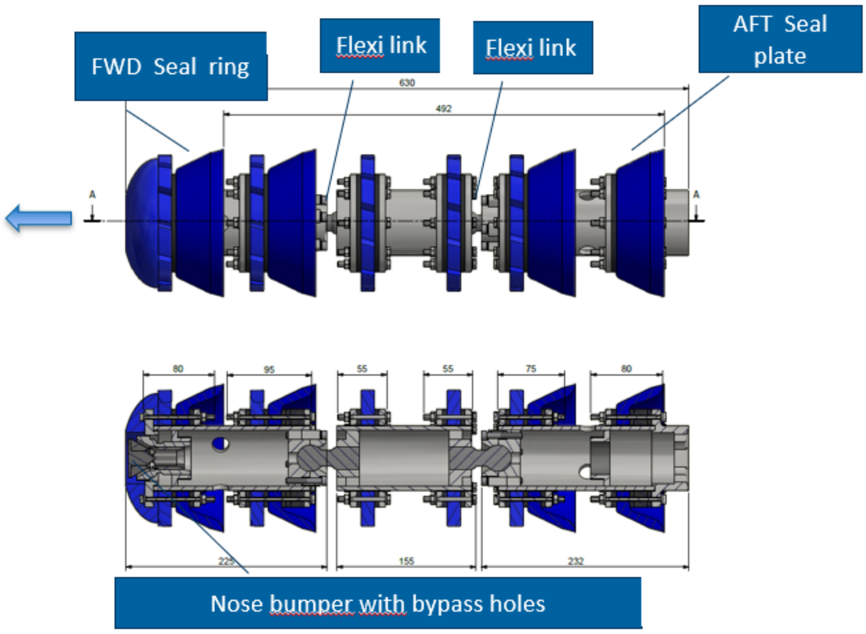


Figure 2.2: The pig's dimensions and composition of parts. Obtained from [4].

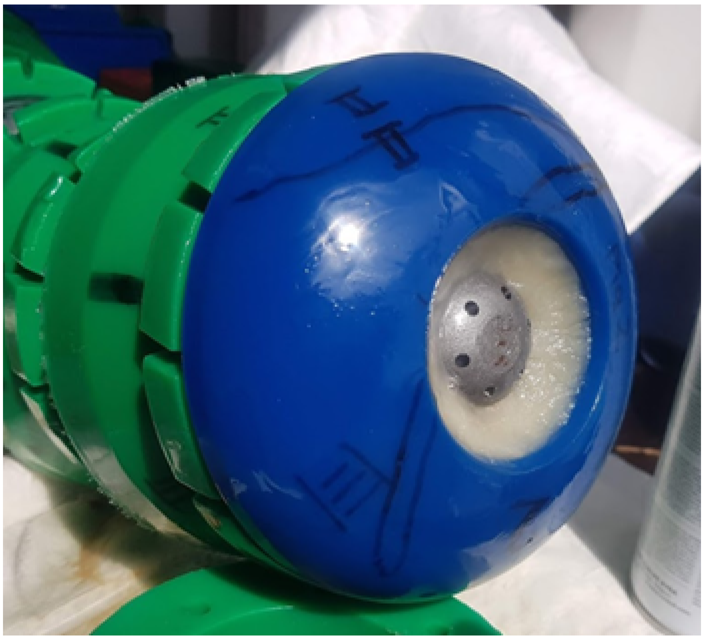


Figure 2.3: Leftover wax collected by the pig during testing. Obtained from [1].

2.1.2 AMPL

An AMPL is a storage unit for the pigs from which the pig is initially launched. The expected lifetime of a pig is around 3-6 months, and an AMPL is necessary to avoid frequent refills of pigs. A 3D-model of an AMPL stacked with pigs is shown in Figure 2.4.

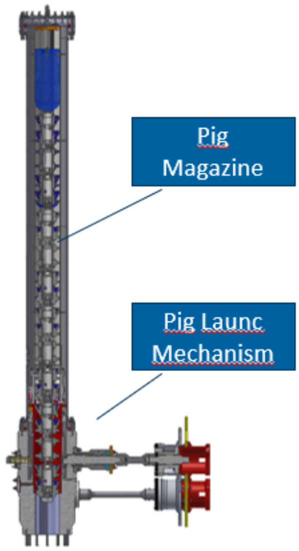


Figure 2.4: A 3D model of an AMPL. Obtained from [4].

2.1.3 Bundled Pipe

A cross-section of a bundled pipe used in the WCS is shown in Figure 2.5. The pig is launched from the AMPL into the pipeline containing hot oil. The pig then travels with the stream of hot oil (represented by red area) while rinsing the pipeline for deposed wax. The bundled pipe is sufficiently long enough for the cooling water to cool the oil down to a temperature ambient. Then, when the oil is cooled down, the pig goes in a loop and continues with the cooled down oil flow. In theory, the wax will not depose to the pipeline wall when the oil temperature is approximately ambient. Therefore, oil is exported without fear of blockage.

Figure 2.6 illustrates the process described above. Initially, the pig is launched from an AMPL located in the leading towhead. The pig travels up to the trailing towhead before turning around and travelling back to the leading towhead.

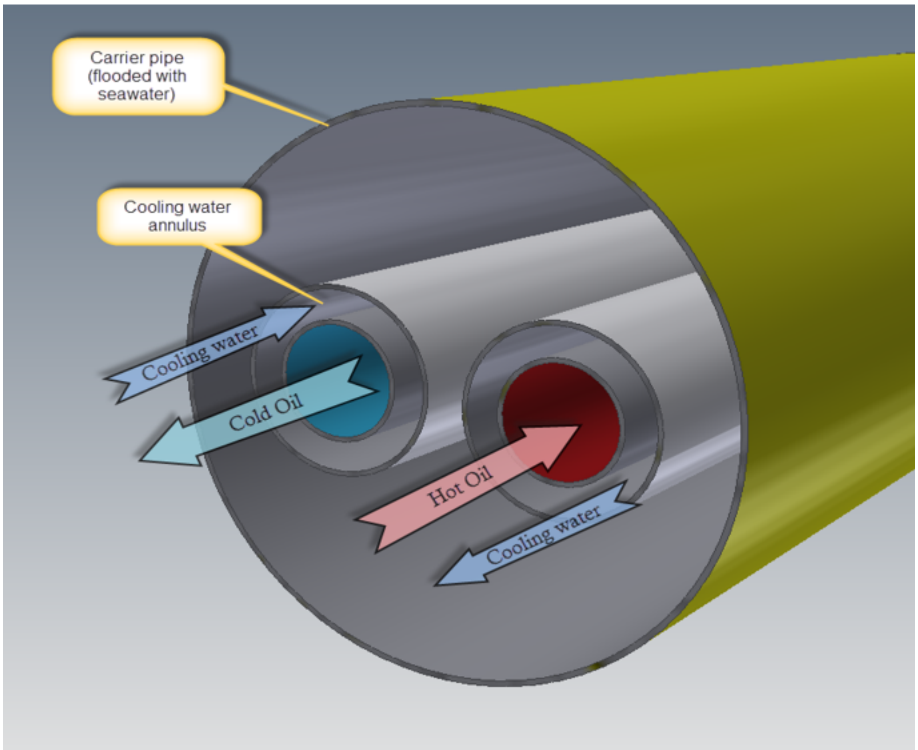


Figure 2.5: A cross-section view of a bundled pipe used in the WCS. Obtained from [4].

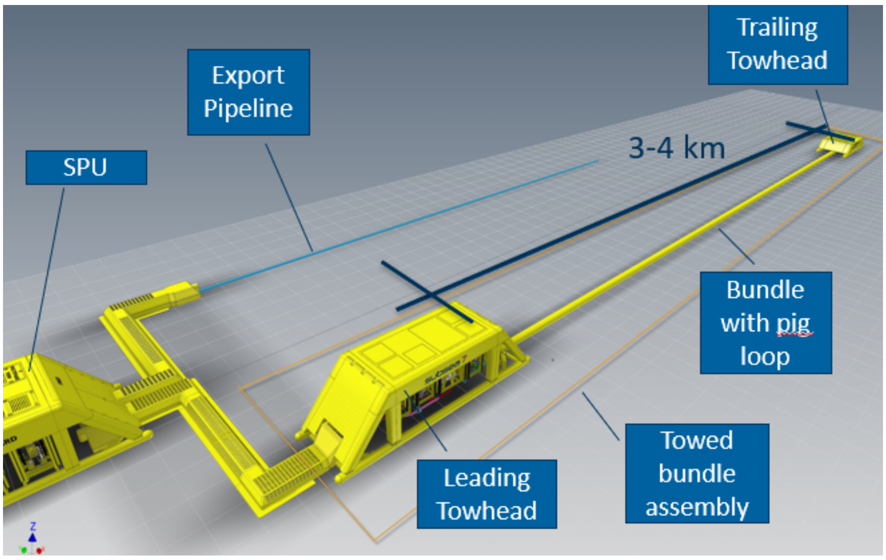


Figure 2.6: The process of cooling oil down to temperature ambient before exportation to process facilities. Obtained from [4].

2.2 Overview

Figure 2.7 shows the Piping and Instrumentation Diagram (P&ID) of the WCS. The whole system is operated remotely with a control system, with multiple sensors to help monitor the system. Hot oil enters the WCS at (1), coming from a Submerged Production Unit (SPU), where most of the sand and water is removed. At this point, the pressure can be increased. Doing so might be necessary for oil transportation over longer distances. A pressure increase will also liquefy the gas present.

The oil flows along the red wave shaped line through a Flow Directional Valve (FDV) (located at (4)), whose job is to change the direction of the flow. The oil then passes the AMPL located at (2). It is from this AMPL the cleaning pig is initially launched.

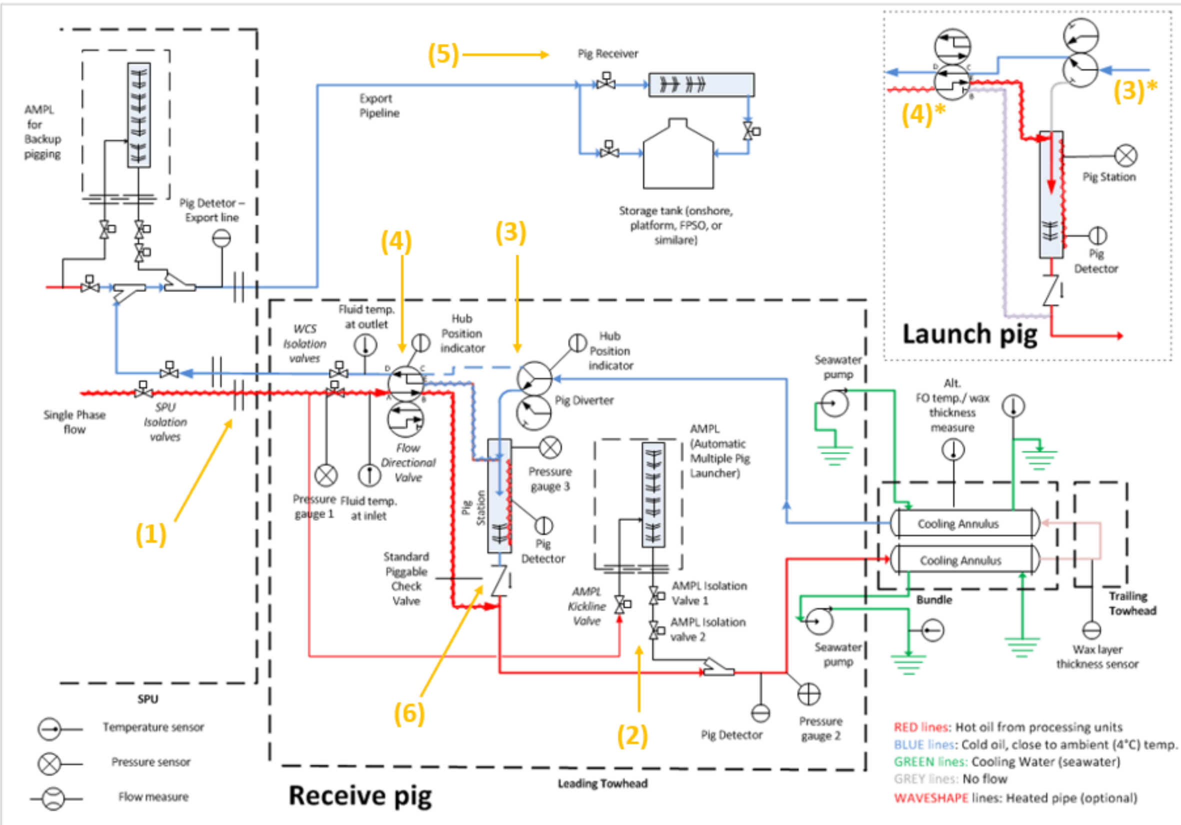
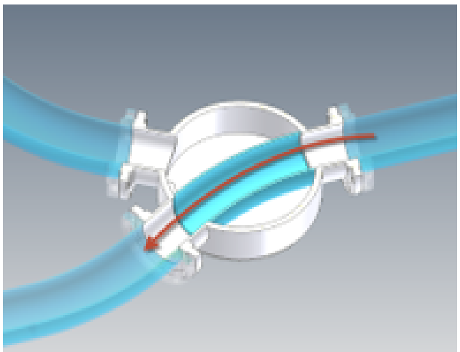
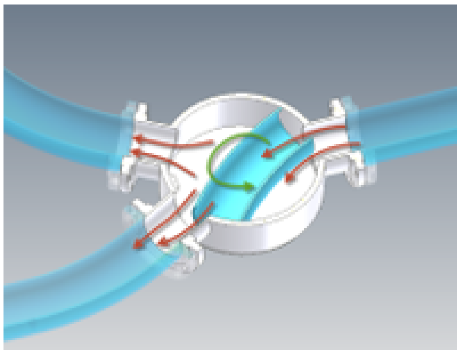


Figure 2.7: The process flow diagram of the WCS. Obtained from [4].

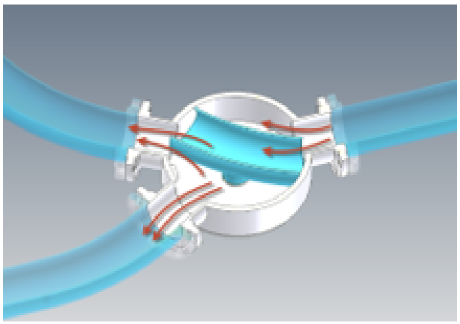
The oil and the pig travel along the red and later blue line, representing hot and cold oil. At (3) the pig and the oil enters a pig diverter. The pig diverter will send the pig and the oil to a pig station if the pig is still usable. The oil will leave the pig station following the blue wave-shaped line. If the pig is at the end of its lifetime, the pig diverter will send the pig and the oil past the pig station and in the export pipeline, via the FDV (shown with (3*). While the export oil is sent to further processing, the pig is sent to a pig receiver (see (5)). The pig diverter is shown in Figure 2.8.



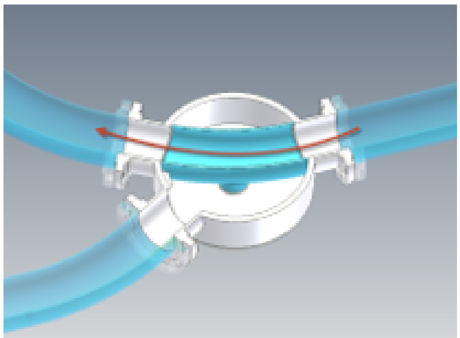
(a) Sending pig to pig station



(b) Rotating counter-clockwise, still allowing oil to flow through



(c) Continuing rotation from (b)



(d) Sending the pig in the export pipeline

Figure 2.8: Functionality of a pig diverter. Obtained from [4].

In the situation where the pig is sent to the pig station, it can be re-entered in the loop at point (6). To do so, the FDV (located at (4)) directs the flow so that it enters the pig station behind the pig. The process of changing the flow direction is shown at (4*). A 3D model of the FDV is shown in Figure 2.9.

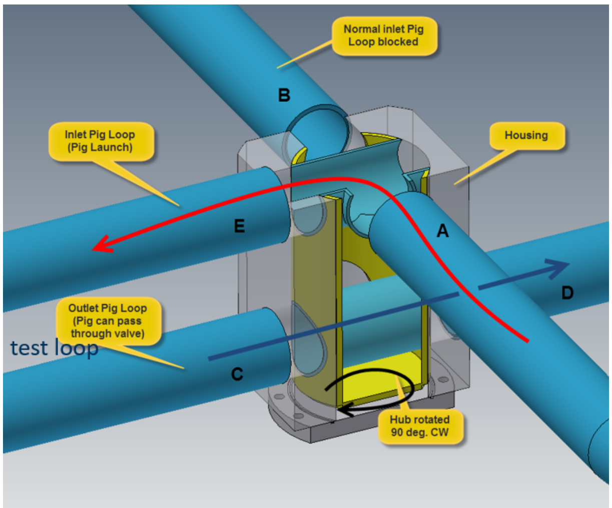
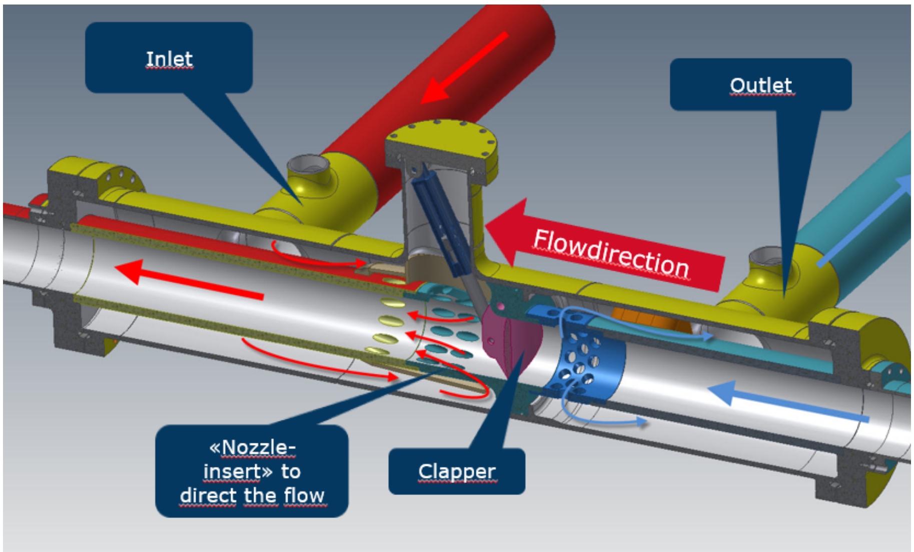
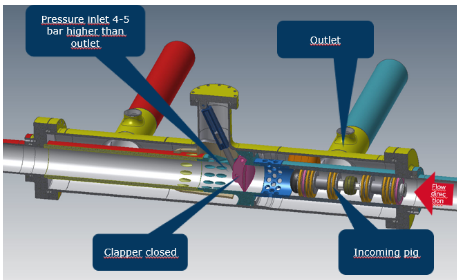


Figure 2.9: A 3D model of the FDV. Obtained from [4].

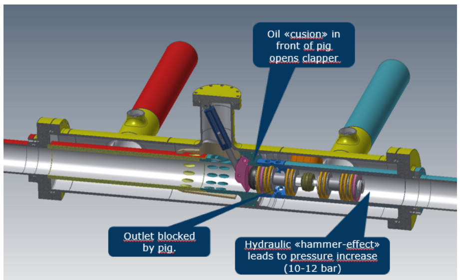
Before re-entering the loop from the pig station, the pig passes through a piggable check valve at (6). This valve is necessary to prevent oil from coming into the pig station from the oil pipeline located below (6) (red-coloured line). An illustration of the pig passing the check valve is shown in Figure 2.10.



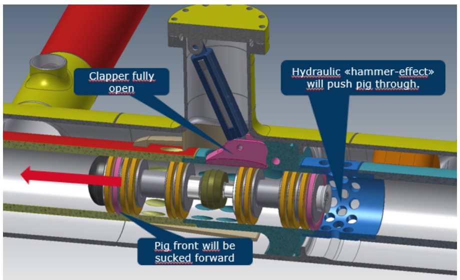
(a) 3D model of the pig check valve. Obtained from [5].



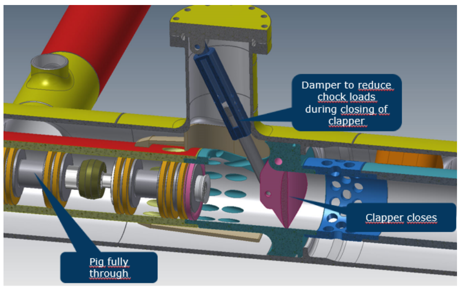
(b) Pig moves towards the clapper



(c) Pig hits clapper



(d) Pig passes clapper



(e) Pig has passed clapper

Figure 2.10: The process when the pig re-enters the loop. Obtained from [5].

Chapter 3

Theory

This chapter discusses relevant theory used in system reliability calculation. Section 3.1-3.7 discusses relevant probability theory and explains some relevant distribution functions. These functions are considered relevant as they are mentioned in risk and reliability literature ([6]–[8]). Section 3.8 discusses relevant reliability theory, in which probability theory is applied.

3.1 Fundamental Probability Theory

A good starting point would be defining the fundamental rule within probability. The sum of all probabilities *(p)* sum to one for all n probabilities related to an event [9].

$$\sum_{i=1}^{n} p_i = 1.$$

Likewise, the probability of an event not happening (here, denoted q_1) is simply one minus the sum of the event happening. This is called the complementary rule [9].

$$q_i = 1 - \sum_{i=1}^{n} p_i.$$

Further, two mutually non-exclusive events (A and B) have some probabilities of outcome in common, denoted $P(A \cap B)$. The probability that either event A or B takes place is denoted $P(A \cup B)$. As $P(A \cap B) = P(A) * P(B)$ for independent events, equation 3.1 can be used to show the relationship. This equation is called the addition rule and is often used in system reliability calculation. A Venn diagram is shown in Figure 3.1 to illustrate this relationship [9].

$$P(A \cup B) = P(A) + P(B) - P(A \cap B),$$
$$P(A \cap B) = P(A \cup B) - P(A) - P(B). \tag{3.1}$$

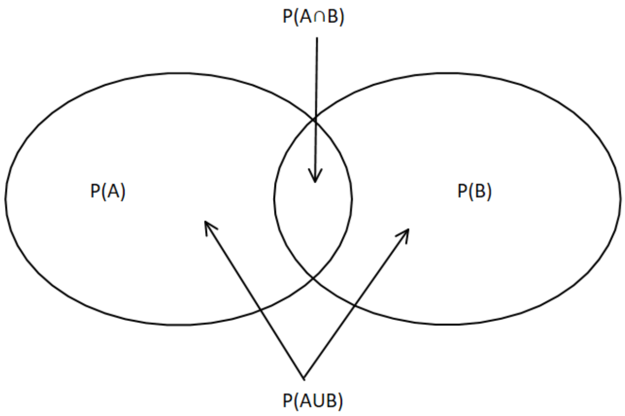


Figure 3.1: Venn diagram of mutually non-exclusive events A and B

As a final notice on fundamental probability theory, we have conditional probability and Bayes' theorem. Bayes' theorem states that the probability for event A to occur, given that event B has occurred ($P(A \mid B)$) is given by [9]–[11]

$$P(A \mid B) = \frac{P(B \mid A) * P(A)}{P(B)}. \tag{3.2}$$

In words, based on some prior knowledge, combined with updated data, we get a posterior probability/distribution. Bayes's theorem is useful when event A and B are dependent on each other. To show why one must first introduce what is called the chain rule [9]–[11].

$$P(A \cap B) = P(A) * P(B \mid A). \tag{3.3}$$

When inserting equation 3.3 in the numerator on the right side of equation 3.2, one will get that [9]–[11]

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \tag{3.4}$$

which is the definition of conditional probability. We see that if event A and B were independent (i.e. $P(A \cap B) = P(A) * P(B)$), one could factorize *P(B)*, and we are left with *P(A) = P(A)*.

3.2 Probability distributions of Random Variables

A probability distribution is the distribution of a random stochastic variable's outcome when the experiment is repeated multiple times. The distribution is a function whose input is a possible value (i.e. a value in the random variables' sample space) and whose output is a probability of occurrence [10].

One generally talks about two different types of probability distributions, namely discrete and continuous probability distributions for discrete and continuous random variables, respectively. These types of random variables are different in the way that discrete random variables have a set of countable values. For example, the number of ways a system can fail. Continuous random variables on the other hand can take any value. For example, the Brent oil price at a given time or the oil flow capacity at a given time) [10].

The distribution function for a discrete random variable (denoted f(x) or f(t)) is called a Probability Mass Function (PMF), while for a continuous random variable it is called a Probability Density Function (PDF). Both have a Cumulative Distribution Function (CDF), however the process of getting them is different. In the case with discrete random variables, the CDF (denoted F(x)) is defined as [10].

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t).$$

For a continuous random variable, the CDF is defined as

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt.$$

Presenting this difference helps explain in better detail the usage of discrete versus continuous random variables. For discrete random variables, one typically wants to find the probability for exact variables. For continuous random variables, one wants to find the probability that the random variable is within an interval [10]. More examples of differences and similarities are shown in Table 3.1. The remaining part of this section presents some relevant PMFs and PDFs and briefly explains under which conditions these are useful.

Table 3.1: Differences and similarities between discrete and continuous random variables

Discrete random variables	**Continuous random variables**
$P(X = a) = b$	$P(X = a) = 0$
$P(X \leq a) \neq P(X < a)$	$P(X \leq a) = P(X < a)$
$f(x) \geq 0$	$f(x) \geq 0$ when $x \in \mathbb{R}$
$\sum_x f(x) = 1$	$\int_{-\infty}^{\infty} f(x)\, dx = 1$
$f(x) = P(X = x)$	$\int_a^b f(x)\, dx = P(a < X < b)$

3.2.1 Probability Mass Functions

Poisson Distribution

The Poisson distribution is applicable when one is interested in seeing how many times an event occurs in a specified area or time interval. The PMF for the Poisson distribution is defined as [10], [11]

$$\begin{aligned} f(x) &= p(x;\ \mu), \\ f(x) &= \frac{\mu}{x!}\, e^{-\mu},\ x = 0,\ 1,\ 2,\ \ldots\ n. \end{aligned} \tag{3.5}$$

Where:

- $\mu = \lambda t$;
- λ: The intensity/expected number of events (in specified time interval/area);
- μ: The expectation (E(X)) and variance (Var(X)).

Binomial Distribution

The binomial distribution is useful in situations when a specified number of trials (n) is repeated. Each trial is either "success" with probability p, or "not success/failure". All trials are also independent of each other, and the output is the probability that in the n

trial, there are x successes. The PMF for the binomial distribution is defined as [10], [11]

$$\begin{aligned} f(x) &= b(x;\ n,\ p), \\ f(x) &= \binom{n}{x} p^x\,(1-p)^{n-x},\ x = 0,\ 1,\ 2,\ \ldots\ n.. \end{aligned} \tag{3.6}$$

Where:

- The expectation (E(X)) $= n * p$ (the number of trails multiplied by probability of success for each trial);
- The variance (Var(X)) $= n*p*(1-p)$ (the expectation multiplied by the probability of failure for each trial).

Negative Binomial Distribution

The negative binomial distribution can be viewed as a special case of the binomial distribution. The output from the negative binomial distribution is the probability that x number of trials from n number of trials are needed before success number k is recorded. The PMF for the negative binomial distribution is defined as [10], [11]

$$\begin{aligned} f(x) &= b^*(x;\ k,\ p), \\ f(x) &= \binom{x-1}{k-1} p^k\,(1-p)^{x-k},\ x = k,\ k+1,\ \ldots\ . \end{aligned} \tag{3.7}$$

Where:

- The expectation (E(X)) $= \frac{k}{p}$ (the number wanted successes over the probability of success for each trial);
- The variance (Var(X)) $= k * \frac{1-p}{p^2}$ (the expectation multiplied by the probability of failure for each trial over the probability of success for each trial).

Geometric Distribution

Finally, the geometric distribution. The geometric distribution is a special case of the negative binomial distribution where k equals one. The output is the probability that x number of trials are needed until the first success. The PMF for the geometric distribution is given by [10], [11]

$$
\begin{aligned}
f(x) &= g(x;\ p), \\
f(x) &= p\,(1-p)^{x-1},\ x = 1,\ 2,\ 3,\ \dots\ n.
\end{aligned}
\tag{3.8}
$$

3.2.2 Probability Density Functions

Normal Distribution

The normal distribution, characterised by its bell-shaped curve, can describe a wide variety of real-life phenomena. In [10, p. 192], the normal distribution is described as "The most important continuous probability distribution in the entire field of statistics..." A reason for this might be that it can be used to describe many phenomena alone. Another reason supporting this statement is that many distributions (both PMFs and PDFs) can under specific conditions be approximated by this distribution. These conditions and how to make these approximations are shown in [10], [11].
The PDF for the normal distribution is given by [10], [11]

$$
\begin{aligned}
f(x) &= n(x;\ \mu,\ \sigma), \\
f(x) &= \frac{1}{\sqrt{2\pi}\sigma}\ \exp\,(-\frac{1}{2}\,\frac{(x-\mu)^2}{\sigma^2}),\ -\infty < x < \infty
\end{aligned}
\tag{3.9}
$$

Where:

- The expectation (E(X)) $= \mu$;
- The variance (Var(X)) $= \sigma^2$.

Integrating and finding the CDF can be quite complicated. Therefore, a transformation is done. The theory used for this transformation is presented in [10, Chapter 7]. In short, let $Z = \frac{X-\mu}{\sigma}$. Then, when this is solved for X and inserted in equation 3.9, one will see that many parameters cancel out, which yields that

$$\frac{1}{\sqrt{2\pi}\sigma}\ \exp\,(-\frac{1}{2}\,z^2).$$

Where:

- Z (the random variable) $\sim$ N(0,1) (is standard normally distributed with expectation $(\mu) = 0$ and variance $(\sigma^2) = 1$).

A table for the standard normal distribution can then be used to find $P(Z < \frac{a-\mu}{\sigma})$.

Log-normal Distribution

The log-normal distribution is a variant of the normal distribution. The random variable X is log-normal distributed in the case when *ln(X)* is normally distributed [10]. According to [6, p. 45], the log-normal distribution can be used to model the repair/maintenance time for a component.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp(-\frac{1}{2} \frac{(ln(x) - \mu)^2}{\sigma^2}), \ for \ x > 0. \tag{3.10}$$

Where:

- The expectation (E(X)) $= e^{\mu + \frac{\sigma^2}{2}}$;
- The variance (Var(X)) $= e^{2(\mu+\sigma^2} - e^{2\mu+\sigma^2}$.

Inverse Gaussian Distribution (Wald Distribution)

If one were to plot the failure rate function as a function of time t (see Section 3.3 on finding these), one would see that the failure rate increases towards a specific time t_1. After this, in the interval $\langle t_1, \infty \rangle$, the failure rate decreases monotonically towards zero [6], [12].

It is suggested in [6, p. 50] that this is an unreasonable assumption, as a decreasing failure rate towards zero is not the case in practical situations. Thus, it is also suggested in [6, p. 50] to use the inverse Gaussian distribution in place of the log-normal distribution to model repair/maintenance time. Like the log-normal distribution, the failure rate function increases towards a value t_1. Unlike the log-normal distribution, this failure rate function decreases towards a limit in the time interval $\langle t_1, \infty \rangle$, which depends on the parameters. [6] presents the following PDF for the inverse-Gaussian distribution.

$$f(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp -((\frac{\lambda}{2\mu^2})(\frac{(x-\mu)^2}{x})), \ for \ x, \ \lambda, \ \mu > 0. \tag{3.11}$$

Where:

- The expectation (E(X)) $= \mu$;
- The variance (Var(X)) $= \frac{\mu^3}{\lambda}$.

Exponential Distribution

The exponential distribution is often used in situations when one wishes to model various lifetimes, and its PDF is defined as [10], [11]

$$f(x) = \frac{1}{\beta}\, e^{-\frac{x}{\beta}},\ for\ x \geq 0,$$
$$f(x) = \lambda\, e^{-x\lambda},\ for\ x \geq 0. \tag{3.12}$$

The latter expression shows the PDF often used in reliability calculation, where $\lambda = \frac{1}{\beta}$. Further it can be shown that:

- The expectation (E(X)) $= \beta$;
- The variance (Var(X)) $= \beta^2$.

The exponential distribution has a property that makes it unique among the continuous probability distributions, namely the memoryless property. This is best shown by using the convolution formula (shown in equation 3.4). Similar proof of the memoryless property for the exponential distribution can also be found in [11, p. 288].

$$\begin{aligned}
P(T > s + u \mid T > u) &= \frac{P((T > s + u \mid T > u) \cap (T > u))}{P(T > u)} \\
P(T > s + u \mid T > u) &= \frac{P(T > s + u)}{P(T > u)} \\
P(T > s + u \mid T > u) &= \frac{e^{-\frac{(s+u)}{\beta}}}{e^{-\frac{u}{\beta}}} \\
P(T > s + u \mid T > u) &= e^{\frac{-s}{\beta}} \\
P(T > s + u \mid T > u) &= P(T > s).
\end{aligned} \tag{3.13}$$

In words, given that a component is working after a time u, the probability of it working after a time s+u is the same as the probability of it working after a time u. The component is time-independent. It does not deteriorate or improve over time.

Another property can be found when comparing the PMF of the Poisson distribution and the PDF of the exponential distribution. They look quite similar, and their relation is this: The time between events in a Poisson process are exponentially distributed [10], [11].

Gamma Distribution

The gamma distribution is another distribution used to model lifetimes. Its PDF is defined as [10], [11]

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-\frac{x}{\beta}}\cdot,\ for\ x \geq 0. \tag{3.14}$$

Where:

- α: Shape parameter;
- β: Scale parameter;
- The expectation (E(X)) $= \alpha * \beta$;
- The variance (Var(X)) $= \alpha * \beta^2$;
- $\Gamma(\alpha)$: The gamma function.

As a final note, some relation between the gamma distribution and other distributions. Firstly, for $\alpha = 1$, one gets the exponential distribution. Secondly, the gamma distribution is used to describe the time until event number k in a Poisson process. When using this feature, $\alpha = k$ [10], [11]. Finally, another relation between the gamma distribution and exponential distribution is that in the case with n independent exponentially distributed random variables, the sum of those (i.e. $Y = \sum_{i=1}^{n} X_i$) has a gamma distribution with parameters $\alpha = n$ and β. This is further explained in [10, Ch. 7]

Chi-Squared Distribution

If one let $a = \frac{\nu}{2}$ and $\beta = 2$ in a gamma distribution, one will obtain what is called the chi-squared (χ^2) distribution [10].

$$f(x) = \frac{1}{2^{\frac{\nu}{2}}\,\Gamma(\frac{\nu}{2})}\, x^{\frac{\nu}{2}-1}\, e^{-\frac{x}{2}},\ for\ x > 0 \tag{3.15}$$

Where:

- ν: Degrees of freedom;
- The expectation (E(X)) $= \nu$;
- The variance (Var(X)) $= 2 * \nu$.

The χ^2-distribution is added here because some of the data presented in the OREDA-handbook are the upper 95% and lower 5% percentile in the χ^2-distribution. Also, notice that for $\nu = 2$, one will get the exponential distribution with $\beta = 2$.

Weibull Distribution

The Weibull distribution is just like the exponential and gamma distribution used to model lifetimes. Its PDF is given by [10]

$$f(x) = \alpha\beta x^{\beta-1} \, e^{-\alpha x^{\beta}}, \; for \; x \geq 0. \tag{3.16}$$

Where:

- α: Scale parameter;
- β: Shape parameter;
- The expectation (E(X)) = $\alpha^{\frac{-1}{\alpha}} * \Gamma(1 + \frac{1}{\beta})$;
- The variance (Var(X)) = $\alpha^{\frac{-2}{\beta}} * (\Gamma(1 + \frac{2}{\beta} - \Gamma(1 + \frac{1}{\beta})^2)$.

Likewise, as for the gamma distribution, when the shape parameter (here, β) = 1, one gets the exponential distribution.

Birnbaum-Saunders Distribution (Fatigue Life Distribution)

Finally, regarding relevant continuous distribution functions, there is the Birnbaum-Saunders Distribution. It is also known as the fatigue life distribution, as it can be used in situations when a component is exposed to cyclic stress. Its PDF is defined as [6]

$$f(x) = \frac{\sqrt{\beta x} + \frac{1}{\sqrt{\beta x}}}{2\sqrt{2\pi}\alpha x} \; \exp(-\frac{(\sqrt{\beta x} - (\frac{1}{\sqrt{\beta x)}}))^2}{2\alpha^2}), \; for \; x > 0. \tag{3.17}$$

Where:

- α: Shape-parameter;
- β: Scale-parameter;
- The expectation (E(X)) = $\frac{1}{\beta}(1 + \frac{\alpha^2}{2})$;
- The variance (Var(X)) = $\frac{\alpha^2}{\beta^2}(1 + \frac{5\alpha^2}{4})$.

3.3 Hazard Rate

A distribution function's hazard rate (in the following called the failure rate) is a measure used to help determine a components deterioration or wear over time. The failure rate function is used to say something about the probability of failure at a point in time *t*, given that the component has survived so far [10], [11]. The failure rate function *Z(t)* is defined as

$$
\begin{aligned}
Z(t) &= lim_{\Delta t->0} \frac{F(t+\Delta t) - F(t)}{\Delta t} \frac{1}{R(t)} \\
Z(t) &= \frac{F'(t)}{R(t)} \\
Z(t) &= \frac{f(t)}{R(t)} \\
Z(t) &= \frac{f(t)}{1 - F(t)}.
\end{aligned}
\tag{3.18}
$$

Where:

- *Z(t)*: The failure rate function;
- *f(t)*: The PDF for the distribution;
- $F(t) = P(T \geq t)$: the CDF, often called the failure function;
- $R(t) = P(T < t)$: the survival function.

As an example, it can be derived that the Weibull distribution has the failure rate function [10, p. 225]

$$Z(t) = \alpha\beta t^{\beta-1}.$$

As previously shown in this subsection, $\beta = 1$ in the Weibull distribution yields the exponential distribution. As a result, one sees that the failure rate function for the exponential distribution is a constant (α), and hence time-independent.

3.4 Extreme Variables and Order Variables

[11]
Consider a set of random variables X_1, X_2, $\ldots$, X_n. These are assumed to be independent and identically distributed. When these are sorted in increasing order (i.e. X_1, X_2, $\ldots$, X_n, given that $X_1 \leq X_2 \leq \ldots \leq X_n$), they are called order variables [11].

Further, extreme variables are those that are minimum or maximum values from the set of ordered variables.

$$X_1 = min(X_1, X_2 \ldots, X_n)$$

$$X_n = max(X_1, X_2, \ldots, X_n)$$

From this, the distribution functions and PDF (for continuous random variables) can be determined for X_1 and X_n. These functions are presented below. How these functions are derived can be found in [11, pp. 388-389]. The probability function for X_1 and X_n respectively is given by

$$F_1(x) = 1 - [1 - F(x)]^n, \ -\infty < x < \infty,$$
$$F_n(x) = [F(x)]^n, \ -\infty < x < \infty$$

The PDF for X_1 and X_n can be found by deriving $F_1(x)$ and $F_n(x)$, respectively.

$$f_1(x) = n\, f(x)\, [1 - F(x)]^{n-1}, \ -\infty < x < \infty,$$
$$f_n(x) = n\, f(x)\, [F(x)]^{n-1}, \ -\infty < x < \infty$$

These results can be used to derive equations for calculating system reliability for series and parallel structures.

3.5 Estimation

Generally, parameters in various distribution functions must be estimated. This can be done from relevant data sets, and multiple methods exist to do this.

3.5.1 Confidence Intervals

Often when estimates are found, there will be some degree of uncertainty in the estimate. Confidence interval is a useful tool that can give indications about the degree of uncertainty in the estimate. The interpretation of the confidence interval is that when an experiment is repeated infinitely many times, the unknown parameter will fall within this interval $(1-\alpha)$*100% of the times. For instance, a 95% confidence interval contains the value of the unknown parameter 95% of the times [10].

$$P(\theta_L < \theta < \theta_U) = 1 - \alpha,$$

is the (1-α)100% confidence interval for the interval $[\hat{\theta_L}, \hat{\theta_U},]$, where $\hat{\theta_L}$, and $\hat{\theta_U}$, is the lower and upper limit in the interval, respectively.

The general procedure for finding a confidence interval for the unknown parameter θ is as follows: Assuming that the estimator for θ is found (i.e. $\hat{\theta}$) and a expression $h(\hat{\theta}$,$\theta)$ that contains both θ and $\hat{\theta}$), with a known distribution. This distribution is usually found by using transformations and linear combinations for functions of random variables (see [10, Ch. 7]. Then [10]:

1. $P(w_{1-\frac{\alpha}{2}} \leq h(\hat{\theta}, \theta) \leq w_{\frac{\alpha}{2}}) = 1 - \alpha$;
2. Solve inequality with respect to θ, so that: $P(\theta_L < \theta < \theta_U) = 1 - \alpha$.

Where:

- $w_{1-\frac{\alpha}{2}}$: The lower quantile, found in tables;
- $w_{\frac{\alpha}{2}}$: The upper quantile, found in tables;
- α: probability that the parameter will not be in the interval;
- $1.\alpha$: probability that the parameter will be in the interval.

3.5.2 Maximum Likelihood Estimation

Maximum Likelihood Estimation (MLE) is an estimation technique where the idea is to first find the distribution's likelihood function with respect to the unknown parameter and then maximize this function. Following a step-by-step process, one will get an expression for an estimator that maximizes the probability that the distribution function produce data that is observable in real life [8], [10]. The step process is shown below [13]–[15]:

1. Find the distribution's likelihood function with respect to the parameter in question ($L(\theta) = \Pi_{i=1}^{n} f(x_i, \theta)$).
2. Find the log-likelihood function by taking the logarithm of the likelihood function with respect to the parameter ($l(\theta)$).
3. Take the derivative of the log-likelihood function with respect to the parameter ($\frac{\partial l(\theta)}{\partial \theta}$),
4. Solve the equation from step 4 to find an expression for θ (now $\hat{\theta}$, indicating that θ now is an estimator).
5. Finally, one should check if $\hat{\theta}$ indeed is a maximum and not a minimum. This is done by taking the second derivative of the log-likelihood function and see if it is less than zero ($\frac{\partial^2 l(\theta)}{\partial \theta^2} < 0$).

Furthermore, some comments regarding the MLE. It is emphasised in [10] that it is necessary with knowledge about the distribution when applying the MLE method. It is shown in [10] through examples that the estimator is not necessarily unbiased in the limit or asymptotically, even though bias goes to zero as the sample goes to infinity. An estimator $\hat{\theta}$ is unbiased if $\mathrm{E}(\hat{\theta}) = \theta$ [10, p. 332].

Finally, some properties for the maximum likelihood estimator [13]–[15]:

- The estimator has the property of asymptotic normality, meaning that $\frac{\hat{\theta}-\theta}{\sqrt{Var(\hat{\theta})}} \approx N(0,1)$.
- Its variance can be approximated as: $Var(\hat{\theta}) \approx \frac{1}{J(\theta)}$, where $J(\theta) = \frac{\partial^2 l(\theta)}{\partial \theta^2}$.
- The above-mentioned properties can be used to construct an approximate confidence interval, in this case called a Wald interval: $(\hat{\theta} - z_{\frac{\alpha}{2}}\sqrt{\frac{-1}{J(\hat{\theta})}}, \hat{\theta} + z_{\frac{\alpha}{2}}\sqrt{\frac{-1}{J(\hat{\theta})}})$.
- Consistency. It is more likely that the estimate is close to the true value, the bigger the sample size is: $\mathrm{Var}(\hat{\theta}) \rightarrow 0$ and $\mathrm{E}(\hat{\theta}) \rightarrow \theta$, when $n \rightarrow \infty$.
- Invariance. let $\mathrm{g}(\theta)$ be strictly decreasing or increasing. Then: The maximum likelihood estimator of $\mathrm{g}(\theta) = \mathrm{g}(\hat{\theta})$.

As shown in the second bullet point, approximations can be made to create confidence intervals for the maximum likelihood estimator. These approximations make it much easier to create confidence intervals, as one only needs the log-likelihood function $\mathrm{l}(\theta)$.

3.5.3 Bayesian Statistics

Using Bayesian statistics makes it possible to combine some subjective beliefs (for example, based on knowledge from similar situations) with data. This updated distribution function that now combines knowledge and data is called a posterior distribution [8], [10], [16]. Recall equation 3.2. The posterior distribution can be found by

$$p(\theta \mid x) = \frac{f(x \mid \theta) * p(\theta)}{p(x)}. \tag{3.19}$$

Where:

- $p(\theta \mid x)$: Posterior distribution;
- $f(x \mid \theta)$: Distribution for the observed data;
- $p(\theta)$: Prior distribution, distribution based on knowledge;
- $p(\theta \mid x) \propto f(x \mid \theta)p(\theta)$
- $p(x) = \int_\theta f(x \mid \theta)p(\theta)\, \mathrm{d}\theta$: Often a constant;
- θ: Parameter in question/relevant parameter.

In the case when multiple observations exist, equation 3.19 can be written as [10], [16]

$$\begin{aligned} p(\theta \mid data) &= \frac{f(data \mid \theta) * p(\theta)}{p(data)}, \\ p(\theta \mid data) &= L(\theta) * p(\theta). \end{aligned} \tag{3.20}$$

Where $L(\theta)$ is the likelihood function of the prior distribution. The Bayes estimator for the parameter θ is defined as the expectation of the random variable in the posterior distribution [10], [16].

$$\begin{aligned} \hat{\theta}_{Bayes} &= E(\theta \mid data), \\ \hat{\theta}_{Bayes} &= \int_{\theta} \theta p(\theta \mid data) \, \mathrm{d}\theta. \end{aligned} \tag{3.21}$$

Finally, the posterior distribution can, in turn, be used to create a credibility interval (Bayesian interval). Unlike a confidence interval, the interpretation of a credibility interval is that there is a subjective probability that the parameter is in the interval $[\theta_L, \theta_U]$ [10], [16].

$$P(\theta_L < \theta < \theta_U \mid data) = 1 - \alpha.$$

Once the posterior distribution is found, the method of finding a credibility interval is the same as for confidence intervals.

3.6 Queueing Systems

A queueing system is a system in which someone or something arrives or fail and then join a queue of other arrived people or components. The length of the queue is determined by the number of servers or repair crews, the number of arrivals and the time it takes for the servers to complete the service. A typical type of queue is the policy of "first come, first serve" (i.e. the first in line is served first). Other types of queues also exist. For example, one queue policy can be based on priority (i.e. something or someone has higher priority over another) [11], [17].

The notation used in queue theory is shown in Figure 3.2. In this figure, the continuous probability distribution discussed previously is sorted according to typical applications, but the distributions are not solely limited to the mentioned applications [11], [17].

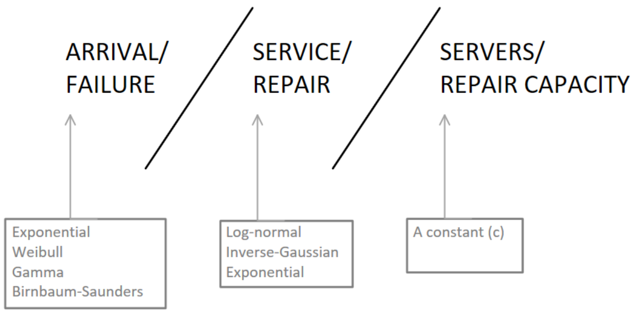


Figure 3.2: Notation in queue theory, and how the above-mentioned continuous probability distributions relate to this

This notation indicates the the "first come, first serve" policy is used, and different notations are used to describe different types of queue systems. Below are some examples of this notation, with explanations [11]:

- M/M/1;
- G/GI/2;
- GI/D/c.

Where:

- M: Exponentially distributed times between arrivals and or service;
- G: General distributed; times between arrivals and or service
- GI: General independent; distributed times between arrivals and or service;
- D: Deterministic (constant and non-random) times between arrivals and or service;
- c: Constant (servers, repair crews etc.)

Finally, a queue is considered stable if the following is true [11].

$$E(T_i) > \frac{E(C_i)}{c},$$
$$\lambda < c * \mu.$$

3.7 Markov Chains

A Markov chain is a tool used to study the state of someone or something in the future, which is only based on the present state and independent of the past. More formally, one can say that a Markov chain is a stochastic process with finite or infinite state space. The future state of a random variable (denoted X_{n+1} or X_{t+1}) is solely based on the present state of that random variable (denoted X_n or X_t) [11], [18].

There are a total of four different types of Markov chain, as shown in Table 3.2. Only the two cases with discrete state space will be discussed in the following. The continuous-time Markov chain will be emphasized the most, while the discrete-time Markov chain will be used more as a further introduction through an example.

Table 3.2: The four different types of Markov chain

	State space	
Time	(1) Discrete time Discrete state space	(2) Discrete time Continuous state space
	(3) Continuous time Discrete state space	(4) Continuous time Continuous state space

As a final note in this introduction, several descriptions are used to classify Markov Chains. For instance, the chains and their states can be described as irreducible, recurrent, transient, communicating, absorbing, and have a period. These descriptions will not be explained in detail in this thesis, but some will be explained in the discussion about steady-state probabilities. The reader is referred to [11, Chapter 12, pp. 528-559] for a detailed explanation of these terms.

3.7.1 Discrete Time-Discrete State Space Markov Chains

The introduction to Markov chains, presented above, can be mathematically expressed as

$$P(X_{n+1} = j \mid X_n = i,\, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0),$$
$$P(X_{n+1} = j \mid X_n = i).$$

Let $\mathcal{S}$=0, 1, 2,... denote the state space (i.e. all the possible states for the Markov chain). This can be both infinite or finite. Assume then that probability of moving from state i to j (the transition probability) for all $i,\, j \in \mathcal{S}$ is the same for all values of $n = 0,\, 1,\, 2\, 3,\, \ldots,$. Here, n denotes the number of steps the random variable X has, and one can say that the transition probabilities in such a case are stationary. Further, the following notation is used for the transition probabilities.

$$p_{ij} = P(X_{n+1} = j \mid X_n = i).$$

Where p_{ij} is the probability of moving from state i to j. These transition probabilities are grouped in what is called a transition probability matrix. Below is a transition probability matrix with a finite state space ranging to n [11], [17]–[19].

$$P = \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0n} \\ p_{10} & p_{11} & \cdots & p_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n0} & p_{n1} & \cdots & p_{nn} \end{bmatrix}. \tag{3.22}$$

Consider a case where the oil price changes from day 0 to day 1. Assume then that one is interested in whether or not the brent oil price is above or below 50$ per barrel.

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix}.$$

From this matrix, the interpretation of the transition probabilities is as follows:

- P_{00}=P(oil price<50$ tomorrow | oil price<50$ today)
- P_{01}=P(oil price>50$ tomorrow | oil price<50$ today)
- P_{10}=P(oil price<50$ tomorrow | oil price>50$ today)
- P_{11}=P(oil price>50$ tomorrow | oil price>50$ today)

Notice that the sum of each row equals one (i.e. $\sum_{j=0}^{\infty} p_{ij} = 1$). Finally, a useful tool to better understand the transition probability matrix is a transition graph. The transition graph for this transition probability matrix is shown in Figure 3.3.

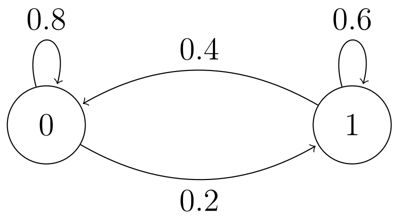


Figure 3.3: Transition graph for the example with brent oil pirce changes

In this situation, the brent oil price change for one day at a time is considered. If one wishes to find the price two or m days in the future, one can multiply the transition probability matrices [11].

$$P^m = P * P * P \ldots = \begin{bmatrix} p_{00}^m & p_{01}^m & \cdots & p_{0n}^m \\ p_{10}^m & p_{11}^m & \cdots & p_{1n}^m \\ \vdots & \vdots & \ddots & \vdots \\ p_{n0}^m & p_{n1}^m & \cdots & p_{nn}^m \end{bmatrix}.$$

Here,

$$p_{ij}^m = P(X_{n+m} = j \mid X_n = i).$$

3.7.2 Continuous Time-Discrete State Space Markov Chains

In the case with the discrete-time Markov chain, A random variable jumps in countable steps to new states with a probability. In this case, however, the interpretation is slightly different. The random variable *X(t)* (for $t \geq 0$) with state-space $\mathcal{S}$=0, 1, 2,... will enter a state at a time t and remain in that state for a period of time according to an exponential distribution [11], [17], [18].

Again, assuming that transition probabilities are stationary, then for s and t>0, the continuous-time Markov chain can be expressed as [11], [17], [18]

$$\begin{aligned} P(X(s+t) = j \mid X(s) = i,\, X(u) = x_u,\ for\ 0 \leq u < s), \\ = P(X(s+t) = j \mid X(s) = i). \end{aligned}$$

From this, one can see that the Markovian property holds. For the points in time u and s, where u>s and $s \geq 0$, the future state at *X(u)* only depends on the present state at *X(s)* [11]. Further, let T denote the spent in one state i before it enters a new state *j*. Then [11]

$$\begin{aligned} P(T > s+t \mid T > s) &= P(X(u) = j,\, s < u \leq s+t \mid X(s) = i,\, X(\tau) = i,\, 0 \leq \tau < s), \\ P(T > s+t \mid T > s) &= P(X(u) = j,\, s < u \leq s+t \mid X(s) = j) = P(T > t). \end{aligned}$$

Here, the expression $X(\tau) = i,\, 0 \leq \tau < s)$ is considered irrelevant. This is due to the fact the process is Markovian. Further, one can recognise the expression *P(T>s+t| T>s)* from the discussion about the exponential distribution and its memoryless property. This indicates that the time spent in each state is exponentially distributed and that a continuous-time Markov chain is modelled as a Poisson process [11].
To finalise this discussion, for discrete-time Markov chain, p_{ij} was denoted the probability of moving from state i to *j*. This notation is used for continuous-time Markov chains as well, but it is defined somewhat different. Here it is not defined as just a number between 0 and 1 [11].

$$p_{ij} = \frac{q_{ij}}{\nu_i},\ j \neq i.$$

Where:

- q_{ij}: The specific transition rate (the rate of going from state i to j);
- ν_i: The total rate out from state i (ν_i=$\sum_{j \in \mathcal{S}, j \neq i} q_{ij}$).

The total rate out can, in turn, be used to find the expected time spent in each state: $E(T) = \frac{1}{\nu_i}$.

The Kolmogorov Forward equation

For the continuous-time Markov chain, one might be interested in finding the probability of moving from state i to j at a time t ($p_{ij} = P(X(s+t) = j \mid P(X(s) = i)$. This is analogous to p_{ij}^n in the case of discrete-time (i.e. probability of moving from i to j in n number of steps). To find this, one can utilise the Kolmogorov forward equation [11], [18].

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{ij}(t) = \sum_{k \neq j} q_{kj} p_{ik}(t) - \nu_j p_{ij}(t), \; i, \, j \in \mathcal{S}. \tag{3.23}$$

The initial conditions and boundary conditions used when solving the Kolmogorov forward equation is that [11]

$$\begin{aligned} p_{ii}(t=0) &= 1, \\ p_{ij}(t=0) &= 0, \; when \; i \neq j, \\ t &\geq 0. \end{aligned}$$

Proof of the Kolmogorov forward equation can be found in [11, pp. 570-571].

3.7.3 Steady-State Probabilities

The steady-state probability is the probability of being in a state j as time goes to infinity. This probability is denoted pi_j as $lim_{t \to \infty} p_{ij}(t)$. Steady-state probabilities exist if [11]:

- All states in the Markov chain communicate with each other (i.e. there is a probability of moving from state i to j and from j to i for all states). The chain is then said to be irreducible.
- When in a state, return to that state is guaranteed, and the number of transitions to do so is finite. The chain is then said to be positive recurrent.
- For discrete-time Markov chains: The period (i.e. the number of transitions before returning to the same state) is one. The chain is aperiodic.

Steady-State Probabilities: Discrete Time Markov Chains

In the case with discrete-time, one can observe that when the transition probability matrices P is multiplied by itself m-times as $m \to \infty$, the starting point (P^1) decreases in importance. It is presented in [11, p. 555] that the steady-state probabilities can be found by

$$P(X_{n+1} = j) = \sum_{i=0}^{\infty} P(X_{n+1} = j \mid X_n = i)\, P(X_n = i),$$
$$P(X_{n+1} = j) = \sum_{i=0}^{\infty} p_{ij}\, P(X_n = i).$$

When $lim_{n\to\infty} P(X_n = j)$. When n goes to infinity, one will end up with the equation

$$\pi_j = \sum_{i=0}^{\infty} p_{ij}\pi_i.$$

On matrix form, this can be written as

$$\Pi = P^T \Pi,$$
$$\begin{bmatrix} \pi_0 \\ \pi_1 \\ \vdots \\ \pi_n \end{bmatrix} = \begin{bmatrix} p_{00} & p_{10} & \cdots & p_{n0} \\ p_{01} & p_{11} & \cdots & p_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ p_{0n} & p_{1n} & \cdots & p_{nn} \end{bmatrix} * \begin{bmatrix} \pi_0 \\ \pi_1 \\ \vdots \\ \pi_n \end{bmatrix}.$$

Where:

- P^T: The transposed transition probability matrix;
- $\pi_0 + \pi_1 + \ldots + \pi_n = 1$.

Steady-State Probabilities: Continuous Time Markov Chains

As for the case with continuous-time and discrete state space, the steady-state probabilities can be found by utilising the Kolmogorov forward equation [11].

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{ij}(t) = \sum_{k\neq j} q_{kj} p_{ik}(t) - \nu_j p_{ij}(t),\ i,\, j \in \mathcal{S}.$$

Then, when $t \to \infty$, one will get that

$$0 = \sum_{k\neq j} q_{kj}\pi_k - \nu_j \pi_j,$$
$$\sum_{k\neq j} q_{kj} = \pi_k - \nu_j \pi_j.$$

The interpretation of this is that the steady-state probabilities are found by balancing the transition rates out of state j with the transition rate to state j. For a system, one will get a set of equations that can be solved with respect to π_0, $\pi_1, \ldots, \pi_n$. The fact that the sum of π_0, $\pi_1, \ldots, \pi_n = 1$ can simplify these calculations.

3.7.4 Further In-depth Theory

This subsection briefly presents two topics that can give a deeper understanding of aspects related to reliability calculations.

Renewal Theory

Renewal theory is the theory used to describe when components should be repaired and the expected number of repairs needed in a component's lifetime, amongst other things. An in-depth description of renewal theory, renewal processes and examples of situations where various assumptions are made can be found in [6, pp. 247-277] and [18, Ch.3 3]. Here, only some equations used will be presented. This is done to show the usage of renewal theory, and renewal processes [6, p. 248].

The time until renewal number n is given by

$$S_n = T_1 + T_2 + T_3 + \ldots + T_n = \sum_{i=1}^{n} T_i.$$

This can be used to find the number of renewals in a time interval from 0 to and including t (i.e. (0, t]).

$$N(t) = max(n : S_n \leq t).$$

N(t) is, in turn, used to define the renewal function.

$$W(t) = E(N(t)).$$

The renewal density is then given by

$$w(t) = \frac{d}{dt} W(t).$$

Finally, the following equation can be used to describe the mean number of renewals in the time interval (t_1, t_2).

$$W(t_2) - W(t_1) = \int_{t_1}^{t_2} w(t)dt.$$

Semi-Markov Processes

A semi-Markov process is different from a Markov chain in the way that the time it takes before a state change from i to j has its own distribution. This indicates that the markovian property (i.e. future state only depends upon the present and not past) is not present. Hence, the exponential distribution (and its memoryless property) is no longer the only distribution in use [18].

More detailed, for a continuous-time Markov chain, the queue notation used is $M/M/c$, indicating that failure and repair are exponentially distributed, and time between events is described as a homogeneous Poisson process. For a semi-Markov process, however, the queue notation can be presented as $(G/GI)/(G/GI)/c$, indicating that any general probability distribution can be used to describe both failure and repair time. As a consequence, the time between events is now described by an inhomogeneous Poisson distribution. More thoroughly theory about semi-Markov processes can be found in [18, Sec. 4.8]. Here, and in [6, p. 354] it is shown that the distribution of the time spent in a state i before moving to a state j is

$$F_i(t) = \sum_{j=0,\, j\neq i}^{r} P_{ij}\, F_{ij}(t).$$

Where:

- P_{ij}: Transition probability from state i to j;
- $F_{ij}(t)$: Distribution function describing the time until next transition from state i to j, when being in state i.

Also, the expected time in a state i is defined as

$$\mu_i = \int_0^\infty \mathrm{d}F_i(t).$$

Finally, it can be shown that if the semi-Markov process is positive recurrent and irreducible, then

$$\begin{aligned} \pi_j &= \sum_i \pi_i\, P_{ij}, \\ P_i &= \frac{\pi_i \mu_i}{\sum_j \pi_j \mu_j}. \end{aligned} \tag{3.24}$$

Where P_i denotes the proportions of transitions into state i. This is shown in [18, p. 215] to be equal to the long-run proportion of time in state i.

3.8 Reliability

Aven defines reliability as "... the ability of a component or a system to perform a specific function" [8, p. 5]. To get an overview and calculate the reliability of a system, one often uses a Reliability Block Diagram (RBD). In such diagrams, each component in a system is represented by a block, with one assigned probability, multiple probabilities (if the component has multiple states), or a CDF. Examples of various diagrams are shown in Figure 3.4-3.6 and explained in this section [7], [8].

Some notations that will be used throughout this section and thesis is presented below [7], [8]:

- h: System reliability;
- $g = 1 - h$: System unreliability;
- p_i: Reliability of component i;
- $q_i = 1 - p_i$: Unreliability of component i.

3.8.1 Series Structures

A series structure is a structure where all components must function in order for the system to function. An example of such as system is shown in Figure 3.4. This system's reliability is calculated by multiplying the reliability (i.e. probability that the component is functioning) for each component [7], [8], [11], [17].

$$\begin{aligned} h &= p_1 * p_2 \ldots * p_n, \\ h &= \prod_{i=1}^{n} p_i. \end{aligned} \tag{3.25}$$

A proof of equation 3.25 can be found in [11, Example 9.3, pp. 375-376]. This is done using the theory about extreme variables, presented in Section 3.4.

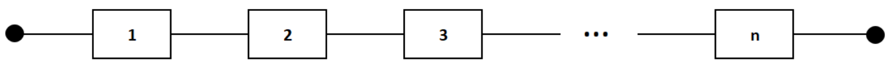


Figure 3.4: Series structure in a RBD. Inspired by [8, Fig. 3.1, p. 88].

3.8.2 Parallel structures

Unlike a series structure, a parallel structure functions as long as at least one component is functioning. An example of a parallel structure is shown in Figure 3.5. The reliability for a parallel structure is found by [7], [8], [11], [17]

$$
\begin{aligned}
h &= 1 - (q_1 * q_2 \ldots * q_n) \\
h &= 1 - \prod_{i=1}^{n} q_i \\
h &= 1 - \prod_{i=1}^{n} (1 - p_i)
\end{aligned}
\tag{3.26}
$$

A proof of equation 3.26 can be found in [11, Example 9.3, pp. 375-376]. This is done using the theory about extreme variables presented in Section 3.4.

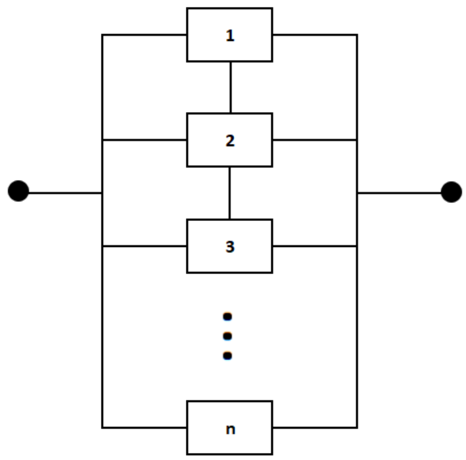


Figure 3.5: Parallel structure in a RBD. Inspired by [8, Fig. 3.2, p. 89].

3.8.3 K-out-of-n Structures

The k-out-of-n structure is a structure where both series and parallel is present. The system functions if and only if k-out-of-n structures functions. An example of a k-out-of-n structure is shown in figure Figure 3.6. This system functions if and only if at least one of the following is true:

- Component a and b functions;
- Component b and c functions;
- Component a and c functions;

- Component a, b and c functions.

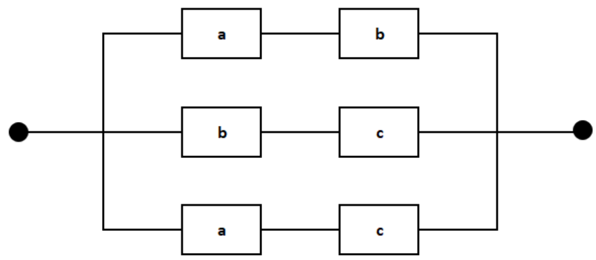


Figure 3.6: an example of a k-out-of-n structure in a RBD. This system functions if and only if 2-out-of-3 components function. Inspired by [8, Fig. 3.3, p. 90].

As it happens, in the case where all components have the same probability of functioning, one can calculate the system using the binomial distribution. The equation for calculating the reliability for a k-out-of-n structure is shown in equation 3.27. Notice the summation sign. In the situation as shown in Figure 3.6, one must summarise in order to include the probability that all components are functioning (n-out-of-n) [7], [8].

$$h = \sum_{i=k}^{n} \binom{n}{i} p^x (1-p)^{n-x} \tag{3.27}$$

3.8.4 Time-dependent Components

In many cases, whether or not a component functions at a given time is dependent of time. In such cases, a CDF is used. As shown in Section 3.2, a CDF for a continuous random variable is the integral of the PDF. As a result,

$$CDF = P(X \leq x),$$
$$CDF = P(T \leq t).$$

In reliability however, we are interested in finding the probability that the variable T will function after a specific time t. In system reliability calculation, one therefore uses the survival function *R(t)* and failure function *F(t)* in place of p_i and q_i, respectively.

$$F(t) = P(T \leq t),$$
$$R(t) = 1 - F(t),$$
$$R(t) = P(T > t).$$

3.8.5 Minimal Cut Sets/Vectors

Minimal Cut sets

For calculation of complex systems, it is often useful to identify the systems minimal cut sets. A minimal cut set is the minimum set of component that causes system failure. For example, in the series structure in Figure 3.4, the minimal cut sets would be 1, 2. Only component 1 or 2 would be necessary to cause system failure. For the parallel structure, the minimal cut set would be 1, 2,...,n. Every component would need to fail in order for the system to fail [7], [8], [17].

Minimal Cut Vectors

Minimal cut vectors are used when the components have multiples states, and changes based on the delivery output level c. Given that all other components are working at maximum possible output except for component i, the minimal cut vectors is the set of outputs that causes the system to deliver an output one level lover than a benchmark level c [7], [8], [17].

The usage of minimal cut vector is best understood through an example. Imagine a series structure system as shown in Figure 3.4 consisting of three components. The delivery level c at point b is 600 liters of crude oil per day. Component A, B and C all have four states each. They can process 0, 400, 500 or 800 liters of oil. The minimal cut vectors in this case are listed below, with the top entry showing the format:

- {state component A, state component B, state component C, state component D};
- {800, 800, 800, 500};
- {800, 800, 500, 800};
- {800, 500, 800, 800};
- {500, 800, 800, 800}.

For a series structure, the system output can not be higher than the lowest processing state in the system. For a parallel structure however, the opposite holds. To find the output delivery from a parallel structure, one must add the delivery states for each component to find the total delivery output.

3.8.6 Methods for Calculating System Reliablity for Complex Structures

This subsection discusses two different methods used to calculate system reliability represented by complex RBDs.

Factoring Algorithm

The factoring algorithm (or pivot-decomposition method) is used to calculate system reliability, by pivoting on one arbitrary component i. By doing so, the system transforms from one complex system to simpler series and parallel structures. The factoring algorithm is given by [7], [8]

$$h(p) = p_i \, h(1_i, p) + (1 - p_i) \, h(0_i, p). \tag{3.28}$$

Where:

- *h(p)*: System reliability;
- p_i: Reliability of component i;
- $h(1_i, p)$: System reliability *given* that component i is in its best state (i.e. reliability=1);
- $(1 - p_i) = q_i$: Unreliability of component i;
- $h(0_i, p)$: System reliability *given* that component i is in its worst state (i.e. reliability=1).

As one may see from equation 3.28 and the explanation, the factoring algorithm is closely related to the chain rule presented in equation 3.3 [7], [8].

The Inclusion-Exclusion Method

Another method of calculating the system reliability is by the Inclusion-Exclusion method. Here, one uses the minimal cut sets and assumes that the components are independent. This method is best described by an example for binary components. However, this method is also applicable for components with multiple states [7], [8], [17].

Consider the system shown in Figure 3.7. One identifies the minimal cut sets as {1, 2, 3} in subsystem A, {4, 5} in subsystem B and {6, 7, 8} in subsystem C. Malfunction for all components in one of these sets causes system failure.

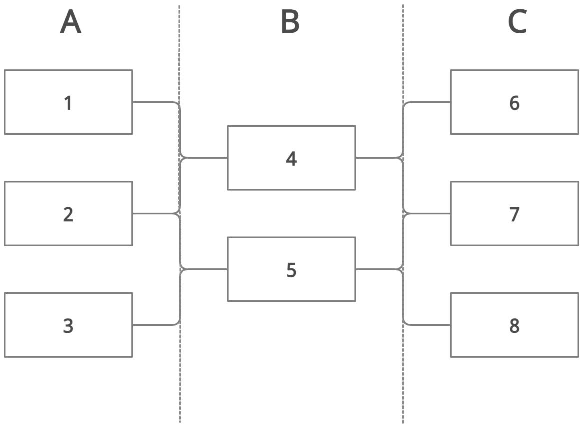


Figure 3.7: Example structure for explaining the inclusion-exclusion method

The unreliability of subsystem A is therefore equal to $q_1 * q_2 * q_3$. The unreliability of subsystem B is $q_4 * q_5$ and the unreliability of system C is $q_6 * q_7 * q_8$. The relation between component A, B and C can be illustrated using a Venn diagram as shown in Figure 3.8. A Venn diagram can be used as the subsystems are mutually non-exclusive.

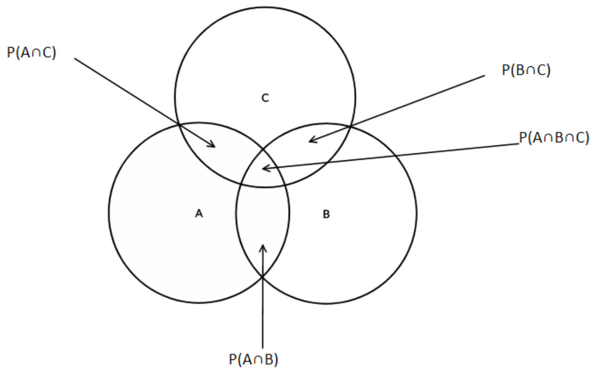


Figure 3.8: Relation between subsystem A, B and C

Figure 3.8 illustrates that when summarising the unreliability of subsystem A, B and C, one also includes the elements in $P(A \cap B), P(A \cap C)$ and $P(B \cap C)$ two times, and P($A \cap B \cap C$) three times in the calculation. One must therefore subtract $P(A \cap B)$, $P(A \cap C)$ and $P(B \cap C)$. Then, when these are subtracted, one must re-add $(A \cap B \cap C)$ again, as this was removed in the process.

This is done by first multiplying the minimal cut sets with each other in pairs (i.e. $A * B$, $A * C$, and $B * C$), summarizing these and subtracting this from $P(A)+P(B)+P(C)$. Then, $P(A) * P(B) * P(C)$ is added at last. The equation for the system unreliability now looks like

$$
\begin{aligned}
g = (q_1 * q_2 * q_3 + q_4 * q_5 + q_6 * q_7 * q_8) \\
- (q_1 * q_2 * q_3 * q_4 * q_5 + q_1 * q_2 * q_3 * q_6 * q_7 * q_8 + q_4 * q_5 * q_6 * q_7 * q_8) \\
+ (q_1 * q_2 * q_3 * q_4 * q_5)
\end{aligned}
$$

Then, by using the complementary rule ($h = 1 - g$), one will get an equation for the system reliability.

3.8.7 Importance Measures

When configuring and changing the design of a system to improve the reliability, a good starting point is to find the most critical components (i.e. the biggest contributors to system unreliability). To identify the most critical components, one can look at which has the highest improvement potential, or which is the most important for the system's functionality [7], [8].

Improvement Potential

The improvement potential is an importance measure used to find the components' potential for improvement in terms of reliability. It is calculated as shown in equation 3.29 [7], [8].

$$I_i^A = h_i - h. \tag{3.29}$$

Where:

- I_i^A: Improvement potential for component i;
- h_i: The system reliability when component i is in the best state (reliability=1);
- h: The system reliability.

Birmbaun's Measure

Birmbaun's measure is a measure used to find the components' importance in the system [7], [8]. To define it, one should first recall the factoring algorithm (from equation 3.28) used to calculate system reliability for complex systems.

$$h(p) = p_i\, h(1_i, p) + (1 - p_i)\, h(0_i, p). \tag{3.30}$$

Birmbaun's measure (I_i^B) can then be found by partial differentiate *h(p)* with respect to p_i [7], [8].

$$
\begin{aligned}
I_i^B &= \frac{\partial h}{\partial p_i}, \\
I_i^B &= h(1_i, p) + h(0_i, p).
\end{aligned}
\tag{3.31}
$$

3.8.8 Availability

The reliability theory discussed so far is related to non-repairable systems. This subsection discusses the case when the system or components are repairable.

Availability is defined as the portion of time a component i is functioning in the long run. When one talks about repairable systems, one generally uses the terms availability and unavailability (A and $\overline{A}$ in place of h and g, respectively). Availability is also used for calculating on component level. The method for calculating system availability is the same as for calculating system reliability. The (limiting) availability for component i is defined as [7], [8]

$$A_i = \frac{MTTF_i}{MTTF_i + MTTR_i}. \tag{3.32}$$

Where:

- $MTTF_i$: Mean Time to Failure. The expected portion of time component i is in operational state;
- $MTTR_i$: Mean Time to Repair. The expected portion of time component i is being repaired;
- A_i Availability for component i.

Notice that one uses expected values to estimate the limiting availability. More specifically, the MTTF is the inverse of the failure rate ($\frac{1}{\lambda}$) and the MTTR is the inverse of the repair rate ($\frac{1}{\mu}$). Inserting these definitions in equation 3.32 yield the following.

$$A_i = \frac{\mu_i}{\mu_i + \lambda_i}$$

A proof that equation 3.32 holds can be found in [18, Example 5.4(A), pp. 242-243], [6, Example 8.5, pp. 313-314] and [17, Example 6.11, pp. 371-372]. Here, the Kolmogorov forward equation is used to show that this equation holds.

Further, one define the limiting unavailability for component i ($\overline{A_i}$) as

$$\overline{A_i} = 1 - A_i,$$
$$\overline{A_i} = \frac{MTTR_i}{MTTF_i + MTTR_i}. \tag{3.33}$$

3.8.9 Redundancy

A parallel structure might not only be used to distribute work across multiple similar components, it might also be useful in situations when redundancy is more convenient. With redundancy, a component is installed as a backup and connected to the system when the main component fails [6]. Figure 3.9 illustrates the concept of redundancy in a system.

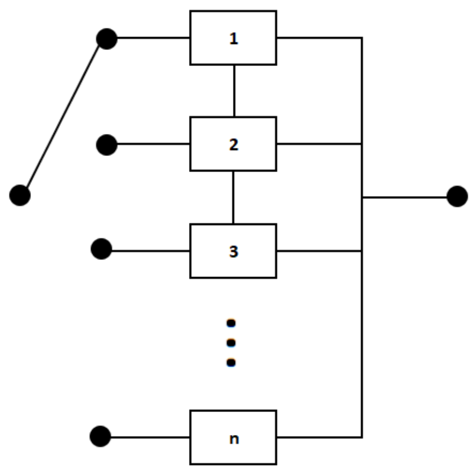


Figure 3.9: A standby system with n components. Inspired by [6, Fig. 4.14, p. 174].

The system MTTF and survival function for such a structure will vary depending on the assumptions made. As an example, in [6, p. 339] presents a case where the switch from main to backup component is perfect, and when there is no repair. Then, the lifetime and the MTTF for the system is simply the sum of each component [6].

$$T = \sum_{i=1}^{n} T_i,$$
$$MTTF = \sum_{i=1}^{n} MTTF_i.$$

Further, if it is also assumed that the components are exponentially distributed, then the survival function of the system is [6]

$$R_S(t) = \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!}\, e^{-\lambda t},$$

which one can recognise as the Poisson distribution.

3.8.10 Maintenance and Replacement

This subsection presents three replacement models discussed in literature that are used for maintenance of components in systems.

Age Replacement model

When using the age replacement model, components are replaced after a specified age u or at failure. This depends on what occurs first; age or failure [6], [20].

The expected replacement cost per unit of time (*K(u)*) is shown in equation 3.34. As one can see, K(u) is defined as expected cost *E(cost)* over the expected time between replacements (*E(T)*) [6], [20].

$$\begin{aligned} K(u) &= \frac{k_c\, F(u) + k_p\, R(u)}{\int_0^u R(t)\, dt}, \\ K(u) &= \frac{E(cost)}{E(T)}. \end{aligned} \tag{3.34}$$

Where:

- *K(u)*: Expected replacement costs per unit (u) of time;
- k_c: Corrective replacement cost;
- k_p: Preventive replacement cost;
- $E(cost) = k_c * P(corrective) + k_p * P(preventive)$;
- $E(cost) = P(T \leq U)$.

Block Replacement Models

In a block replacement model, components are replaced at fixed units of times (u) and when they fail to function. The expected replacement cost per unit of time for a block replacement strategy is shown in equation 3.35 [6], [20].

$$K(u) = \frac{k_c \, E(N(u)) + k_p}{u} \tag{3.35}$$

Here, *E(N(u))* denotes the expected number of failures.

A version of the block replacement model is the minimal repair block replacement model. Like the block replacement model, components are replaced after fixed units of times (u). However, in the minimal repair block replacement model, components are not replaced when they fail between the time blocks. They are instead repaired to the same state as it was just before the failure occurred (minimal repair). The expected replacement cost per unit of time for this model is shown in equation 3.36 [6], [20].

$$\begin{aligned} K(u) &= \frac{k_m \int_0^u \lambda(t)dt + k_p}{u}, \\ K(u) &= \frac{k_m \, E(N(u)) + k_p}{u}. \end{aligned} \tag{3.36}$$

Here, k_m is the cost associated with the minimal repair. Further, a non-homogeneous Poisson process (i.e. variable rate parameter $\lambda(t)$) is used to describe when the failures between the blocks u occur [6], [20].

Unlike the age replacement model, the block replacement model makes it easier to keep an overview of the components, as every component is replaced at fixed time intervals rather than after individual ages. One disadvantage is of course that seemingly new and fully functional components are replaced nevertheless [6], [20].

Condition-Based Maintenance

A final replacement model/policy worth mentioning is the condition-based maintenance policy. Here, maintenance is based on indications or measurements from sensors and other instruments. This policy can be quite useful in cases where the system is not easily viewable or accessible [6].

Chapter 4

Methodologies

Several methodologies exist within the field of risk and reliability. Depending on the context, some are more favourable than others. This chapter discusses some relevant methodologies presented in the literature, how they are used, any relation between them and how they differ from one another. Then, a third section is devoted to presenting the basics behind the Monte Carlo simulation technique, a popular tool used in reliability calculation. A fourth section is then used to present data sources used in this thesis. Finally, a conclusion on the choice of methodologies to be used in the analyses is given in the last section, in addition to a summary of the mentioned data sources used.

Risk, reliability, and probability are closely related. One may therefore argue that some of the theory mentioned in Chapter 3 can be seen as methodologies. More specifically, RBD and Markov chains are definitely methods or tools that could be included in this chapter. However, the theory regarding Markov chains is quite extensive and is therefore discussed in Chapter 3. For the case with RBD, most of the theory regarding reliability is easier to understand and present when the concepts can be related to something visual and concrete like a RBD. It therefore seemed more natural to include RBD in Chapter 3 than in this chapter.

Aven divides various risk analyses into three main categories; simplified risk analysis, standard risk analysis, and model-based risk analysis [21, Tab. 1.1, p. 2]. Relevant types of standard and model-based analyses are discussed separately in the two first sections, respectively.

4.1 Standard Risk Analysis

4.1.1 Structured What-If Technique and Hazard and Operability Studies

Two methods used to identify hazards or threats are Hazard and Operability Study (HAZOP) and Structured What-If Technique (SWIFT). These two methods are closely related because they are both qualitative analysis tools used to systematically uncover potential threats [21].

HAZOP analysis is often utilised in the planning phase and is widely used in the oil and gas industry. The analysis uses guidewords to identify adverse scenarios. After the scenarios are identified, their causes and consequences are studied. Typical guidewords are "as well as", "more of/less of", "no/not", and "other than" [7], [21].

For instance, one might state that "the cleaning pig does not remove wax *as well as* in simulations". Causes might be unrealistic approximations during simulation or that tolerances are not taken into consideration. Consequences might be that the pig gets stuck and blocks the oil flow.

A SWIFT analysis differ from a HAZOP analysis in the way that a SWIFT analysis uses the sentence "what if" throughout the analysis to identify deviations. Another difference is that a SWIFT analysis is somewhat more flexible than a HAZOP analysis [21].

4.1.2 Failure Mode and Effects Analysis

A Failure Mode and Effects Analysis (FMEA) or Failure Mode, Effects and Criticality Analysis (FMECA) is an inductive, quantitative or qualitative analysis tool used for the identification of potential failures of a system. Failure modes are uncovered by studying each component individually. Each possible way the component possibly can fail is noted, in addition to their respective consequences and corrective measures. One might also rank the failure modes after severity and include the probability of occurrence for each event. Doing so would make the analysis quantitative, which might be useful for the implementation of model-based analysis tools [6], [7], [21].

There are some strengths and weaknesses related to a FMEA. First of all, each component is studied individually. Hence, it is not taken into account cases where a combination of components cause system failure, and not just one. On another note, a FMEA gives a good overview of a broad spectre of possible failure modes for the system, even though creating one might be rather time-consuming and extensive [7], [21].

4.2 Model-Based Risk Analyses

Model-based analyses are to a large extent primarily quantitative (i.e. by the use of probabilities) but might as well be used for qualitative analyses. Three different models are described below [21].

4.2.1 Fault Tree Analysis

A Fault Tree Analysis (FTA) is a top-down, deductive analysis tool that can either be quantitative or qualitative. The top event describes a failure mode (identified from a FMEA). Then, one moves downwards and sets up combinations of basic events that can cause the top event, using "AND" or "OR" gates. These combinations are minimal cut set, as failure in one of the sets cause the top event to occur [6], [7], [21].

Figure 4.1 shows both simple fault trees and how they relay to a Reliability Block Diagram (RBD). At the top of Figure 4.1 one sees an "OR" gate and the corresponding RBD. This gate indicated that the top event would occur at least component 1, 2 or 3 fails. Below is a parallel structure and the corresponding fault tree using an "AND" gate. Here, the top event will occur only if component 1, 2 and 3 fails. Finally, at the bottom of Figure 4.1 is a combination of an "AND"gate and an "OR" gate. Here, the top event occurs if component 1 fails *or* if component 2 *and* 3 fails [7], [21].

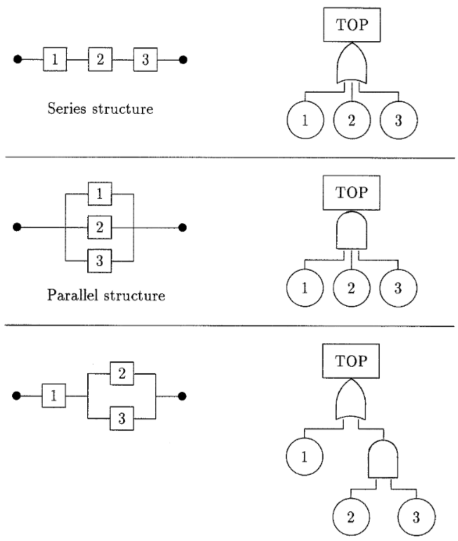


Figure 4.1: An example of a fault tree and how it relates to a RBD. From [21, Fig. 6.5, p. 73].

4.2.2 Event Tree Analysis

While a FTA illustrates what causes an event, an Event Tree Analysis (ETA) illustrates the consequences of the event. Moreover, unlike the FTA, the ETA is an inductive analysis tool. [6], [21]

Figure 4.2 shows an example of an ETA. Starting from an event A, event B represents a consequence if event A occurs, while $\overline{B}$ represents the case when event B does not occur. Further, if event B occurs ($P(B \mid A)$), event C might happen ($P(C \mid B)$) or not ($P(\overline{C} \mid B)$. Finally, Y_0, Y_1 and Y_1 represents the the outcome if event B does not occur ($\overline{B}$), event C does not occur ($\overline{C}$) and if event C occur, respectively [21].

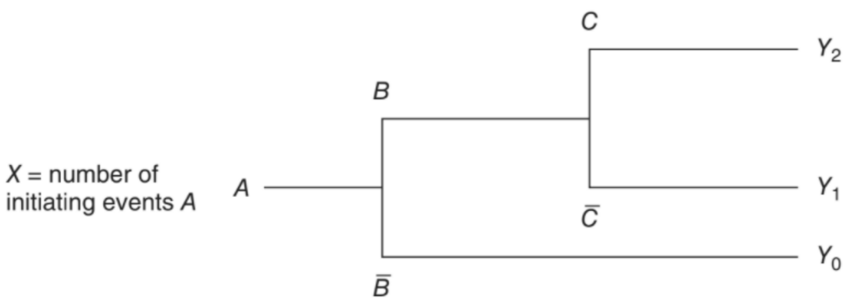


Figure 4.2: An example of an event tree. From [21, Fig. 6.9, p. 78].

For example, the cleaning pig gets stuck in the pipeline (event A). It may or may not block the oil flow, depending on the design of the bypass system (event B or $\overline{B}$). If it blocks the oil flow, pressure increase might (or not) damage the components, sensors or the system (event C or $\overline{C}$). The outcome of these events can be varying costs necessary to solve the problem (Y_0, Y_1 or Y_1).

4.2.3 Bow-tie diagram

A bow-tie diagram can be used to get an overview of the risk picture associated with an event. It is composed of three parts. The left side illustrates what causes an undesirable event to occur, the centre describes the event, and the right side illustrates the consequences of the event [21].

Since a FTA illustrates what causes a failure and an ETA illustrates the consequences of a failure, both analysis can be used to illustrate the risk picture [21]. Further, recall that the top events are identified using a FMECA, HAZOP or SWIFT analysis. A bow-tie diagram is shown in figure 4.3 and is meant to show how the above-mentioned methodologies all relate.

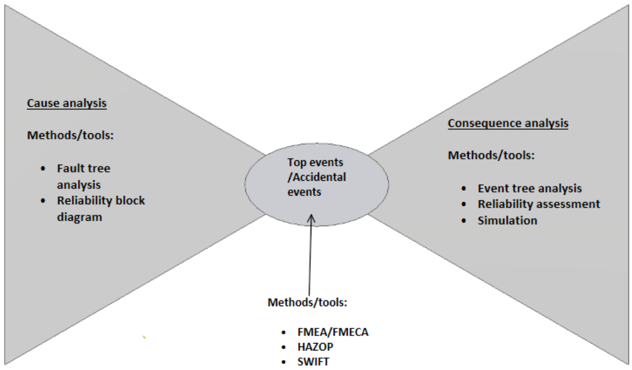


Figure 4.3: A top event can combine a fault and event tree in a bow-tie diagram. Inspired by [21, Figure 1.1, p. 2], [6, Fig. 1.3, p. 9].

4.3 Monte Carlo Simulation

An alternative to analytical methods using calculations is a Monte Carlo simulation. This is a quantitative method where the idea is to run multiple simulations with various probability distributions over time, based on a model of the system [7]. The modelling of the system in question can, for instance, be done with a RBD.

Even though a Monte Carlo simulation generates a large amount of data and can be a fairly good representation of the system, it has some disadvantages. Firstly, if the simulation and the model are not set up correctly, the simulation can do more damage than good. One might end up with a result indicating that the system is far more reliable than it actually is. Secondly, setting up and running the simulations can be time-consuming and expensive if special software is required [7]. When it comes to software, performing a Monte Carlo simulation can also be done in Microsoft Excel or programming languages such as Python.

4.4 Data

Data used in this thesis are primarily obtained from Subsea 7. Some data are also obtained from the OREDA handbook. The OREDA handbook is an extensive collection of failure modes and data for various components used in subsea installations worldwide. One drawback with these data is that the failure rates presented are constant. This is because during a product's useful life phase, its failure rate is often approximately constant. This is an issue because some equipment might have a burn-in phase and a wear-out phase in their lifetime, at which the failure rate is not constant (see Figure 4.4). However, according to the OREDA handbook, a good qualification process as well as maintenance and replacement policy will reduce the burn-in and wear-out phase, respectively. Consequently, the failure rate in the component's useful life phase is approximately constant [22]. The concept of a constant failure rate in the component's useful life phase is also discussed in [8, p. 265], [7, p. 253] and [6, p. 21].

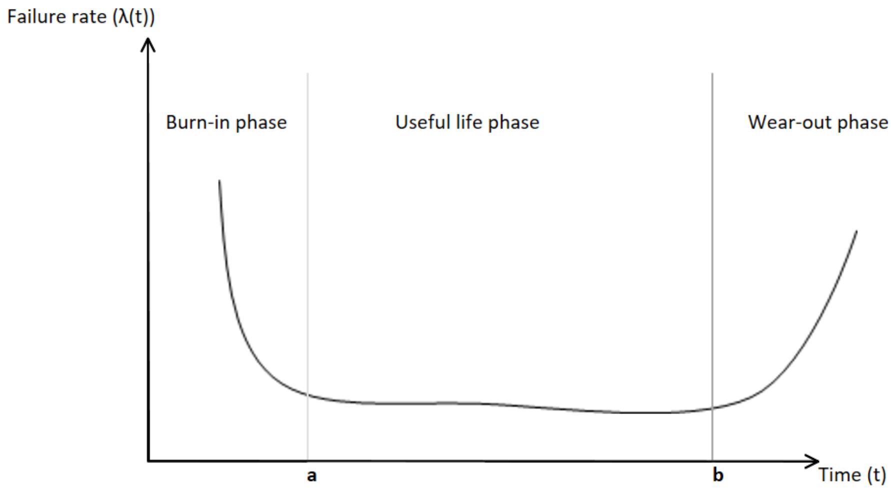


Figure 4.4: Typical development of the failure rate for some components. The failure rate is non-constant in the burn-in and wear-out phase, and assumed to be approximately constant in the useful life phase. Inspired by [22, Fig. 3, p. 24].

4.5 Methodologies Used in the Rest of this Thesis

As it happens, Subsea 7 has already made an extensive FMECA, evaluating a total of 105 different failure modes. It therefore seems unnecessary to create a new FMECA, or even a HAZOP or a SWIFT. Input from the FMECA may be used in various case studies constructed for the analyses.

As Subsea 7 still is in the qualification process, there is currently little to no failure data about the components in the system. In fact, the components have not yet failed during testing. Even though test data from the equipment would be the best data for analysis, some data from the OREDA handbook are somewhat related to some of the components. Therefore, data from OREDA is a good starting point, even though the failure rates presented are constant. Estimates can later be updated when more test data is available. Doing so can, for example, be done by the use of Bayesian statistics.

Furthermore, two RBDs will be made, based on the P&ID shown in Figure 2.7. From these, Markov chains and, in turn, simulations and graphical representations of the differences in system configuration will be made using Python.

One of the objectives of this thesis is to evaluate and compare two different configurations in light of costs, availability and maintainability. In relation to this, qualitative analyses seem at first glance inexpedient. However, as for the second objective of this thesis, the qualitative methodologies might be useful to clarify the arguments being made. In such a case, fault trees will be constructed. The reason for choosing fault trees is mainly its link and transferability from and to RBDs.

Some arguments supporting this decision are that quantitative fault trees and event trees would require a probability calculation on component level. However, the data required to do so does not exist at the time being. Secondly, these methods (RBD, FTA, in addition to Monte Carlo simulation) are suggested in DNVGL's recommended practice for technology qualification, which Subsea 7 uses [2].

Finally, the existing tools and methods presented in 3.8 will be used in discussions for two of the components in the WCS. Trying to apply these tools could unveil whether or not the existing tools fall short. Moreover, Comparing these results to an analysis when Markov chains are used instead could give some indications on the usefulness of Markov chains in reliability analyses.

Chapter 5

RAM Analysis: AMPL

5.1 Introduction to the analysis: Reliability and Importance

5.1.1 System Configurations

The two RBDs considered are shown in Figure 5.1 and 5.2. The case when the cleaning pig is first released from the AMPL and completes one loop is the situation being considered in this chapter. Notice that the cleaning pig itself is not included in the RBDs. A discussion regarding this is given in Chapter 6.

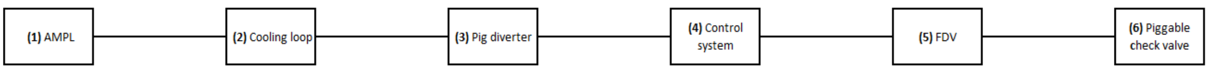


Figure 5.1: RBD of the WCS, based on the PI&D

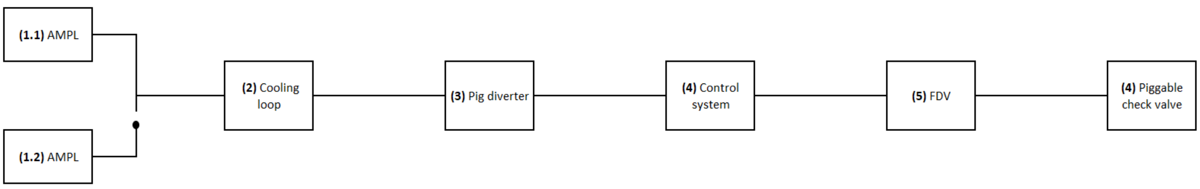


Figure 5.2: Alternative configuration of the WCS, based on the PI&D

The functionality flow related to the two figures are as follows:

1. The cleaning pig is released from the AMPL.
2. The cleaning pig travels with the oil flow through the cooling loop. It then passes the pig diverter (in its initial position), which diverts the pig into its station.
3. The control system sends a signal to the FDV, which turns the FDV. The oil enters from behind the pig in the pig station.

4. Pig re-enters the loop through a check valve.
5. The control system turns the FDV back to its initial position.
6. The control system turns the pig diverter back to its initial position.

Further, in the paper ahead of OTC 2021, Subsea 7 lists the following performance claims for six components [1] in the system and the system as a whole:

- AMPL: MTTF>20 years;
- Cooling loop: MTTF=240 years;
- Pig diverter: MTTF=240 years;
- FDV: MTTF=240 years;
- Cleaning pig: MTTF=2,5 years;
- Cleaning pig: min/max. lifetime=3/6 months;
- Control system: MTTF>240 years;
- System lifetime: 25 years.

These can best be interpreted as design criteria. There are currently (25.05.21) no testing data available to verify if these estimates are reliable. Anyway, the MTTFs give some indication on how Subsea 7's believes of the robustness of the components.

The system will be designed for operating in 25 years on the seabed. The degree to which the components are repairable is both reflected in the performance criteria and described in the OTC paper. All components, except for the AMPL is (to a large extent) non-repairable. Even though some components can be inspected, none can be individually retrieved from the WCS, as they are integrated components in the WCS [1].

5.1.2 Importance measures

From the performance claim/design criteria, one sees that the AMPL is not designed to stay on the seabed longer than the WCS' lifetime without any repair. This makes sense, as the AMPL holds four cleaning pigs in the magazine at the time and must therefore be refilled and looked after roughly every 1-2 years. A consequence of the rather frequent retrieval and refill of cleaning pig is that it does affect the AMPL's importance measures.

Recall the formula for calculation of improvement potential and Birmbaun's importance measure from equation 3.29 and 3.30. From the improvement potential,

$$I_i^A = h_i - h,$$

one sees that the more h_i increases when h remains constant, the higher is the improvement potential.

Also, recall Birmbaun's importance measure,

$$I_i^B = h(1_i, p) + h(0_i, p).$$

Here, $h(1_i, p)$ is the same as h_i in the improvement potential formula. Further, for a series system, $h(0_i, p) = 0$. In short, the AMPL has both the highest improvement potential and is the most important component. This, in addition to the fact that the AMPL is the easiest component to maintain and assumed to be the least available component, serves as strong arguments to why one should consider placing two AMPLs in parallel (as shown in Figure 5.2).

Besides the difference in the number of AMPLs implemented in the systems, the configurations are equal. For that reason, the remaining part of this chapter shall only compare the use of one versus two AMPLs.

5.2 Availability and Maintainability: One versus Two AMPLs

The following notations will be used when analysing availability and maintainability:

- λ_F: Failure rate for the AMPL ($\lambda = \frac{1}{MTTF}$);
- λ_E: Empty rate for the AMPL ($\lambda = \frac{1}{MTTF}$);
- μ_F: Repair rate for theAMPL ($\mu = \frac{1}{MTTR}$).
- μ_E: Refill rate for the AMPL ($\mu = \frac{1}{MTTR}$);

It will also be assumed that a condition-based maintenance strategy is chosen, mainly because the system is inaccessible.

5.2.1 Estimation of the Failure Rate for the AMPL

Recorded data for "pig launcher" from the OREDA handbook is presented in Table 5.1 [22, p. 126].

Table 5.1: Pig launcher data from OREDA

	10^6 hours of operation							
OREDA 2015	No. units	No. of failures	Lower	Mean	Upper	SD	n/t	Mean repair time (hours)
Pig launcher	2	0	0,02	4,51	17,34	6,38	0	0

As one can see, no failures are observed. For components with no recorded failures, OREDA presents the following approach in order to obtain a lower, mean and upper value [22, p. 30]:

1. $\hat{\lambda}_P$: The failure rate estimate. This is obtained from one level up in the taxomony hierarchy.
2. τ: The total time in service (operational or calendar) for the component
3. $\alpha = \frac{1}{2}$ and $\beta = \frac{1}{2\hat{\lambda}_P} + \tau$.
4. The α and β is used to estimate a failure rate: $\hat{\lambda} = \frac{\alpha}{\beta}$.
5. Standard deviation: $SD = \sqrt{\frac{\alpha}{\beta^2}}$.
6. A 90% uncertainty interval: $(\frac{1}{2\beta} z_{0.95,2\alpha}, \frac{1}{2\beta} z_{0.05,2\alpha}) = (\frac{0.002}{\beta}, \frac{1.9}{\beta})$.

When using the data in Table 5.1 and known lifetime of the system (25 years), one will get:

1. $\hat{\lambda}_P = 4,51$
2. τ=25 years
3. $\alpha = \frac{1}{2}$, $\beta = \frac{1}{2*4,51} + 25 = 25,11$
4. $\hat{\lambda} = \frac{0,5}{25,11} = 0,019$
5. $SD = \sqrt{\frac{0,5}{25,11^2}} = 0,028$
6. $(7,965 * 10^{-5}, 0,0756)$

This 90% confidence interval will be used when comparing the different configurations.

5.2.2 Markov Chains for Different Configurations and Maintenance Strategies

In this subsection, transition graphs and transition probability matrices based on Markov chain theory will be constructed for the different system configurations and maintenance strategies. Consequently, it will be assumed that the occurrence of failures occur according to a Poisson process (i.e. the random variables are Poisson distributed). This is a reasonable assumption, as the number of events (failures) in the time interval (here, 25 years) are independent. One may argue that this assumption does not hold in practice, as it may depend on the quality of the repair. However, it is assumed that after the repair, the AMPLs are in a state as good as when they were new. This assumption of perfect repair can be justified by the fact that the AMPL is transported to shore regularly for refilling the magazine with cleaning pigs. When doing so, it makes sense that general inspection is done simultaneously. By assuming a Poisson process, the times between the failures are exponentially distributed. Another consequence of assuming that failures follow a Poisson process is that two events can not occur simultaneously. In a Poisson process, the probability of two failures in a small time interval Δt is zero [6], [10]. Both in [7, Appendix D.4, pp. 280-283] and [6, Example 8.2, pp. 307-309], it is mentioned examples of Markov chains where two components in parallel fail simultaneously, but because of a common-cause failure. This type of failure has its own failure rate and can therefore be thought of as one single event.

When it comes to repairing, it is assumed that there is capacity enough to repair and refill two AMPLs and that these repair and refill crew is adjusted after the situation. As a consequence, the repair and refill rate for one AMPL is the same as for two AMPLs (i.e. the queue notations are M/M/1 and M/M/2 for one single AMPL and a parallel structure, respectively). As a final note, failure is here defined as not being able to function on demand, for instance, due to a malfunction, external forces or that the cleaning pig gets stuck inside the AMPL when it is intended to be released into the oil pipeline.

One AMPL

Figure 5.3 shows the transition graph describing a configuration with only one AMPL. A description of each of the three states is given in Table 5.2.

Table 5.2: Meaning behind the three states used in the transition graph for the Markov process in the case with one AMPL in use

State No.	AMPL
2	Function
1	Empty
0	Fail

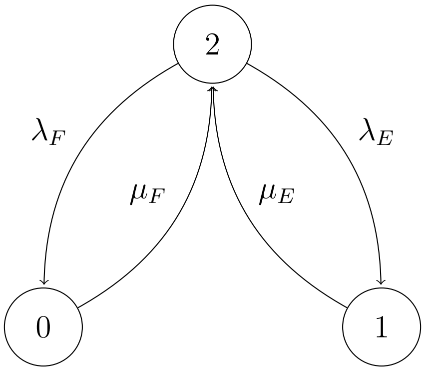


Figure 5.3: Transition graph for a singleAMPL

Figure 5.3 should be interpreted as follows:

- At *t=0*, the AMPL is in state 2.
- At a time *t*, the AMPL is emptied (transition from state 2->1 with empty rate λ_E).
- At a time *t*, the AMPL malfunctions (transition from state 2->0 with failure rate λ_F).
- At a time *t*, the AMPL is repaired and reinstalled (transition from state 0->2 with repair rate μ_F).
- At a time *t*, the AMPL is refilled and reinstalled (transition from state 1->2 with refill rate μ_E).

From the transition graph in Figure 5.3, one can obtain the following transition matrix.

$$P_s = \begin{pmatrix} -\mu_F & 0 & \mu_F \\ 0 & -\mu_E & \mu_E \\ \lambda_F & \lambda_E & -(\lambda_F + \lambda_E) \end{pmatrix} \tag{5.1}$$

Two AMPLs in parallel

In this subsection, the case when two identical and independent AMPLs are connected in parallel is studied. In one of the situations, it is assumed that repair is done in the same order as they fail (i.e. repaired once one has failed). In the other paragraph, it is assumed that repair is done only when both AMPLs have failed. Table 5.3 describes the meaning behind the different states being used in the transition graphs. The interpretation of the descriptions in each state is that "when one AMPL is ..., the other is /... For example, for state 5, when one AMPL is in operation, the other is standby. The switch from an empty/broken AMPL to a stationary is considered perfect (i.e. no delay). Also, it is assumed that an AMPL can malfunction even though it is in standby mode. This can, for instance, happen due to external subsea forces.

Table 5.3: Meaning behind the five states in the transition graph for a parallel structure

State No.	AMPL 1	AMPL 2
5	Operation/Standby	Standby/Operation
4	Empty/Operation	Operation/Empty
3	Failed/Operation	Operation/Failed
2	Empty/Failed	Failed/Empty
1	Empty/Empty	Empty/Empty
0	Failed/Failed	Failed/Failed

The interpretation of the transition graphs presented below is as follows:

- At time t=0, the system is in state 5.
- From state 5, one of the AMPLs have failed (Transition from state 5→3 with failure rate $2\lambda_F$).
- From state 5, the initially operating AMPL is emptied (Transition from state 5→4 with the empty rate λ_E).
- The standby AMPL that was introduced because the other was empty is now also empty (Transition from state 4→1 with the empty rate λ_E).
- The standby AMPL that was introduced because the other was empty has failed to operate (Transition from state 4→2 with failure rate λ_F).
- The standby AMPL that was introduced because the other failed has failed to operate (Transition from state 3→0 with failure rate λ_F).
- The standby AMPL that was introduced because the other failed is now empty (Transition from state 3→2 with failure rate λ_E).
- When both AMPLs have failed, they are sent to repair (Transition from state 0→5 with repair rate μ_F).
- When both AMPLs are empty, they are sent to refill (Transition from state 1→5 with refill rate μ_E).

- When one AMPL is empty, and one has failed, they are sent to refill and repair (Transition from state 2→5 with repair rate μ_F).

- **Only for one situation:** When one AMPL fails, it is sent to repair (Transition from state 3→5 with repair rate μ_F).

Assuming Reparation or Refill Process Starts Once One Has Failed or Both Are Emptied

Figure 5.4 shows a transition graph for the situation where the AMPL is sent to repair once one has failed, once both has failed or when both are empty.

It might seem strange that even though the maintenance strategy states that repair should be done once one of two has failed, there are still rates from 3->2 and 3->0. However, the repair rate μ_F is much greater than both λ_E and λ_F (see Subsection 5.2.3). As a result, the path from 3->5 will be used the most in the long run, making this graph a good representation of the maintenance policy "repair once one has failed". Blue lines represent repair or refill, while red lines represent a failure or that the AMPLs are emptied.

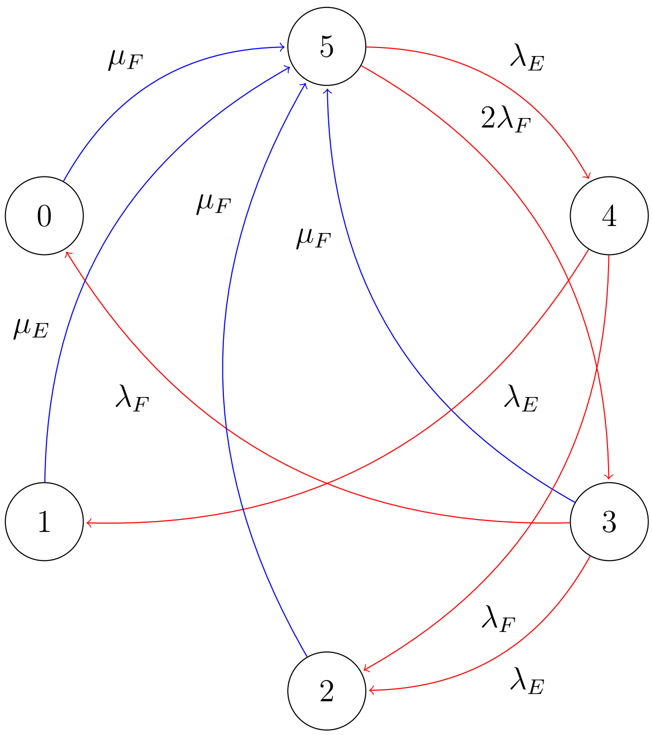


Figure 5.4: Transition graph for a parallel structure where repair is done once one has failed or both are empty

From Figure 5.4 one can obtain the following transition matrix.

$$\begin{pmatrix} -\mu_F & 0 & 0 & 0 & 0 & \mu_F \\ 0 & -\mu_E & 0 & 0 & 0 & \mu_E \\ 0 & 0 & -\mu_F & 0 & 0 & \mu_F \\ \lambda_F & 0 & \lambda_E & -(\lambda_E+\lambda_F+\mu_F) & 0 & \mu_F \\ 0 & \lambda_E & \lambda_F & 0 & -(\lambda_E+\lambda_F) & 0 \\ 0 & 0 & 0 & 2\lambda_F & \lambda_E & -(\lambda_E+2\lambda_F) \end{pmatrix}. \tag{5.2}$$

Assuming Reparation or Refill Process Starts Once Both Has Failed or Are Emptied

The other maintenance policy being considered is shown in the transition graph in Figure 5.5. Here, the AMPLs are only sent to refill or repair once both of them has failed. Just as with the other maintenance policy, blue lines represent refill or repair, and red lines represent a failure or that the AMPLs are emptied.

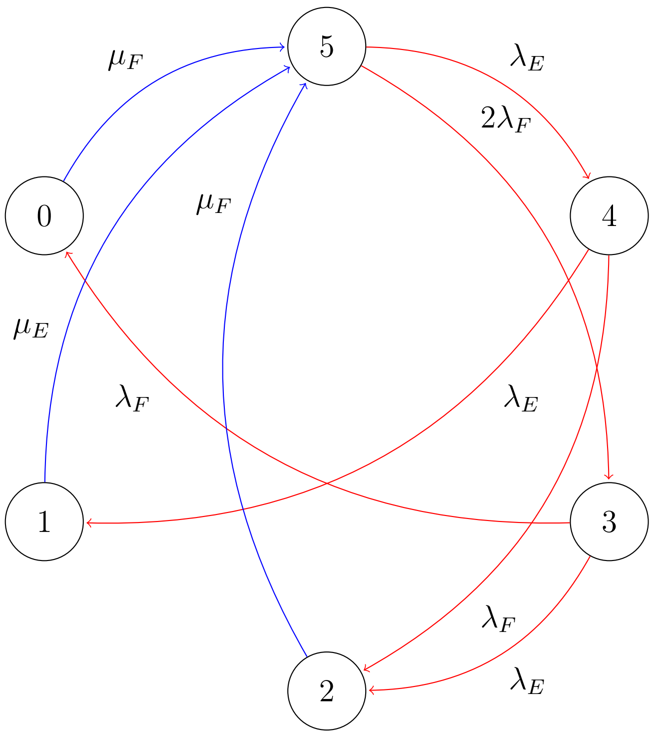


Figure 5.5: Transition graph for a parallel structure where repair is done once both has failed or are empty

From Figure 5.5 one can obtain the following transition matrix.

$$\begin{pmatrix} -\mu_F & 0 & 0 & 0 & 0 & \mu_F \\ 0 & -\mu_E & 0 & 0 & 0 & \mu_E \\ 0 & 0 & -\mu_F & 0 & 0 & \mu_F \\ \lambda_F & 0 & \lambda_E & -(\lambda_E+\lambda_F) & 0 & 0 \\ 0 & \lambda_E & \lambda_F & 0 & -(\lambda_E+\lambda_F) & 0 \\ 0 & 0 & 0 & 2\lambda_F & \lambda_E & -(\lambda_E+2\lambda_F) \end{pmatrix} \tag{5.3}$$

5.2.3 Comparison of the Different Configurations

The different configurations and maintenance strategies will be compared using 2D and 3D plots. The functions are the steady-state probabilities found from solving the equation sets from each Markov chain. These sets of equations are identified using the transition probability matrices and transition graphs and are presented below.

In the case where only one AMPL is used, the steady-state probability of being in state 2 (i.e. π_2) is identified. As for the parallel structure, the system will function as long as the system is in state π_5, π_4 or π_3. These probabilities are added to find the probability of being in either of the states. The equation sets are solved with respect to the desirable steady-state probabilities using a script made in Python (see Appendix A.1) The scripts for the 2D and 3D plot can be found in Appendix A.2.

The set of equations for the configuration with one AMPL:

$$\pi_2\lambda_F = \pi_0\mu_F$$

$$\pi_2\lambda_E = \pi_1\mu_E$$

$$\pi_0\mu_F + \pi_1\mu_E = \pi_2\lambda_F + \pi_2\lambda_E$$

$$\pi_0 + \pi_1 + \pi_2 = 1$$

The set of equations for the configuration with two AMPLs, when the repair is done as soon as one fails or both are emptied:

$$\pi_3\lambda_F = \pi_0\mu_F$$
$$\pi_4\lambda_E = \pi_1\mu_E$$
$$\pi_4\lambda_F + \pi_3\lambda_E = \pi_2\mu_F$$
$$\pi_5 2\lambda_F = \pi_3\mu_F + \pi_3\lambda_F + \pi_3\lambda_E$$
$$\pi_5\lambda_E = \pi_4\lambda_E + \pi_4\lambda_F$$
$$\pi_3\mu_F + \pi_2\mu_F + \pi_1\mu_E + \pi_0\mu_F = \pi_5\lambda_E + \pi_5 2\lambda_F$$
$$\pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 = 1$$

The set of equations for the configuration with two AMPLs, when the repair is done when both fails or are emptied:

$$\pi_3\lambda_F = \pi_0\mu_F$$
$$\pi_4\lambda_E = \pi_1\mu_E$$
$$\pi_4\lambda_F + \pi_3\lambda_E = \pi_2\mu_F$$
$$\pi_5 2\lambda_F = \pi_3\lambda_F + \pi_3\lambda_E$$
$$\pi_5\lambda_E = \pi_4\lambda_E + \pi_4\lambda_F$$
$$\pi_2\mu_F + \pi_1\mu_E + \pi_0\mu_F = \pi_5\lambda_E + \pi_5 2\lambda_F$$
$$\pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 = 1$$

When making a comparison, the following assumptions have been made:

1. Steady-state probabilities exist (the chains are both irreducible and positive recurrent) and are used when comparing. Using steady-state probabilities is reasonable because of the long-planned time of operation (t=25 years). For example, in [18, Example 5.4(A), pp. 242-243] and [6, Example 8.5, pp. 313-314], the Kolmogorov forward equation are utilised on a simple structure. When the differential equations are solved, one will get an expression which is a function of time *t*. When inserting, for instance, t=25, one will see that the expression simplifies, as parts of the expression now are ≈ 0. One is then left with the steady-state probabilities.

2. The repair rate μ_F is difficult to predict, as it depends on many factors:

- Distance from the operation site to the service location;
- Complexity of the repair;
- Weather conditions in the area (may complicate the retrieval of the AMPL);
- Time available until the cleaning pig currently in operation is considered worn out, and a new pig must be put in operation.

Because of this, the steady-state probabilities will be plotted as a function of the repair rate $\pi_i(\mu_F)$, with the repair rate ranging from three months ($\mu_F = 4$, MTTR=$\frac{1}{4}$) to one week ($\mu_F = 52$, MTTR=$\frac{1}{52}$)

3. The refill rate μ_E is also challenging to predict. It makes sense that refill goes faster than repair, but it is unknown how much faster. For that reason, the refill rate is set equal to the repair rate, so that $\mu_E = \mu_F = \mu$.

4. The empty rate λ_E is set to be equal to 0,8. It is initially five cleaning pigs in the AMPL, each with a minimum lifetime of three months. As a consequence, the AMPL will need refill after minimum 5*3=15 months=1,25 years (MTTF=1,25, $\lambda_E = \frac{1}{1,25} = 0,8$).

5. A total of three 2D plots will be made as a function of the repair rate:

   - One plot is based on the upper limit in the 90% confidence interval for the failure rate λ (here denoted λ_F). This upper limit was found to be 0,0756 (MTTF=13,2 years). For the 2D-plot, this value is modified to 0,1 (MTTF=10 years).
   - A second plot will base the λ_F on the performance claim presented in the OTC paper. Here, AMPL is claimed to have a MTTF>20 years. The second plot will therefore use the limit of 20 years (λ_F=0,05).
   - The third plot will use the estimated failure rate, based on the data from the OREDA-handbook. This estimate was found to be equal to 0,019 (MTTF=52,63 years).

6. In reality, the failure rate may vary from field to field, as it will depend on factors like the conditions where it is installed and the quality of the build, to mention a few. Therefore, a contour plot is made by varying the repair rate μ from three months to one week and varying failure rate λ_F. The failure rate values will be based on the 90% confidence interval, although somewhat modified. Firstly, the upper limit is adjusted from 0,0756 to 0,1. Secondly, it is unrealistic to assume that the failure rate will ever have a value equal to the lower limit of the confidence

interval ($\lambda = 7,965 * 10^{-5}$, MTTF=12554,9 years). This lower limit is therefore adjusted for the sake of plotting from to $\lambda_F = 5 * 10^{-3}$, (MTTF=200 years).

The above-mentioned plos are also done with the empty rate λ_E equal to 0,4 (MTTF= 2,5 years; cleaning pig lifetime is set to six months instead of three). The reason for doing so is to get a picture of the role the empty rate has on the AMPLs availability. Intuitively, one might think that this would increase the availability. Although this might be true, increasing the time between refill and inspection also increases the possibility of failure while in operation.

Finally, the 2D plots will be compared to the plots when the limiting availability formula ($\frac{MTTF}{MTTF+MTTR}$) is used. These plots are meant as a discussion the usefulness of the Markov chain. As a consequence, these comparing plots will be presented in Chapter 8 instead of Chapter 7.

Chapter 6

Discussion About and Analysis of a Non-Stationary Component

When discussing system reliability in literature (such as in [6]–[8], [21]), one generally discuss cases where the components represented in the RBD are stationary components. For instance, a pump, valve, or vessel that performs work without moving. However, in the case with the WCS, the cleaning pig is continuously moving around in the system and passing through other components.

In the first section, standard and model-based risk analysis methods are applied on the cleaning pig. Then, a discussion is presented concerning how these reliability concepts in some areas might fall short in an analysis of a system with a non-stationary component. In the second section, two case studies are presented. These studies are intended to show if and how Markov chain theory might be used as tools to analyse or describe the wear of the cleaning pig.

The two sections seem, at first glance, not particularly related, as the first is more of a discussion and the second is more analytical. However, the idea behind the second section is to present an alternative method to analyse a typical failure mode related to the cleaning pig, identified during qualification testing and listed in Subsea 7's FMECA. The reason for doing so is related to the first section, namely that the existing concepts or tools in some areas might fall short.

6.1 Application of Existing Methods and Concepts

Consider the fault trees presented in Figure 6.1 and 6.2. These are transformed from the RBDs in Figure 5.1 and 5.2, respectively. The only difference is that the cleaning pig is excluded from the RBDs and included in the fault trees. The interpretation is the same for these fault trees as for the RBDs. The top event "system failure" occurs if either the AMPL(s), the piggable check valve, the FDV, the control system, the cooling loop or the cleaning pig fails. An important note one can take from these fault trees is that the order in which the components are meant to operate is not considered.

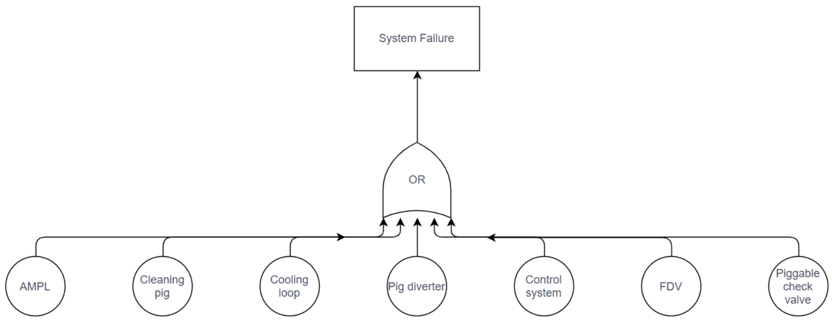


Figure 6.1: Fault tree for the configuration with one AMPL

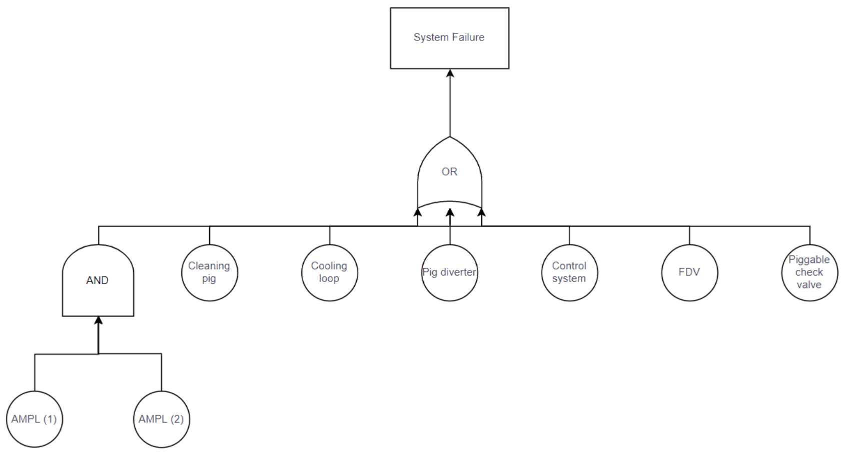


Figure 6.2: Fault tree for the configuration with AMPLs in parallel

This order is easier to present with a RBD. However, some components like the FDV and the pig diverter are put in operation more than once in a cleaning cycle. Consequently, they might have a higher probability of failure the first time than the second time. For instance, it is reasonable to think that the pig diverter will have a higher probability of failure when roataing as thr pig passes versus when it just rotates back to its initial position. One of the reasons for this is that the cleaning pig is present the first time and not being present the second time.

In an ideal setting, a FMEA or FMECA has uncovered all potential failure modes associated with each component. The probability of failure would then be the sum of all the component's failure mode probabilities. This raises a problem regarding a non-stationary component. If a failure mode is identified as "the cleaning pig gets stuck in the cooling loop, pig diverter or piggable check valve", where should the corrective measures be done? Is the cleaning pig's design or its materials the problem, or is the problem the conditions caused by the cooling loop/pig diverter/piggable check valve?

The above-mentioned issues regarding the cleaning pig can be easier understood when it is illustrated. Consider the series structure represented in Figure 6.3, somewhat similar to a RBD but with a few differences. One of them is that the order matters. Imagine that the cleaning pig completes one loop, first released by the AMPL, and re-entered in the loop after passing the piggable check valve. At a time t_1, the cleaning pig is in the cooling loop. Because of the deposed wax at that time, the cleaning pig will have a survival function $R_1(t_1)$. After completing the cooling loop, the pig enters the pig diverter at t_2. Because of the tolerances in the pig diverter and the weld joints used to join the bundled pipe and the diverter, the cleaning pig will have another unique survival function $R_2(t_2)$. Further, imagine that because of the tolerances, surface roughness, or the piggable check valve, some parts of the cleaning pig have fallen off. When the cleaning pig now re-enters the cooling loop, it will have a new, unique survival function $R_3(t_3)$.

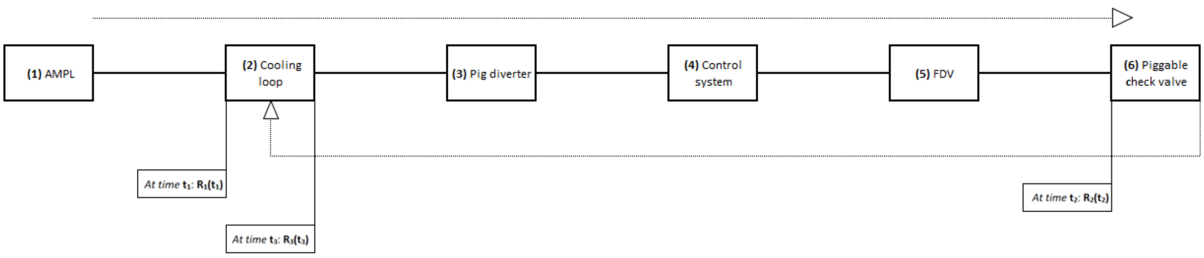


Figure 6.3: The cleaning pig has different survival functions, depending on where it is in the loop

The situations described above indicates that $R_1(t) \neq R_2(t) \neq R_3(t)$. One might think that this is quite dramatic with multiple different survival functions, that the differences between these are negligible and that the cleaning pig could just as well be described using a Weibull distribution, where the failure rate is not constant. This might be true, as the explanation above Figure 6.3 also indicates that failure rate at t_1, t_2 and t_3 are different from one another. However, the possibility of failure might both increase and decrease on the same lap. For instance, the deposed wax at the end of the cooling loop will be much less than in the beginning, indicating that the failure rate here is decreasing over time. Then, the pig enters the pig diverter, a new environment where different factors such as surface roughness, weld joints, and old wax rests can increase the probability of failure.

A solution to bypass this problem could be to exclude the cleaning pig from the reliability calculation and rather increase the failure rate of each component. One strong argument for this solution is that each component will have its own unique survival function. Moreover, the probabilities P(cleaning pig broken down) and P(cleaning pig gets stuck in this component) are also covered by these survival functions. Doing so will include the failure modes associated with the cleaning pig by shifting from "what if the cleaning pig gets stuck/broken in this component" to "what if the component causes the cleaning pig to get stuck/be broken".

Removing the cleaning pig, however, has some drawbacks. Firstly, how much should the failure rate of the other components be increased, and how should one proceed to find that increase? Secondly, when removing the moving component, how does one find its importance or improvement potential?

It was shown in Chapter 5 that the AMPL had the highest improvement potential and importance measure when the cleaning pig was excluded. As a consequence, the AMPL was chosen for further analysis. The reason why the AMPL had the highest improvement potential and importance measure was that it had the lowest performance claim related to MTTF. However, according to the OTC paper, the cleaning pig has an even lower performance claim with a MTTF of 2,5 years. Using the same argument as with the AMPL, one may argue that the cleaning pig is the most important component in the system.

What makes this case particularly interesting is that the cleaning pig's improvement potential is largely dependent on the other components. This is quite intuitive, as the cleaning pig's wear depends on the surface roughness of the components it passes through.

Without performing calculations, assuming that the cleaning pig is the most important component does make sense. Firstly, the cleaning pig's lifetime affects how frequently another component (the AMPL) must be sent for refill, inspection or even repair. Secondly, the whole idea behind the WCS, and the main purpose of the cleaning pig, is to remove deposed wax from the oil pipe.

This deposition of wax can be illustrated using a transition graph. Consider the transition graph shown in Figure 6.4, where each state represents percentage blockage by deposed wax of the cross-sectional area at a random location in an oil pipeline. The wax is deposed (i.e. a state changes) with a rate λ. This rate depends on multiple factors, like wax present in the crude oil, the temperature differences between oil and oil pipe or already deposed wax. When there is n percentage blockage in the oil pipeline, the cleaning pig is sent out to remove the wax so that there is 0% (ideally) deposed wax (represented with repair rate μ). When n is equal 100, there is no oil stream passing through, as the cross-sectional area is completely blocked by wax.

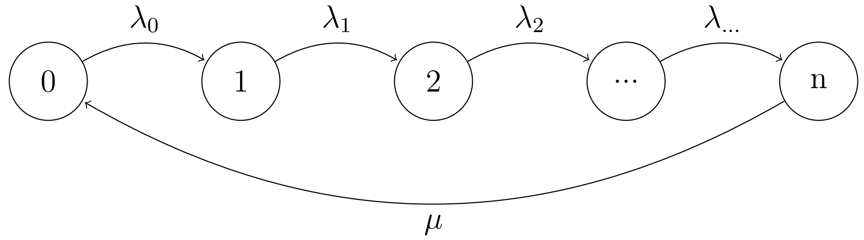


Figure 6.4: Percentage wax blockage over time

6.2 Case Studies and Examples: Analysis of Wear Using Markov Chain Theory

In this section, the issue regarding the wear of the cleaning pig is addressed. The material used on the cleaning pig will be worn over time, with a rate that will depend on multiple factors that can vary from one installation to another. Some common factors causing wear for each installation could typically be:

- Wax content;
- Surface roughness and weld joints;
- Tolerances;
- Crude oil quality and chemical composition in oil mixture.

From the accelerated testing, it was reported that:

- After 1000 completed cleaning laps of 120 meters, 3,2% of its initial weight was lost due to wear and tear;
- After 1457 completed cleaning laps of 120 meters, 4,2% of its initial weight was lost due to wear and tear;

In an actual installation field, the loop will typically be 6-8 km. However, it must be stressed that these data are from an accelerated testing. Although the cooling loop distance is much longer in practice, the internal surface will typically be much smoother, which will reduce the wear of the cleaning pig.

Such a deviation from testing to an actual system makes it even more difficult to estimate a wear rate. Because of this will case studies and examples presented below be thought of as a continuation of the accelerated testing in the qualification test rig. Although not directly applicable to a real-life WCS, the purpose of the case studie is to show how and if Markov chain theory can be used as an analysis tool. The Python scripts created for these case studies are given in Appendix A.3.

6.2.1 Case Study 1: Degradation of the Cleaning Pig

Consider the case where the cleaning pig is worn over time. This process can be described as a continuous-time Markov chain represented by the transition graph in Figure 6.5. Here, each state represents the percentage mass reduced due to wear over time until there is no mass left (i.e. 100% mass reduced in state 100). Such a chain where the repair rates (μ) are zero is commonly referred to as a pure birth process in literature, such as in [17, p. 372]. For each state starting the wear process from time t=0, it can be of interest to graphically illustrate the time it takes before $P_{ij} > P_{ii}$ for each state (i.e. the time it takes before the probability of a mass reduction is higher than the probability of not losing mass).

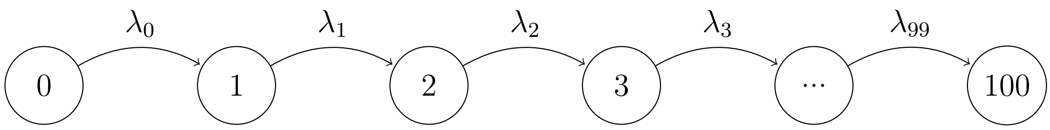


Figure 6.5: Percentage mass reduction of the cleaning pig over time

Results from testing indicated that between 0 and 1000 completed cleaning loops, the mass reduction was 3.2%. After another 457 laps, the mass was reduced by another 1%. Table 6.1 shows the data obtained from testing and calculation that are used in this analysis.

Table 6.1: Data and calculations obtained from the qualification testing

Cleaning pig velocity [m/s]	1,6
Cooling loop length [m]	120
Total completed loops	1457
Total mass reduced	4,2%
Total distance traveled [m]	1457*120=174840
Total time spent cleaning [s]	$\frac{174840}{1,6} = 109275$
Total time spent cleaning [hrs]	$\frac{109275}{60*60} = 30,35$

From this data, one can also calculate the average wear rate (λ_W). $\lambda_W = \frac{4,2\%}{30,35} = \frac{84}{607} \approx 0,1384$. To analyse this process, the first thing that must be done is to use Python and the Kolmogorov forward equation to find the transition probabilities for each of the 100 states. Solving these equations is necessary as the Markov chain is neither positive recurrent nor irreducible, meaning that steady-state probabilities do not exist.

A consequence of the fact that this is a pure birth process is that setting up the differential equations in Python is quite manageable. It is shown in [17, p. 373] that the Kolmogorov forward equation will be on the following form.

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{ii}(t) = -\lambda_i p_{ii}(t),$$

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{ij}(t) = \lambda_{j-1} p_{i,j-1}(t) - \lambda_j p_{ij}(t).$$

Where, j in this case, equals *i+1*. Furthermore, recall the initial conditions.

$$p_{ii}(t = 0) = 1,$$

$$p_{ij}(t = 0) = 0, \ when \ i \neq j.$$

Each of the 100 wear rates will be drawn randomly from an interval ranging from [0.1384-0.1*0.1384, 0.1384+0.1*0.1384]. The reason for drawing the wear rate randomly from an interval rather than assigning all of them to the average value is to include some variation in the wear rates. It is unlikely to be constant. The p_{ii}s and p_{ij}s are solved and then plotted as a function of time t

One can also do this type of analysis with, for instance, 10 states (i.e. mass reduction per 10th%) or with 1000 states (i.e. mass reduction per mille (‰)). When doing so, one must also change the wear rates.

Since the wear process for each of the 100 states here starts from time t=0, there is also another possible interpretation of the results: Assume that one is interested in finding the time it takes until the probability of a state change p_{ij} is larger than p_{ii} for state 0 only. The results (i.e. the plot) then shows relationship between p_{01} and p_{00} with varying wear rate for 100 different tests.

Example for Further Research 1: Degradation after Time *t* Hours Spent Cleaning in Qualification Test Rig

Based on the case study in Subsection 6.2.1, an idea for further analysis could also be to find out the typical wear/mass reduction on the cleaning pig after time t spent cleaning.

To do so, one must in some way create dependencies between the different solved probabilities $P_{ii}(t)$ and $P_{ij}(t)$. In the case study, each $P_{ii}(t)$ and $P_{ij}(t)$ starts simultaneously. However, for this analysis to make sense, conditions saying that for instance $P_{ii}(t_1)$ is defined as long as $P_{ii}(t_1) > P_{ij}(t_1)$, and that the next $P_{ii}(t_2)$ is defined from a time t_2, where $P_{ii}(t_2) > P_{ij}(t_1)$ for $t_2 > t_1$ and $j = i + 1$. After a specified time t, one would then have one or two functions (transient and stationary probability functions), related to a specific state that can be returned.

This process can then be done over multiple iterations, and the state the cleaning pig is in the most (or moving from/to the most) after a time t can be returned and counted. The most visited state at time t would then be the state with the highest relative frequency in a histogram.

An interpretation of the result from such a histogram would be that if the testing continued for time t, and the same test was completed multiple times (i.e. for multiple pigs), how much mass would typically be lost due to wear per cleaning pig.

6.2.2 Case Study 2: Probability of Being in State 8 at Time t

Assume that it has been decided that the cleaning pig shall be replaced once its mass is reduced by 8% (i.e. it has reached state 8). A transition graph illustration this process is shown in figure 6.6.

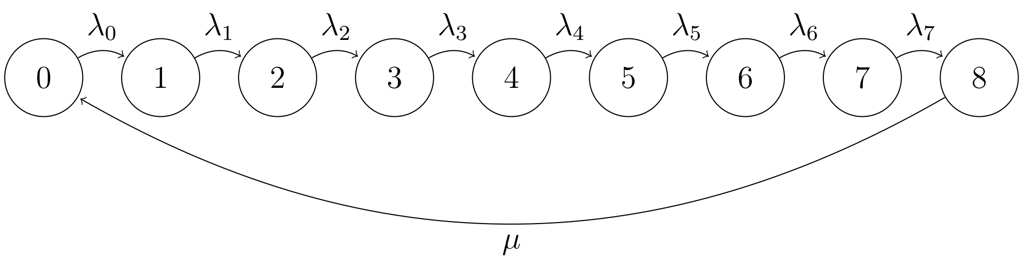


Figure 6.6: Percentage mass reduction and renewal of the cleaning pig after 8% mass reduction

Here, a repair rate μ is introduced to the Markov chain. This repair rate connects state 8 to state 0 and represents the process of releasing a new cleaning pig from the AMPL. One might see that the cleaning pigs stored in the AMPL can be viewed as redundant components. When one fails or is worn out to a specific level, the others are standby and ready to be put in operation (illustrated in figure 6.7).
It could be of interest to find the probability of being in state 8 at a time t. To do so, one can use the fact that the sum of n independent exponentially distributed random variables has a gamma distribution with $\alpha = n$ (here, 8) and β. Recall that β is simply the inverse of the wear rate ($\frac{1}{\lambda_W}$). The assumption that the random variables are independent of each other holds, as this is, in fact, the idea behind the Markov property (i.e. future is independent of the past, only the present). The wear rate used here will be the average

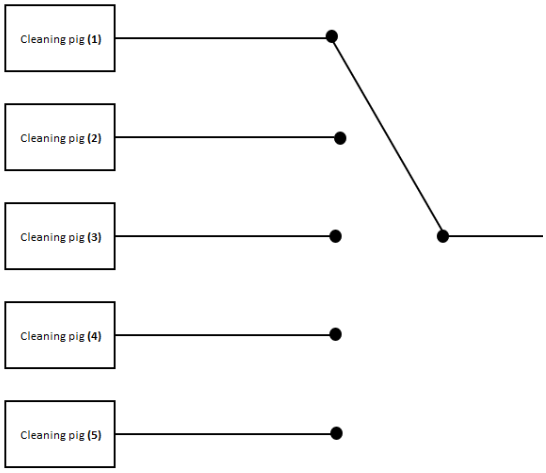


Figure 6.7: The cleaning pigs in the AMPL can be viewed as standby components

wear rate identified in case study 1 in Subsection 6.2.1. The CDF is then plotted as a function of time t.

To look at the usefulness of this, one can also include the plot for the CDF for $\alpha = 4$ and $\beta = \frac{1}{\lambda_W}$. As mentioned, it results from testing showed that the cleaning pig had lost ≈ 4 mass after $\approx$ 30hours. Graphically, the probability after 30 hours should therefore be quite high if this approach works. Using the CDF, the probabilities can be read of from the y-axis.

Example for Further Research 2: Assuming Time t=1000 Hours of Operation in Test Rig, How Many Cleaning Pigs are Typically Needed

This example is based on the case study in Subsection 6.2.2, and acts as a suggestion for deeper analysis of the cleaning pig wear using Markov chain theory. For this analysis as well, it will be assumed that the cleaning pigs are replaced once they reach state 8 (i.e. 8% mass reduction).

The repair rate will typically be very high, as it is based on the time it takes for the AMPL to release a new cleaning pig into the cleaning loop. As a consequence, the cleaning pig will spend little time in state 8. To find how many cleaning pigs that are typically needed over a time period, one can look at how often state 0 is visited during this time period, for instance t=1000 hours.
As with "Example for further research 1", dependencies between the differential equations must be made. When repeating this simulation multiple times, one might, for instance, represents the number of times state 0 is visited using a histogram.

Chapter 7

Results from the Analyses

This chapter presents the results from the analysis of the AMPL in the first section. In the second section, the results from the analysis of the cleaning pig's wear is presented.

7.1 Comparison of AMPL Configurations and Maintenance Strategies

In the two subsections below are graphs showing the availability for different combinations of maintenance strategies and configurations. They are plotted as a function of the unknown repair rate μ and with different values of the failure rate λ_F. Further, in the first subsection, it is assumed that the lifetime of the cleaning pig (which affects the empty rate of the AMPL) is three months. In the second subsection, the lifetime of the cleaning pig is set to six months.

7.1.1 Expected Cleaning Pig Lifetime = three months

In this subsection, a total of four graphs are presented in figure 7.1-7.4. An interpretation of what is shown in the graphs is given at the end of this subsection.

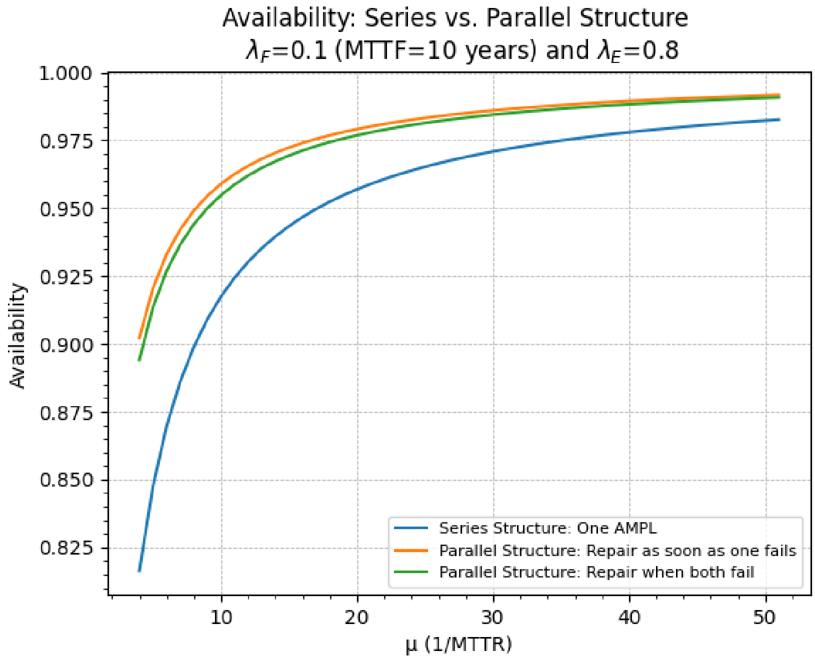


Figure 7.1: Availability as a function of μ for $\lambda_F = 0.1$ and $\lambda_E = 0.8$

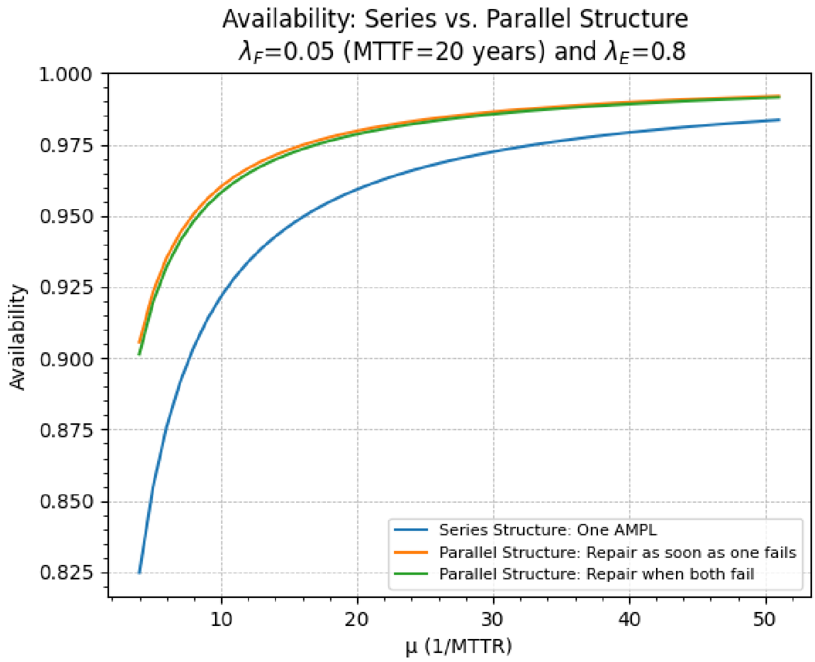


Figure 7.2: Availability as a function of μ for $\lambda_F = 0.05$ and $\lambda_E = 0.8$

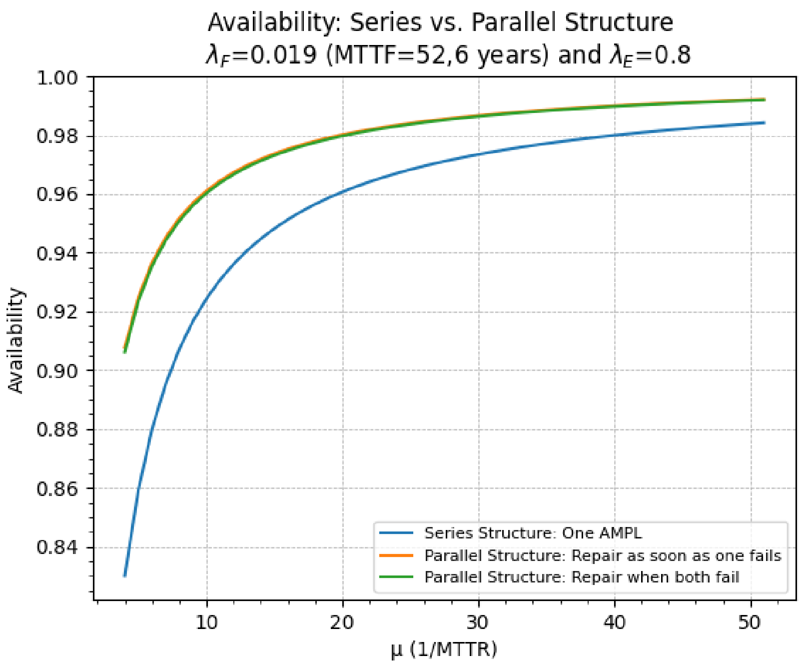


Figure 7.3: Availability as a function of μ for $\lambda_F = 0.019$ and $\lambda_E = 0.8$

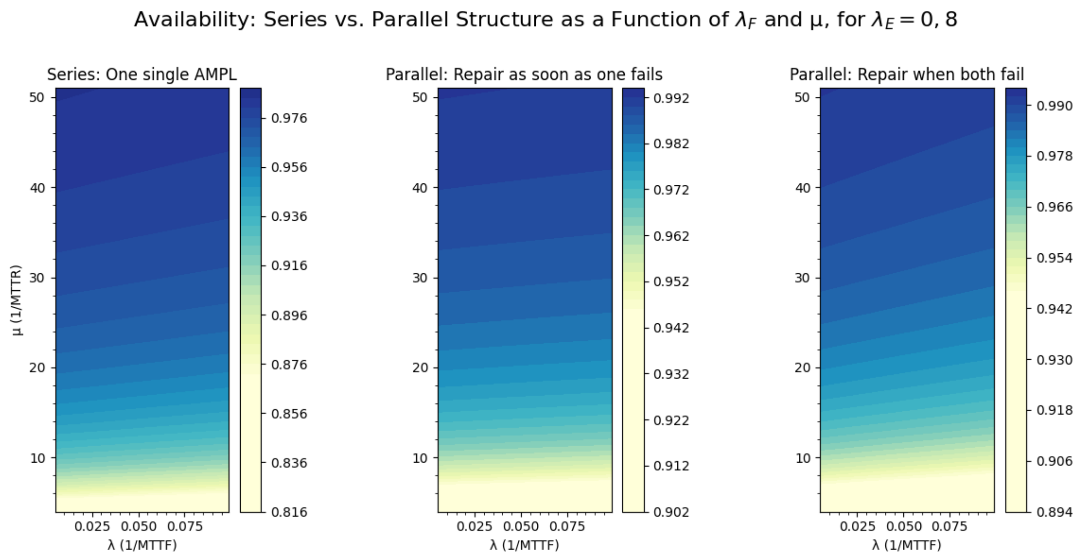


Figure 7.4: Availability as a function of μ and λ_F for $\lambda_E = 0.8$

Not surprisingly, the graphs show that a series structure does not give as high availability as a parallel structure. What is more interesting, however, is the difference in availability. For the longest repair/refill time (i.e. $\mu = 4$, MTTR=three months) an AMPL in a series structure is in the long run 10% less available. Further, as the time it takes to repair/refill decreases towards one week, the difference in availability between the two structures decreases rapidly.

The two graphs for the different parallel structures represent two different maintenance strategies: retrieve and refill once one is empty, or wait until both of the AMPLs are empty. The conclusion that can be drawn by looking at the graphs in figure 7.1-7.3 is that the difference in availability is minimal. A visible difference in availability is only really present in the case where the MTTF is 10 years (i.e. failure rate λ_F=0.1), 10 years shorter than Subsea 7's own performance claim.

Figure 7.4 is meant as a supplement. Instead of showing the difference in availability for point values of the failure rate λ_F, figure 7.4 presents the availabilities for continuous values of μ and λ_F.

7.1.2 Expected Cleaning Pig Lifetime = six months

Figure 7.5-7.8 presents the availability where the expected cleaning pig lifetime is six months. An interpretation of the results from these graphs is given at the end of this subsection.

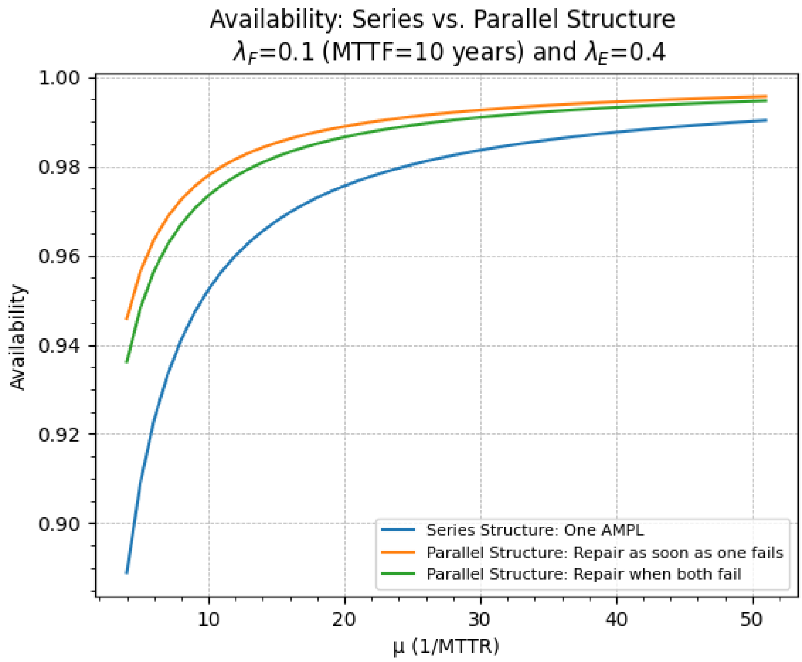


Figure 7.5: Availability as a function of μ for $\lambda_F = 0.1$ and $\lambda_E = 0.4$

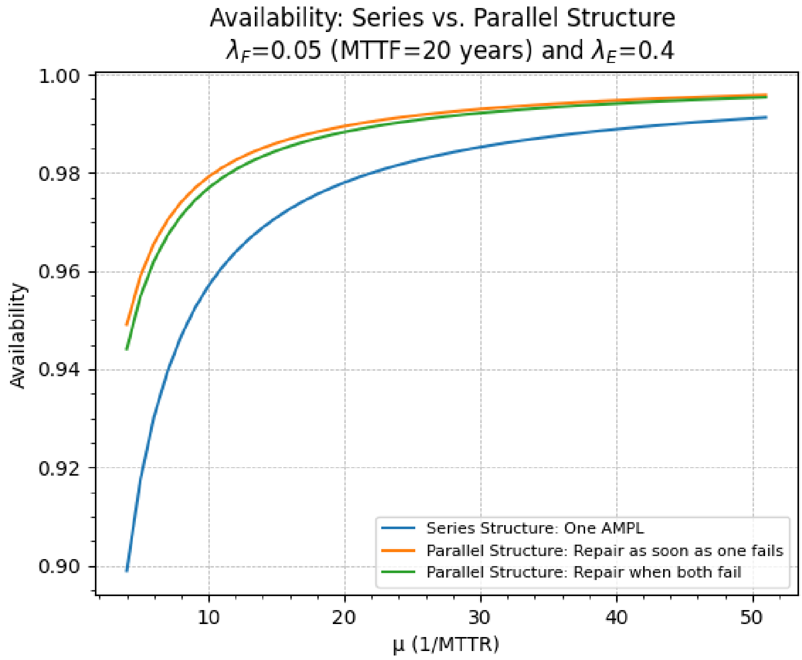


Figure 7.6: Availability as a function of μ for $\lambda_F = 0.05$ and $\lambda_E = 0.4$

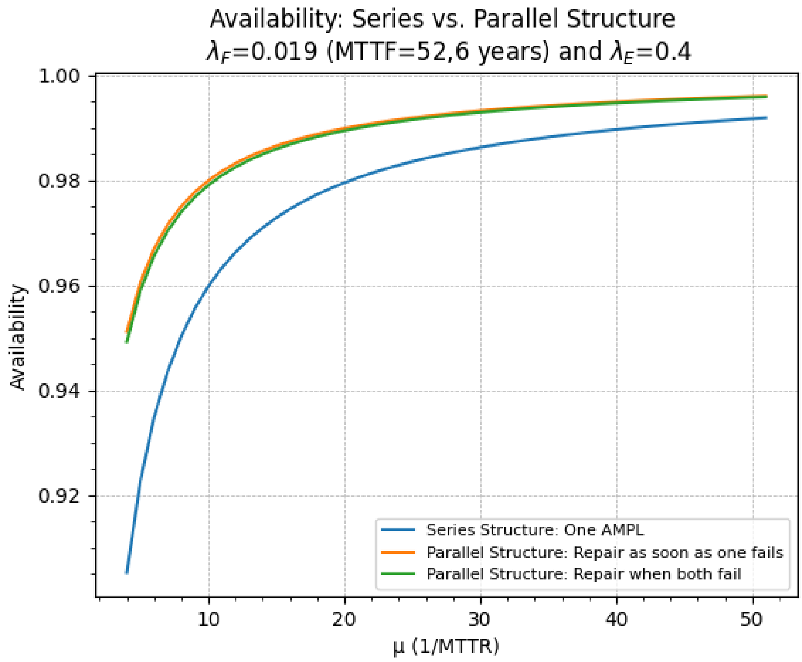


Figure 7.7: Availability as a function of μ for $\lambda_F = 0.019$ and $\lambda_E = 0.4$

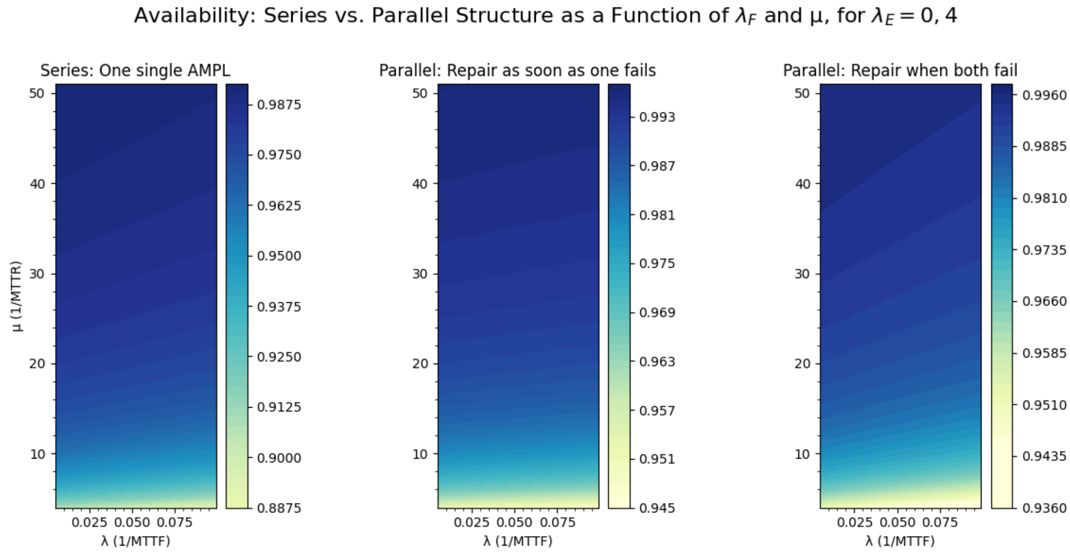


Figure 7.8: Availability as a function of μ and λ_F for $\lambda_E = 0.4$

Primarily two things are worth mentioning that separates the graphs in this subsection from the graphs where $\lambda_E = 0, 8$. Firstly, the availability for both the parallel structure and the series structure is generally higher for all values of the repair/refill rate μ. This is a good indication of what might seem obvious: the AMPLs availability depends on the cleaning pig's lifetime. Secondly, for $\lambda_F = 0, 1$, it seems to be a greater difference in availability between the two different maintenance strategies (for a parallel structure) for the smallest values of μ. However, as λ_F decreases, the difference in availability between the two different maintenance strategies decreases. For $\lambda_F = 0, 019$, the difference in availability is almost negligible.

7.2 Results from Case Studies Built Around the Cleaning Pig

In this section, the results from the case studies concerning the cleaning pig are presented in graphs. Since the case studies presented earlier are on a conceptual level, the graphs presented below open more to a discussion concerning the usefulness of the Markov chain than giving concrete results. As a consequence, this section is rather short.

7.2.1 Results from Case Study 1: Degradation of the Cleaning Pig

Figure 7.9 shows the $p_{ii}(t)$ and $p_{ij}(t)$ with 100 different wear rates. The figure is included primarily to illustrate the relation between the probabilities of being in a state at time

t ($p_{ii}(t)$) and the probabilities of moving to a new state at time t ($p_{ij}(t)$, i.e. that the cleaning pig is worn).

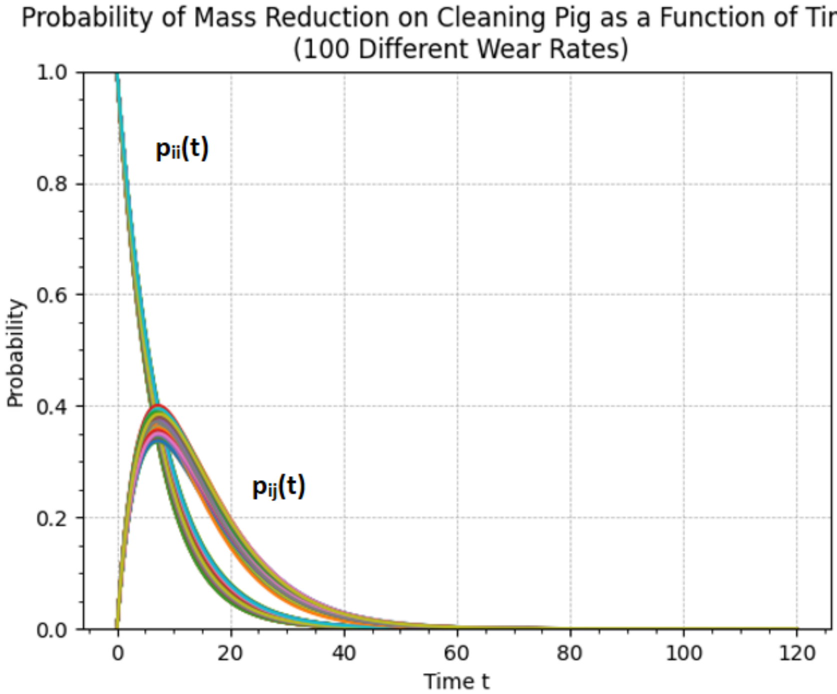


Figure 7.9: Transition probabilities ($p_{ii}(t)$, $p_{ij}(t)$) for 100 random wear rates (based on test data)

In this figure, where the wear rates are based on test data, one can see that in the time interval between 5-10 hours, the probabilities $p_{ij}(t)$ becomes larger than the $p_{ii}(t)$. Further, the $p_{ii}(t)$'s is rapidly decreasing, and after around 10 hours, the probability of a state change is always larger than the probability of staying in a state.

7.2.2 Results from Case Study 2: Probability of Being in State 8 at Time t

Figure 7.10 shows the probability of 4% and 8% wear as a function of time t. It is important to recall the assumptions that were made, namely that the wear is exponentially distributed and that the wear rate here is the same for each percent/state.

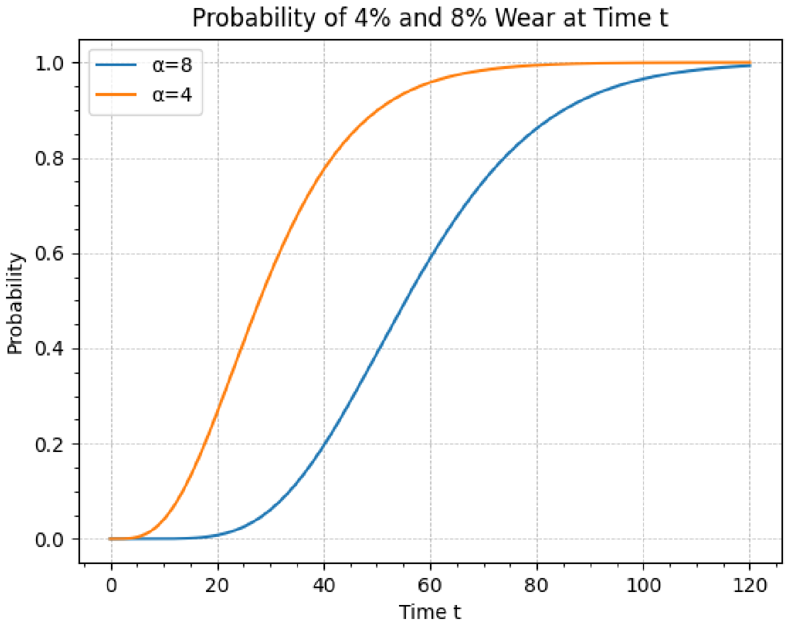


Figure 7.10: Assuming exponential wear, probability of 4% and 8% wear after time t

In this figure, the orange curve represents the probability of 4% mass reduction as a function of time, while the blue curve represents 8% mass reduction. Recall that the total time the cleaning pig spent cleaning in the qualification test rig was $\approx$ 30 hours. The total wear was then 4,2%. From Figure 7.10, one can see that the probability of being in state 4 before 30 hours is $\approx$ 60% ($P(T \leq 30) = 0,6$). Further, the probability of being in state 8 (i.e. 8% mass reduction) before 60 hours of operation is also $\approx$ 60%.

Chapter 8

Discussion

This chapter discusses the results presented in Chapter 7 in light of the thesis objectives and research question. The first section is devoted to thesis objective 1 and 2(a), while the second section is devoted to thesis objective 2(b). Recall that the first section in Chapter 6 is also a discussion devoted to thesis objective 2(b).

8.1 AMPL Configurations and Maintenance Strategies

8.1.1 Decision Making and Selection of Optimal AMPL Configuration

By just looking at the results presented in Chapter 7, it seems obvious that if one wishes to increase the availability of this subsystem, one should put the AMPLs in a parallel configuration. However, there is another factor that should be taken into account in the decision-making process, namely costs. Three relevant cost factors associated with an extra AMPL are:

- Production of an extra AMPL;
- Possible additional maintenance, repair, inspection and refill of cleaning pigs;
- Design: Integration of a new AMPL in the design.

Although the total cost of the WCS or the costs of the components are yet unknown, some estimates exist. It is assumed that the total cost of the WCS is somewhere around 100 million USD (i.e. a system with one AMPL). Even though this is just an assumption, it does give some indication concerning the potential consequences of the decisions that must be made.

This is a costly system, compromised by costly components. If it can not be shown that an extra AMPL increases the system's availability, why include one? In the end, the system is sold to customers, and they must have the funds available and see that they get value for money.

On the other hand, it is reasonable to assume that if something were to happen with such a costly system (for instance, the cleaning pig gets stuck in the AMPL, or that pig magazine is empty when the last cleaning pig is worn long before estimated), the repair would be costly as well. Such events are not only costly to fix, but they could also affect the income, as the oil exportation must be paused. Moreover, these "surprising events" can give an image of an unreliable system, which could damage the reputation of the company who sold it.

This discussion aims not to present the final answer but rather present arguments and concepts worth taking into consideration. One concept that should be included in this discussion is presented by Petroleum Safety Authority Norway (PSA). In their facility regulation, Section 5: "Design of Facilities", they state that:

"Facilities shall be based on the most robust and simple solutions as possible and designed so that
a) no unacceptable consequences will occur if they are exposed to the loads as mentioned in Section 11,
b) major accident risk is as low as possible,
c) a failure in one component, system, or a single mistake does not result in unacceptable consequences,..." [23, Sec. 5, p. 8]

The first part of the quotation can be difficult to use either for or against a parallel structure. The facility design is most likely more simple with one AMPL instead of two. However, the solution is also most likely more robust with two AMPLs instead of one. Moreover, this quotation allover argues in favour of a parallel structure:
a) Less of a chance of "unacceptable consequences" with two AMPLs instead of one.
b) "Major accident risk" is arguably lower with two AMPLs instead of one (i.e. higher availability).
c) Unacceptable consequences are avoided if something happens with one AMPL because another one is standby.

A final concept worth including in this discussion is the ALARP principle. The ALARP principle states that risk should be reduced to a level that is as low as reasonably practicable. In other words, this means that unless it can be demonstrated that the benefits of implementing the risk-reducing measure (here, an extra AMPL) is in gross disproportion to the costs associated with implementing it, it should be implemented [21, pp. 31-32], [24, p. 173].

Assume now that the cost associated with implementing another AMPL in parallel is around 30 million USD. The cost of a repair is twice as much, 60 million USD. From the figures in Chapter 7, The availability with an extra AMPL is overall higher than with a series structure. However, for values of $\mu > 26$ (i.e. MTTR<2 weeks), the availability of a series structure is over 96%, and the differences between a parallel structure and a series structure become smaller and smaller as μ increases. One should therefore have an estimate of how long a repair would take. As the availability differences decrease as the repair time decreases, it can be difficult to argue for adding an extra 30 million USD to the price. For the lowest values of μ, however, arguing for an extra AMPL is easier, as the difference in availability is larger.

This principle can also be used in the discussion regarding the choice of maintenance strategy for parallel structures. From the graphs presented in Chapter 7, one can see that the differences in availability between the two maintenance strategies are small. As the AMPL is assumed to have a rather long MTTF, the difference in resources (and also, costs associated with this type of operation) required to collect the AMPL are also rather small. Repairing once one of the two has failed therefore seems like a maintenance strategy that reduces risk. The costs associated with this maintenance strategy is not in gross disproportion with the risk-reducing measure.

The above-mentioned discussion is, of course, based on the various assumptions made in the analysis. Although one might argue that the assumptions made are reasonable overall, some of the results might be different if more accurate data were used. For example, it was assumed that repair and refill took equal time to complete, which does not seem reasonable. Also, if all of the used parameters were known (i.e. the failure rate λ_F, empty rate λ_E, repair rate μ_F and refill rate μ_E), it would be interesting to see the plot as a function of time t, instead of using steady-state probabilities. Also, it was assumed that an AMPL could fail even though it is in standby mode. While this can be true in practice, the functioning and standby AMPL will have different failure rates, as the one functioning is also affected by the cleaning pig. In the analysis, it was assumed that these failure rates are equal. However, this assumption does make sense. After the functioning AMPL has released a cleaning pig from its magazine, it is about 3-6 months until next time. In the meantime, both AMPLs are only subject to external subsea forces.

It was assumed that recovery and refill of AMPL magazines is only done once both of the AMPL's magazine is empty. Another maintenance strategy could be to recover and refill one AMPL once one is empty. This strategy was not considered in this analysis because one of the benefits with a parallel configuration over a series structure in terms of costs is that recovery and refill do not need to happen every 1-2,5 years (assuming, of course, that no failures occur in that time interval). It could be interesting to see the availability when such a maintenance strategy is used. This is then a suggestion for further research.

As a final note, it was assumed that each pig is replaced after either three or six months in the analysis. While this might be a strategy also used in practice, it could be interesting to see what the availability of a single AMPL would be if each cleaning pig's lifetime is different from one another. Such an analysis can be found in Appendix B.

8.1.2 The Usefulness of Markov Chain Theory as an Analysis Tool

To discuss the usefulness of the Markov chain as an analysis tool of new subsea equipment, it could be of interest to compare the graphs presented in Section 7.1 with the graphs generated when the limiting availability formula is used instead. Recall the limiting availability formula for a component i from equation 3.32, first presented in Chapter 3.

$$A_i = \frac{MTTF}{MTTF + MTTR},$$
$$A_i = \frac{\mu}{\mu + \lambda}.$$

For a parallel structure with identical components, one will get that

$$A_{parallel} = (1 - (1 - \frac{MTTF}{MTTF + MTTR})(1 - \frac{MTTF}{MTTF + MTTR}).$$

In Figure 8.1 and 8.2, A_i and $A_{parallel}$ is plotted as a function of μ together with the functions generated in Chapter 5. The λ in A_i and $A_{parallel}$ is set to be equal to the empty rate λ_E ($\lambda_E = 0,4$ and $\lambda_E = 0,8$). The reason for doing so is that failure is here defined as not being to function on demand. If the AMPL's pig magazine is empty, it will not function. The magazine has an empty rate which is higher than the failure rate λ_F. By using λ_E, one would then get the most pessimistic result. As for the functions generated in Chapter 5, the value λ_F used in these plots for the sake of comparison is 0,05.
As one can see, for the series structure, the graph for the limiting availability formula

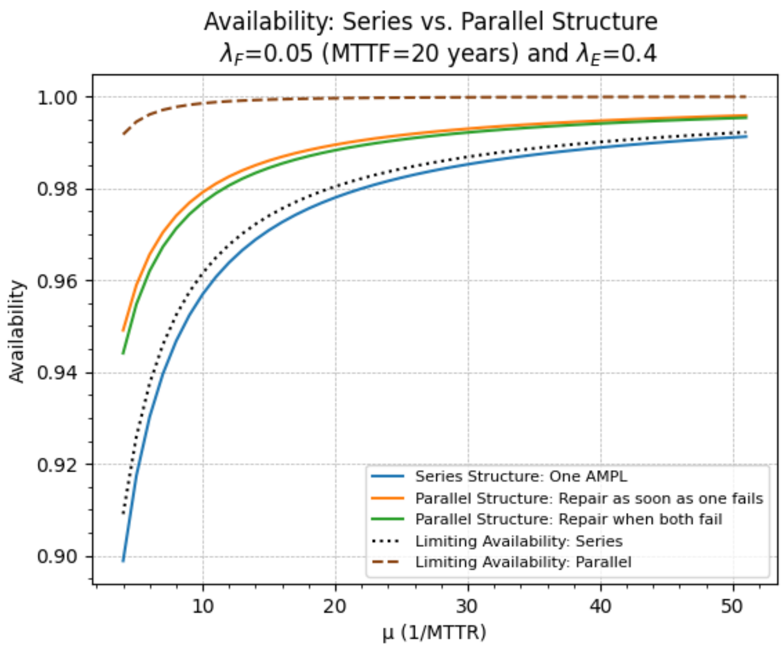


Figure 8.1: Availability using limiting availability formula, compared to functions derived in Chapter 5 ($\lambda_E = 0,4$).

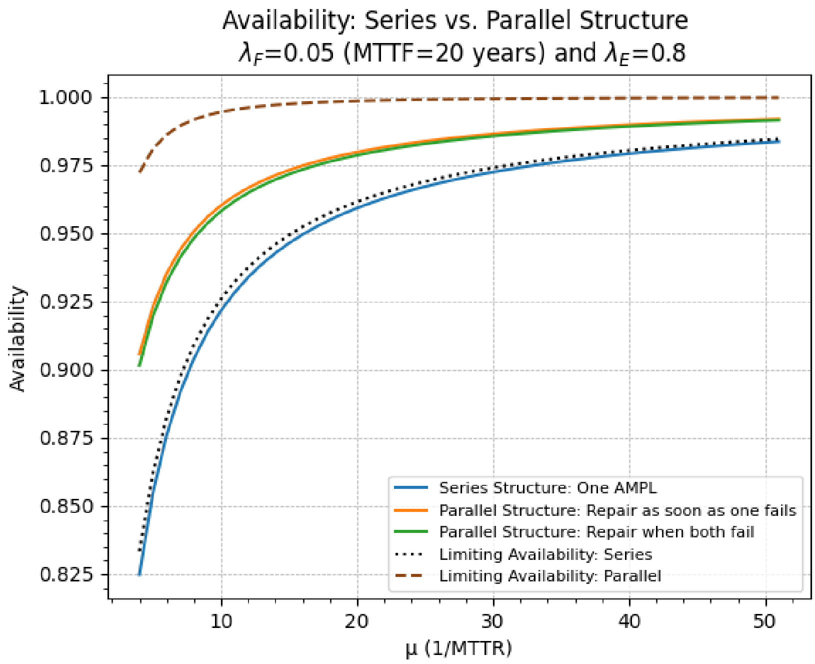


Figure 8.2: Availability using limiting availability formula, compared to functions derived Chapter 5 ($\lambda_E = 0,8$).

is almost equal to the graph generated for a single AMPL using a Markov chain. This makes sense, as the λ_F is so small that the transition graph is approximately equal to the one presented in Figure 8.3, where state 1 represents that the AMPL is functioning, and state 0 represents that the AMPL is in a state of failure. Moreover, this is the transition graph used in [18, Example 5.4(A), pp. 242-243], [6, Example 8.5, pp. 313-314] and [17, Example 6.11, pp. 371-372] to derive the limiting availability formula.

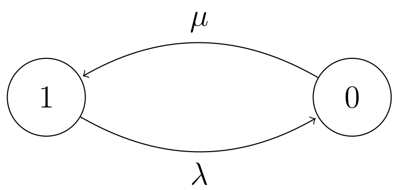


Figure 8.3: Transition graph for a single AMPL

As for $A_{parallel}$, however, there seems to be a higher deviation between the two analysis methods. For almost every value of μ, the availability is approximately equal to one. Also, notice that the availability is assumed to be that high, even with a MTTF as low as 1,25 and 2,5 years.

These graphs open up for two possible conclusions concerning the results: (1) Both the assumptions made and the conducted analysis are wrong. The results are not representative for a real-life AMPL. (2) The limiting availability formula is not a very good approximation for a subsystem consisting of two identical components in parallel, where each component can have more than two states.

Representative or not, some conclusions regarding the usage of the Markov chain can still be drawn from this AMPL analysis. Using the concept of states, it is easy to get an overview of the component, subsystem, or system. The level of detail (and hence, the number of states) one wishes to use is (to some extent) only limited to the imagination. Moreover, combined with, for instance, Python scripts, the analysis can easily be modified if an error is uncovered in the transition graph. One can say that the Markov chain as an analysis tool is quite flexible. The flexibility and level of detail achieved using Markov chain theory are reflected by the fact that dependencies between components can now be identified and modelled quite easily. This is also reflected by the analysis in Appendix B, where an even more detailed Markov chain is presented than in Chapter 5.
The AMPL acts as a good example to demonstrate the usefulness of the Markov chain as a tool. The fact that its availability is dependent of another component's lifetime means that it has states other than just functioning/failure. The number of states can, for instance, be based on how one defines failure. Is the AMPL in a failure state when the magazine is empty? Repair is unnecessary, but one may argue that as long as it is unable to function on demand, it is in a failure state.

8.2 Cleaning Pig Case Studies

While Markov chain seems to fit good as a tool to get an overview of the AMPL, the same can not be said about the cleaning pig. This is shown through the case studies built around the wear of the cleaning pig. It was, for instance, shown in Figure 7.10 that there is approximately 60% chance that the cleaning pig has reduced its mass by 4% before 30 hours of cleaning. However, through testing, it was shown that the mass reduction was indeed $\approx$ 4% after 30 hours. It is, therefore, reasonable to assume that this estimated probability should be higher.

One of the reasons why the analysis deviates from testing might be because of the memoryless property. It is assumed that the future only depends on the present, not the past, in a Markov chain. That does not make much sense when studying wear. The wear rate in the future should be dependent on the rate of wear in the past.

A stochastic process that might be a better fit for this analysis is the semi-Markov process. In this process, the random variable is not restricted to be exponentially distributed. Hence, the memoryless property is avoided. One of the reasons why this stochastic process was not chosen lies in the lack of test data available. For example, an often-used PDF in reliability calculation is the Weibull distribution. Estimators for the shape and scale parameters, identified using the MLE technique, are presented in [6, pp. 284-285], [25, p. 1526, equation (5)] and [8, p. 288] as

$$\hat{\alpha} = \frac{n}{t_0^{\hat{\beta}}},$$
$$\hat{\beta} = \frac{n}{n\,ln(t_0) - \sum_{i=1}^{n} ln(s_i)}.$$

Where:

- $\hat{\alpha}$: Scale parameter estimator;
- $\hat{\beta}$: Shape parameter estimator;
- n: Number of failures;
- t_0: The testing time interval from $t = 0$ to $t = t_0$;
- s_i: The times when failures were observed.

Finding the scale and shape parameter can, in this case, be troublesome, as there are no recorded failures (n) and, as a consequence, no times where failures were observed (s_i).

On the other hand, if one defines failure as "percentage mass reduction", one does have four failures in a time interval of ≈ 30 hours. The trouble is then to find the time after which mass was reduced by 1% and 2%. Only the times for 3% and 4% mass reduction are available. Also, wear is not strictly speaking a failure, as the cleaning pig functions fine after some wear and mass reduction.

Although estimating the Weibull shape and scale parameters is not part of the analysis, an effort has been made to use the Weibull distribution to describe the cleaning pig's wear. This work and the assumptions made can be found in Appendix C

Even though the Markov chain, in this case, does not seem like a good analysis tool, the transition graphs acts can at least act as good illustrations of the wear process. This again shows the flexibility one gets by using Markov chains. One can define states and the relation between these exactly as one wishes. Analysing mass reduction percent can be seen as a coarse analysis. Therefore, it could be interesting to see if a better conclusion could be drawn if the wear was studied per thousandth (‰) instead.

Chapter 9

Conclusion

As an aid to help conclude, recall the research question and the thesis objectives presented in Chapter 1:

1. Analyse, compare, and evaluate two different configurations in light of availability, maintainability, and costs using Markov chains.

2. Use the WCS to construct case studies and discuss how and if Markov chains fits to describe components under conditions not commonly discussed in reliability theory. More specifically:

    (a) One case where the availability of one component might depend on the lifetime of another component, and

    (b) A system with a non-stationary component, where commonly used tools and methodologies within risk and reliability in some areas might fall short.

How do Markov chains fit as a tool in the initial reliability analysis of complex components in a new subsea system?

From the results, one could say that, not surprisingly, AMPLs in a parallel structure would increase the availability compared to a series structure. The analysis looked at availability as a function of the repair rate μ. The reason for doing so was that the retrieve and repair time could vary from installation fields. Other factors should be included in the decision-making process, with one of them being costs. Are the costs associated with implementing an extra AMPL in gross disproportion to the benefits gained? And if not, what about maintenance strategy? From the plots, it seems like there is little difference in availability. Still, the costs of retrieving the AMPLs is here affected by numerous factors, like the AMPL's estimated MTTF and also the estimated lifetime of the cleaning pig.

The main tool used throughout the analyses was the Markov chain. Compared to the limiting availability formula, the use of Markov chains and transition graphs are useful to get an overview of a component or system, especially when a component's functionality depends on multiple external factors. This overview, in turn, can serve as a good basis for a more detailed analysis, which otherwise would not be possible if, for instance, the more well-known limiting availability formula was used. Finally, setting up the analysis in, for example, Python can make the analysis more flexible. By flexible, it is meant that if errors in the analysis are discovered, or changes occur, the analysis and presentation of the results can easily be updated.

One drawback with the Markov chain as a tool is the memoryless property of the exponential distribution. This was shown through the case studies of the cleaning pig's wear as a function of time. Another stochastic process that might be a better fit in cases where it seems obvious that the failure rate is not constant is the semi-Markov process. Although the results from these analyses indicate that the Markov chain indeed has some drawbacks, the transition graphs can be considered useful as they do help with getting an overview. The level of detail in this overview can, for instance, be determined by how one defines failure (for instance, mass reduction per % or ‰). This example then also shows the flexibility.

Many of the examples presented in reliability literature focuses on stationary components. In Chapter 6, it was stressed that applying existing tools and methods on a non-stationary component is not that straightforward. The reason could be that the probability of failure is continuously changing for each completed lap and vary for each component the cleaning pig passes through.

In short, the fit of Markov chain theory as a tool in the initial reliability analysis of new subsea components can be described as good. Initially, it is reasonable to assume that it is important to get a good overview. The transition graphs serve as good tools to do so. The level of detail in the analyses is affected by the analysts and how one defines failure. This shows the flexibility possible to achieve with Markov chains. Finally, it is important to bear in mind that it works as a tool for all applications. Even though one can get a good overview, the property that makes the Markov chain as a tool easy to use could also be the property that makes the analysis in itself oversimplified and unvarnished.

References

[1] Ø. Stangeland, S. Daasvatn, and Y. Novoseltsev, “Qualification of wax control system,” presented at the 52nd Offshore Technology Conference, NRG Park, Houston, Texas, USA, Aug. 16 - 19, 2021, 15.06.21: Yet to be published.

[2] *Technology Qualification*, DNV-RP-A203, Det Norske Veritas AS, July 2013. [Online]. Available: https://rules.dnvgl.com/docs/pdf/dnv/codes/docs/2013-07/rp-a203.pdf.

[3] Subsea 7 Norway, Youtube. *Subsea 7 Webinar Wax Control Technology.* (Jun. 15, 2020). Accessed: Jan. 8, 2021. [Online Video]. Available: https://www.youtube.com/watch?v=t-A3WwpJr3E&feature=youtu.be.

[4] Subsea 7 Norway. (May 27, 2020). Qualification of Wax Control System. [Internal presentation]. Accessed: Jan. 11, 2021. Unpublished.

[5] Subsea 7 Norway. (Jun. 13, 2018). Wax Control System-Qualification Rig. [Internal presentation]. Accessed: Jan. 11, 2021. Unpublished.

[6] M. Rausland and A. Høyland, *System Reliability Theory: Models, Statistical Methods, and Applications*, 2nd ed. Hoboken, New Jersey, USA: John Wiley & Sons, 2004.

[7] T. Aven, *Pålitelighets- og risikoanalyse*, 4th ed. Oslo, Norway: Universitetsforlaget, 2012, vol. 2.

[8] T. Aven, *Reliability and Risk Analysis*, 1st ed. Barking, England: Elsevier Science Publishers Ltd., 1992.

[9] G. G. Løvås, *Statistikk for universiteter og høgskoler*, 3rd ed. Oslo, Norway: Universitetsforlaget, 2015, vol. 2, pp. 88–110.

[10] R. E. Walpole, R. H. Myers, S. L. Myers, and K. Ye, *Probability & Statistics for Engineers & Scientists*, 9th ed. Harlow, Essex, England: Pearson Education, 2016.

[11] S. Ghahramani, *Fundamentals of Probability: With Stochastic Processes*, 3rd ed. Upper Saddle River, New Jersey, USA: Pearson Prentice Hall, 2004.

[12] A. L. Sweet, “On the hazard rate of the lognormal distribution,” *IEEE Transactions on Reliability*, vol. 39, no. 3, pp. 325–328, 1990. DOI: 10.1109/24.103012.

[13] E. L. Lehmann, *Elements of Large-Sample Theory*, ser. Springer Texts in Statistics. New York, New York, USA: Springer, 1999, pp. 451–462, 525–526, 541.

[14] J.T. Kvaløy. (2019). Maximum likelihood estimation. [PDF]. Course Material in STA500-Probability and Statistics 2, University of Stavanger, autumn 2019. Unpublished.

[15] M. J. Panik, *Advanced Statistics from an Elementary Point of View*. San Diego, California, USA: Elsevier Science & Technology, 2005, vol. 9, pp. 411–421, 428–429.

[16] T.S. Kleppe. (2019). Bayesian Statistics. Dated: Aug. 20 2015. [PDF]. Course Material in STA500-Probability and Statistics 2, University of Stavanger, autumn 2019. Unpublished.

[17] S. M. Ross, *Introduction to Probability Models*, 11th ed. St. Louis, Missouri, USA: Elsevier Science & Technology, 2014.

[18] S. M. Ross, *Stochastic Processes*, 2nd ed. Hoboken, New Jersey, USA: John Wiley & Sons, 1996.

[19] T.S. Kleppe and J.T. Kvaløy (2019). Markov Processes. Dated: Aug. 6 2018. [PDF]. Course Material in STA500-Probability and Statistics 2, University of Stavanger, autumn 2019. Unpublished.

[20] J.T. Selvik. (Feb. 21, 2020). Lecture 8: Optimal maintenance intervals & replacement models. [Powerpoint presentation]. Lecture note in RIS510-Reliability and Risk Analysis, University of Stavanger, spring 2020. Unpublished.

[21] T. Aven, *Risk Analysis*, 2nd ed. Chichester, West Sussex, England: John Wiley & Sons, 2015, pp. 1–2, 31–83.

[22] SINTEF and NTNU-Norges Teknisk-Naturvitenskapelige Universitet, *OREDA: Offshore and Onshore Reliability Data*, 6th ed. Trondheim, Norway: OREDA Participants, 2015, vol. 2-*Subsea Equipment*.

[23] P. S. A. Norway, *The Facilities Regulations: Regulations Relating to Design and Outfitting of Facilities, etc. in the Petroleum Activities*, 16 December 2020. [online]. Available: https://www.ptil.no/en/regulations/pdfs-of-regulations/.

[24] T. Aven, *Risk, Surprises and Black Swans: Fundamental Ideas and Concepts in Risk Assessment and Risk Management*, 1st ed. London, England: Routledge, 2014.

[25] O. Gaudoin, B. Yang, and M. Xie, "Confidence intervals for the scale parameter of the power-law process," *Communication in Statistics: Theory and Methods*, vol. 35, no. 8, pp. 1525–1538, 2006. DOI: 10.1080/03610920600637412.

Appendix A

Python Scripts Made for the Analyses

A.1 Equation Set Solvers

Different μ (refill rate $\neq$ repair rate)

```python
import sympy
pi_0, pi_1, pi_2, pi_3, pi_4, pi_5, Lambda_E, Lambda_F, mu_E, mu_F \
    = sympy.symbols("pi_0, pi_1, pi_2, pi_3, pi_4, pi_5, Lambda_E, Lambda_F, mu_E, mu_F")

eq_sum_s = sympy.Eq(pi_0 + pi_1 + pi_2 - 1, 0)
eq_sum_p = sympy.Eq(pi_0 + pi_1 + pi_2 + pi_3 + pi_4 + pi_5 - 1, 0)

series_0 = sympy.Eq(pi_2 * Lambda_F - pi_0 * mu_F, 0)
series_1 = sympy.Eq(pi_2 * Lambda_E - pi_1 * mu_E, 0)
series_2 = sympy.Eq(pi_0 * mu_F + pi_1 * mu_E - pi_2 * Lambda_F - pi_2 * Lambda_E, 0)

parallel1_0 = sympy.Eq(pi_3 * Lambda_F - pi_0 * mu_F, 0)
parallel1_1 = sympy.Eq(pi_4 * Lambda_E - pi_1 * mu_E, 0)
parallel1_2 = sympy.Eq(pi_4 * Lambda_F + pi_3 * Lambda_E - pi_2 * mu_F, 0)
parallel1_3 = sympy.Eq(pi_5 * 2 * Lambda_F - pi_3 * mu_F - pi_3 * Lambda_F - pi_3 * Lambda_E,
0)
parallel1_4 = sympy.Eq(pi_5 * Lambda_E - pi_4 * Lambda_E - pi_4 * Lambda_F, 0)
parallel1_5 = sympy.Eq(pi_3 * mu_F + pi_2 * mu_F + pi_1 * mu_E + pi_0 * mu_F - pi_5 * Lambda_E
- pi_5 * 2 * Lambda_F, 0)

parallel2_0 = sympy.Eq(pi_3 * Lambda_F - pi_0 * mu_F, 0)
parallel2_1 = sympy.Eq(pi_4 * Lambda_E - pi_1 * mu_E, 0)
parallel2_2 = sympy.Eq(pi_4 * Lambda_F + pi_3 * Lambda_E - pi_2 * mu_F, 0)
parallel2_3 = sympy.Eq(pi_5 * 2 * Lambda_F - pi_3 * Lambda_F - pi_3 * Lambda_E, 0)
parallel2_4 = sympy.Eq(pi_5 * Lambda_E - pi_4 * Lambda_E - pi_4 * Lambda_F, 0)
parallel2_5 = sympy.Eq(pi_2 * mu_F + pi_1 * mu_E + pi_0 * mu_F - pi_5 * Lambda_E - pi_5 * 2 *
Lambda_F, 0)

sol1 = sympy.solve([series_0, series_1, series_2, eq_sum_s], [pi_0, pi_1, pi_2])
sol2 = sympy.solve([parallel1_0, parallel1_1, parallel1_2, parallel1_3, parallel1_4,
parallel1_5, eq_sum_p],
                   [pi_0, pi_1, pi_2, pi_3, pi_4, pi_5])
sol3 = sympy.solve([parallel2_0, parallel2_1, parallel2_2, parallel2_3, parallel2_4,
parallel2_5, eq_sum_p],
                   [pi_0, pi_1, pi_2, pi_3, pi_4, pi_5])


def expressions_in_list(sol_1_2_3):
    expressions = list(sol_1_2_3.values())
    if len(expressions) == 3:
        series = expressions[-1]
        return series
    elif len(expressions) == 6:
        parallel = expressions[-1]+expressions[-2]+expressions[-3]
        return parallel
```

Common μ (refill rate = repair rate)

```python
# Filename: steady_state_prob_calculator_common_mu
# File location (folder name): Master

import sympy
pi_0, pi_1, pi_2, pi_3, pi_4, pi_5, Lambda_E, Lambda_F, mu\
    = sympy.symbols("pi_0, pi_1, pi_2, pi_3, pi_4, pi_5, Lambda_E, Lambda_F, mu")

eq_sum_s = sympy.Eq(pi_0 + pi_1 + pi_2 - 1, 0)
eq_sum_p = sympy.Eq(pi_0 + pi_1 + pi_2 + pi_3 + pi_4 + pi_5 - 1, 0)

series_0 = sympy.Eq(pi_2 * Lambda_F - pi_0 * mu, 0)
series_1 = sympy.Eq(pi_2 * Lambda_E - pi_1 * mu, 0)
series_2 = sympy.Eq(pi_0 * mu + pi_1 * mu - pi_2 * Lambda_F - pi_2 * Lambda_E, 0)

parallel1_0 = sympy.Eq(pi_3 * Lambda_F - pi_0 * mu, 0)
parallel1_1 = sympy.Eq(pi_4 * Lambda_E - pi_1 * mu, 0)
parallel1_2 = sympy.Eq(pi_4 * Lambda_F + pi_3 * Lambda_E - pi_2 * mu, 0)
parallel1_3 = sympy.Eq(pi_5 * 2 * Lambda_F - pi_3 * mu - pi_3 * Lambda_F - pi_3 * Lambda_E, 0)
parallel1_4 = sympy.Eq(pi_5 * Lambda_E - pi_4 * Lambda_E - pi_4 * Lambda_F, 0)
parallel1_5 = sympy.Eq(pi_3 * mu + pi_2 * mu + pi_1 * mu + pi_0 * mu - pi_5 * Lambda_E - pi_5
* 2 * Lambda_F, 0)

parallel2_0 = sympy.Eq(pi_3 * Lambda_F - pi_0 * mu, 0)
parallel2_1 = sympy.Eq(pi_4 * Lambda_E - pi_1 * mu, 0)
parallel2_2 = sympy.Eq(pi_4 * Lambda_F + pi_3 * Lambda_E - pi_2 * mu, 0)
parallel2_3 = sympy.Eq(pi_5 * 2 * Lambda_F - pi_3 * Lambda_F - pi_3 * Lambda_E, 0)
parallel2_4 = sympy.Eq(pi_5 * Lambda_E - pi_4 * Lambda_E - pi_4 * Lambda_F, 0)
parallel2_5 = sympy.Eq(pi_2 * mu + pi_1 * mu + pi_0 * mu - pi_5 * Lambda_E - pi_5 * 2 *
Lambda_F, 0)

sol1 = sympy.solve([series_0, series_1, series_2, eq_sum_s], [pi_0, pi_1, pi_2])
sol2 = sympy.solve([parallel1_0, parallel1_1, parallel1_2, parallel1_3, parallel1_4,
parallel1_5, eq_sum_p],
                   [pi_0, pi_1, pi_2, pi_3, pi_4, pi_5])
sol3 = sympy.solve([parallel2_0, parallel2_1, parallel2_2, parallel2_3, parallel2_4,
parallel2_5, eq_sum_p],
                   [pi_0, pi_1, pi_2, pi_3, pi_4, pi_5])


def expressions_in_list(sol_1_2_3):
    expressions = list(sol_1_2_3.values())
    if len(expressions) == 3:
        series = expressions[-1]
        return series
    elif len(expressions) == 6:
        parallel = expressions[-1]+expressions[-2]+expressions[-3]
        return parallel
```

A.2 Availability Plotting

Availability as a Function of the Repair/Refill Rate μ (2D)

```
import sympy as sp
import numpy as np
from Master.steady_state_prob_calculator_commom_mu import sol1, sol2, sol3,
expressions_in_list
import matplotlib.pyplot as plt

mu = sp.Symbol("mu")
Lambda_E = sp.Symbol("Lambda_E")
Lambda_F = sp.Symbol("Lambda_F")

series = expressions_in_list(sol1)
parallel_1 = expressions_in_list(sol2)
parallel_2 = expressions_in_list(sol3)

SERIES = sp.lambdify((mu, Lambda_E, Lambda_F), series, "numpy")
PARALLEL_1 = sp.lambdify((mu, Lambda_E, Lambda_F), parallel_1, "numpy")
PARALLEL_2 = sp.lambdify((mu, Lambda_E, Lambda_F), parallel_2, "numpy")

LAMBDA_E = 0.4
# Lambda_E=0.8: Refill after 5x 3months
# LAMBDA_E=0.4: Refill after 5x 6months

LAMBDA_F = 0.05
# LAMBDA_F=0.1 (MTTF=10 years)
# LAMBDA_F=0.05 (MTTF=20 years)
# LAMBDA_F=0.019 (MTTF=52.6 years)

MU = np.arange(4, 52, 1)
# MU ranging from 4 to 52 (MTTR from 3 months to 1 week)

# Limiting availability comparison (for discussion)
limiting_availability_series = MU/(MU+LAMBDA_E)
limiting_availability_parallel = 1-(1-(MU/(LAMBDA_E+MU)))*(1-(MU/(LAMBDA_E+MU)))

plt.plot(MU, SERIES(MU, LAMBDA_E, LAMBDA_F), label="Series Structure: One AMPL")
plt.plot(MU, PARALLEL_1(MU, LAMBDA_E, LAMBDA_F), label="Parallel Structure: Repair as soon as
one fails")
plt.plot(MU, PARALLEL_2(MU, LAMBDA_E, LAMBDA_F), label="Parallel Structure: Repair when both
fail")
plt.plot(MU, limiting_availability_series, label="Limiting Availability: Series",
linestyle=":", color="k")
plt.plot(MU, limiting_availability_parallel, label="Limiting Availability: Parallel",
linestyle="--", color="saddlebrown")
plt.xlabel("\u03BC (1/MTTR)")
plt.ylabel("Availability")
plt.title("Availability: Series vs. Parallel Structure \n $\u03BB_{F}$=0.05 (MTTF=20 years) "
          "and $\u03BB_{E}$=0.4")
plt.grid(linestyle="--", linewidth=0.5)
plt.legend(prop={"size": 8})
plt.minorticks_on()
plt.show()
```

Availability as a Function of the Repair/Refill Rate μ and Failure Rate λ_F (3D)

```
import sympy as sp
import numpy as np
from Master.steady_state_prob_calculator_commom_mu import sol1, sol2, sol3,
expressions_in_list
import matplotlib.pyplot as plt

mu = sp.Symbol("mu")
Lambda_E = sp.Symbol("Lambda_E")
Lambda_F = sp.Symbol("Lambda_F")

series = expressions_in_list(sol1)
parallel_1 = expressions_in_list(sol2)
parallel_2 = expressions_in_list(sol3)

SERIES = sp.lambdify((mu, Lambda_E, Lambda_F), series, "numpy")
PARALLEL_1 = sp.lambdify((mu, Lambda_E, Lambda_F), parallel_1, "numpy")
PARALLEL_2 = sp.lambdify((mu, Lambda_E, Lambda_F), parallel_2, "numpy")

# LAMBDA_E = 0.4
LAMBDA_E = 0.8
LAMBDA_F = np.arange(5*10**(-3), 0.1, 10**(-3))
MU = np.arange(4, 52, 1)
Lambda_F_meshed, MU_meshed = np.meshgrid(LAMBDA_F, MU)


def do_plot(n, fun, title, vmin, vmax, xlabel=None, ylabel=None):
    plt.subplot(1, 3, n)
    plt.contourf(Lambda_F_meshed, MU_meshed, fun, 50, cmap="YlGnBu")
    plt.title(title)
    plt.clim(vmin, vmax)
    plt.colorbar()
    plt.xlabel(xlabel)
    plt.ylabel(ylabel)
    plt.minorticks_on()


plt.figure()
do_plot(1, SERIES(MU_meshed, LAMBDA_E, Lambda_F_meshed), "Series: One single AMPL", 0.87, 1,
        "\u03BB (1/MTTF)", "\u03BC (1/MTTR)")
do_plot(2, PARALLEL_1(MU_meshed, LAMBDA_E, Lambda_F_meshed), "Parallel: Repair as soon as one
fails", 0.945, 1,
        "\u03BB (1/MTTF)")
do_plot(3, PARALLEL_2(MU_meshed, LAMBDA_E, Lambda_F_meshed), "Parallel: Repair when both
fail", 0.945, 1,
        "\u03BB (1/MTTF)")
plt.suptitle("Availability: Series vs. Parallel Structure as a Function of $\u03BB_{F}$ and
\u03BC, "
             "for $\u03BB_{E}=0,8$", fontsize=16)
plt.tight_layout(pad=1)
plt.show()
```

A.3 Scripts for the Cleaning Pig Case Studies

Script for Case Study 1: Solver for Multiple Coupled Differential Equations

```python
from scipy.integrate import odeint
import numpy as np
import matplotlib.pyplot as plt

lambda_list = []
states = 100
# states = 10
# states = 1000
for i in range(states):
    random = np.random.uniform(0.12456, 0.15224)
    lambda_list.append(random)


def prob(P, t):
    stationary_prob = {}
    for k in range(states):
        stationary_prob[f"P{k}".format(k)] = P[k]

    transition_prob = {}
    count1 = 0
    for l in range(1, states):
        transition_prob[f"P{l-1}_{l}".format(l)] = P[k+l]
        count1 += 1

    diff_equations = []
    for m in range(states):
        dpiidt = -lambda_list[m] * stationary_prob[f"P{m}"]
        diff_equations.append(dpiidt)

    for n in range(1, states):
        dpijdt = lambda_list[n-1] * stationary_prob[f"P{n-1}"] \
                 - lambda_list[n] * transition_prob[f"P{n-1}_{n}"]
        diff_equations.append(dpijdt)
    return diff_equations


initial_conditions = []
for x in range(states):
    initial_conditions.append(1)
for y in range(states-1):
    initial_conditions.append(0)

t = np.linspace(0, 120, 1000)
solve = odeint(prob, initial_conditions, t)

plt.xlabel("Time t")
plt.ylabel("Probability")
axes = plt.gca()
axes.set_ylim([0, 1])

for o in range(2*states-1):
    plot = solve[:, o]
    plt.plot(t, plot)

plt.title(f"Probability of Mass Reduction on Cleaning Pig as a Function of Time t \n ({states}
Different Wear Rates)")
plt.grid(linestyle="--", linewidth=0.5)
plt.minorticks_on()
plt.show()
```

Script for Case Study 2: Gamma Distributed Wear

```
import numpy as np
from scipy.stats import gamma
import matplotlib.pyplot as plt

alpha_wear1 = 8
alpha_wear2 = 4
Lambda = 0.1384
beta = 1./Lambda

t = np.linspace(0, 120, 1000)

y1 = gamma.cdf(t, alpha_wear1, loc=0, scale=beta)
y2 = gamma.cdf(t, alpha_wear2, loc=0, scale=beta)
plt.plot(t, y1, "b", label="\u03B1=8")
plt.plot(t, y2, "r", label="\u03B1=4")
plt.legend()
plt.minorticks_on()
plt.xlabel("Time t")
plt.ylabel("Probability")
plt.title("Probability of 4% and 8% Wear at Time t")
plt.grid(linestyle="--", linewidth=0.5)
plt.show()
```

Appendix B

AMPL Analysis with Varying Cleaning Pig Lifetimes

Figure B.1 shows a Markov chain where the cleaning pigs are replaced at different times. Recall that in the analysis in Chapter 5, all the cleaning pigs were replaced after either three or six months. The interpretation of the transition graph is given under the figure.

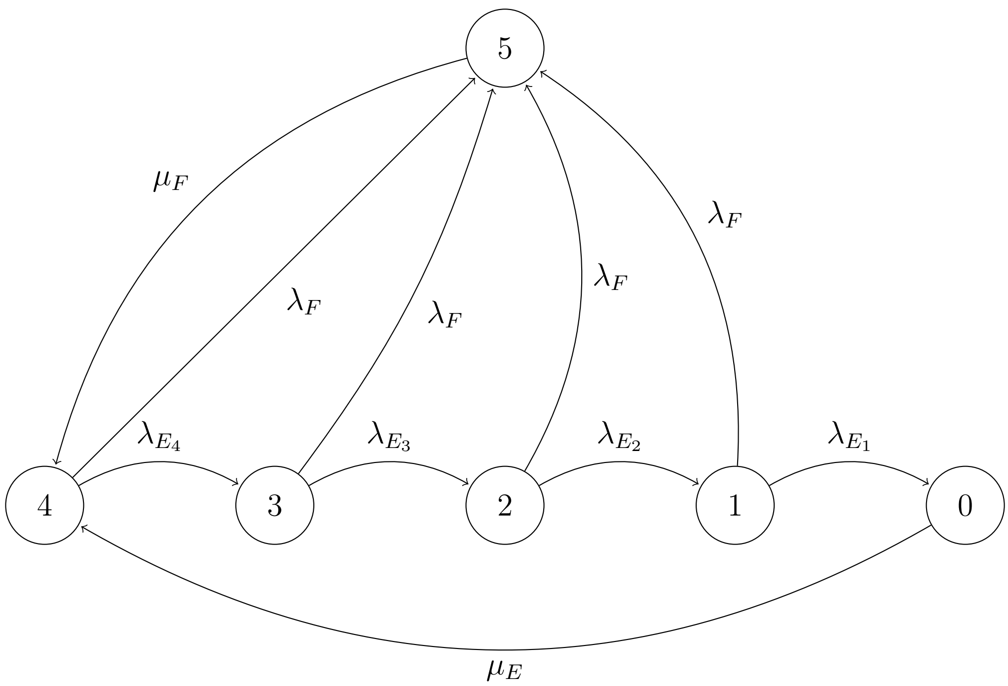


Figure B.1: A Markov chain with different empty rates

Figure B.1 should be interpreted as follows:

- State 4: There are currently 4 cleaning pigs in the AMPL's pig magazine. The AMPL can fail with rate λ_F or emptied with rate λ_{E_4}.

- State 3: There are currently 3 cleaning pigs in the AMPL's pig magazine. The AMPL can fail with rate λ_F or emptied with rate λ_{E_3}.
- State 2: There are currently 2 cleaning pigs in the AMPL's pig magazine. The AMPL can fail with rate λ_F or emptied with rate λ_{E_2}.
- State 1: There are currently 1 cleaning pig in the AMPL's pig magazine. The AMPL can fail with rate λ_F or emptied with rate λ_{E_1}.
- State 0: There are currently none cleaning pigs in the AMPL's pig magazine. The AMPL is retrieved and refilled with refill rate μ_E.
- State 5: The AMPL has failed. it is repaired and refilled with rate μ_F.
- The AMPL will function as long as it is in state 0, 1, 2, 3 or 4, but will be available as long as it is in state 1, 2, 3 or 4 (in state 0, it is sent to refill).

The transition probability matrix for this Markov chain is shown below.

$$\begin{pmatrix} -\mu_E & 0 & 0 & 0 & \mu_E & 0 \\ \lambda_{E_1} & -(\lambda_{E_1}+\lambda_F) & 0 & 0 & 0 & \lambda_F \\ 0 & \lambda_{E_2} & -(\lambda_{E_2}+\lambda_F) & 0 & 0 & \lambda_F \\ 0 & 0 & \lambda_{E_3} & -(\lambda_{E_3}+\lambda_F) & 0 & \lambda_F \\ 0 & 0 & 0 & \lambda_{E_4} & -(\lambda_{E_4}+\lambda_F) & \lambda_F \\ 0 & 0 & 0 & 0 & \mu_F & -\mu_F \end{pmatrix}$$

.

Based on the transition probability matrix and transition graph, the following equation sets can be constructed. It is assumed that steady-state probabilities can be used and that the refill rate (μ_E) is equal to the repair rate (μ_F (just like in Chapter 5).

$$\lambda_{E_1} * \pi_1 = \mu_E * \pi_0$$

$$\lambda_{E_2} * \pi_2 = \lambda_{E_1} * \pi_1 + \lambda_F * \pi_1$$

$$\lambda_{E_3} * \pi_3 = \lambda_{E_2} * \pi_2 + \lambda_F * \pi_2$$

$$\lambda_{E_4} * \pi_4 = \lambda_{E_3} * \pi_3 + \lambda_F * \pi_3$$

$$\mu_F * \pi_5 + \mu_E * \pi_0 = \lambda_{E_4} * \pi_4 + \lambda_F * \pi_4$$

$$\lambda_F * \pi_1 + \lambda_F * \pi_2 + \lambda_F * \pi_3 + \lambda_F * \pi_4 = \mu_F * \pi_5$$

$$\pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 = 1$$

This equation set is solved using a script made in Python (see Figure B.2). The functions is then plotted as a function of refill=repair rate μ. λ_F is set to be 0,05 (MTTF=20 years). λ_{E_1}, λ_{E_2}, λ_{E_3} and λ_{E_4} are drawn randomly from an interval from 2 to 4. This corresponds to a pig lifetime of six months and three months, respectively ($\frac{1}{\frac{3}{12}=4}$, $\frac{1}{\frac{6}{12}=2}$). The availability plot is shown in Figure B.3.

```
import sympy

pi_0, pi_1, pi_2, pi_3, pi_4, pi_5, Lambda_E1, Lambda_E2, Lambda_E3, Lambda_E4, Lambda_F, mu\
    = sympy.symbols("pi_0, pi_1, pi_2, pi_3, pi_4, pi_5, Lambda_E1, Lambda_E2, Lambda_E3,
Lambda_E4, "
                    "Lambda_F, mu")

eq_sum = sympy.Eq(pi_0 + pi_1 + pi_2 + pi_3 + pi_4 + pi_5 - 1, 0)

eq_0 = sympy.Eq(Lambda_E1*pi_1 - mu*pi_0, 0)
eq_1 = sympy.Eq(Lambda_E2*pi_2 - Lambda_E1*pi_1 - Lambda_F*pi_1, 0)
eq_2 = sympy.Eq(Lambda_E3*pi_3 - Lambda_E2*pi_2 - Lambda_F*pi_2, 0)
eq_3 = sympy.Eq(Lambda_E4*pi_4 - Lambda_E3*pi_3 - Lambda_F*pi_3, 0)
eq_4 = sympy.Eq(mu*pi_5 + mu*pi_0 - Lambda_E4*pi_4 - Lambda_F*pi_4, 0)
eq_5 = sympy.Eq(Lambda_F*pi_1 + Lambda_F*pi_2 + Lambda_F*pi_3 + Lambda_F*pi_4 - mu*pi_5, 0)

sol = sympy.solve([eq_0, eq_1, eq_2, eq_3, eq_4, eq_5, eq_sum], [pi_0, pi_1, pi_2, pi_3, pi_4,
pi_5])


def expressions_in_list(sol_1_2_3):
    expressions = list(sol_1_2_3.values())
    while len(expressions) == 6:
        new_series = 1 - (expressions[-1] + expressions[0])
        return new_series
```

```
import numpy
import sympy
import matplotlib.pyplot as plt
from Master.Appendix_vary_lifetime import sol, expressions_in_list

mu = sympy.Symbol("mu")
Lambda_E1 = sympy.Symbol("Lambda_E1")
Lambda_E2 = sympy.Symbol("Lambda_E2")
Lambda_E3 = sympy.Symbol("Lambda_E3")
Lambda_E4 = sympy.Symbol("Lambda_E4")
Lambda_F = sympy.Symbol("Lambda_F")

new_series = expressions_in_list(sol)
NEW_SERIES = sympy.lambdify((mu, Lambda_E1, Lambda_E2, Lambda_E3, Lambda_E4, Lambda_F),
new_series, "numpy")

MU = numpy.arange(4, 52, 1)
LAMBDA_E1 = numpy.random.uniform(2.0, 4.0)
LAMBDA_E2 = numpy.random.uniform(2.0, 4.0)
LAMBDA_E3 = numpy.random.uniform(2.0, 4.0)
LAMBDA_E4 = numpy.random.uniform(2.0, 4.0)
LAMBDA_F = 0.05  # MTTF=20 years

plt.plot(MU, NEW_SERIES(MU, LAMBDA_E1, LAMBDA_E2, LAMBDA_E3, LAMBDA_E4, LAMBDA_F))
plt.xlabel("\u03BC (1/MTTR)")
plt.ylabel("Availability")
plt.title("Availability: Different Cleaning Pig Lifetimes\n $\u03BB_{F}$=0.05 (MTTF=20 years)
"
          "and $\u03BB_{E}$=[2,4)")
plt.grid(linestyle="--", linewidth=0.5)
plt.minorticks_on()
plt.show()
```

Figure B.2: Python scripts made for solving the equation set and plotting the availability of an AMPL when pig lifetime

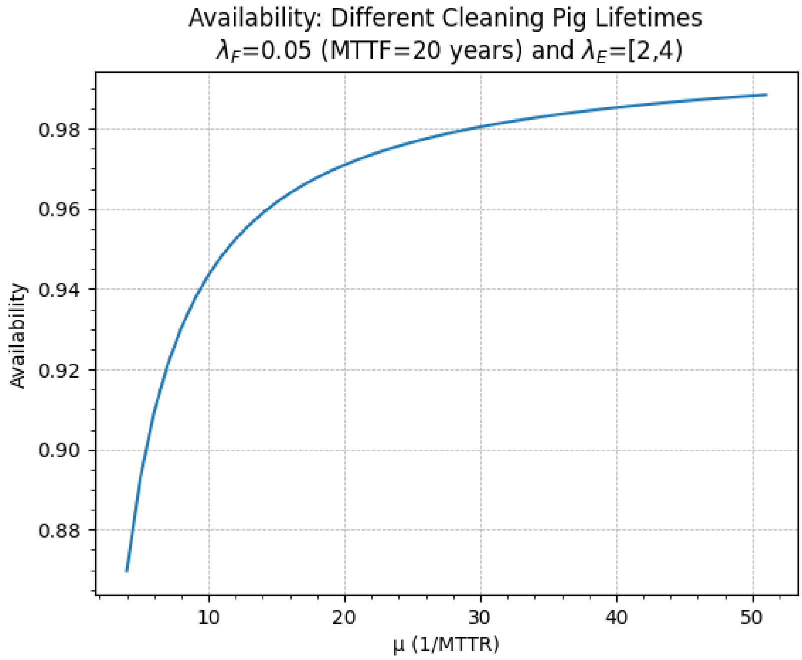


Figure B.3: Availability plot as a function of μ for an AMPL with varying empty rate λ_E

Recall Figure 7.2 and 7.6 from Chapter 7. Compared to these, the results in Figure B.3 is quite similar. when studying the relationship between these figures more thoroughly, one can see that the results in Figure B.3 is greater than in Figure 7.2 where λ_E is based on a pig lifetime of three months. Also, the results are less than in Figure 7.6, where λ_E is based on a pig lifetime of six months. This makes sense, as the λ_Es in this example are drawn from an interval from three to six months.

Appendix C

Cleaning Pig: Using Test Data to Estimate the Weibull Shape and Scale Parameter

Assume now that failure is defined as "the cleaning pig's mass is reduced by 1%". Whenever this happens in a test interval, it is recorded as a failure.
Recall that in Section 8.2, estimators for the Weibull shape and scale parameter could be found using the MLE technique. The scale parameter estimator was defined as

$$\hat{\alpha} = \frac{n}{t_0^{\hat{\beta}}},$$

while the shape parameter estimator were defined as

$$\hat{\beta} = \frac{n}{n\,ln(t_0) - \sum_{i=1}^{n} ln(s_i)}.$$

From the test data, it was shown that:

- n: Number of failures = 4 (4% of the mass is lost in the defined test interval). The actual wear was 4,2% but has been rounded in this example to 4% for the sake of simplicity.

- t_0: The testing time interval from t=0 to t=t_0. This is when 4% mass reduction is recorded, which is after $\approx$ 30 hours of operation.

The final variable that must be identified is s_i. Some assumptions must be made for s_1 and s_2, while s_3 and s_4 are easy to approximate. s_4 was found to be $\approx$ 30 hours. s_3 (i.e. $\approx$ 3% mass reduction) occurred after 1000 completed loops of length 120 meters. With a cleaning pig velocity of 1,6 $\frac{m}{s}$, the time spent before s_3 occurred was

$$\begin{aligned}
\text{distance travelled} &= 1000 * 12, \\
\text{time spent cleaning [s]} &= \frac{\text{distance travelled}}{1,6}, \\
&= \frac{1000*12}{1,6}, \\
&= 75000, \\
\text{time spent cleaning [hours]}(s_3) &= \frac{75000}{60 * 60}, \\
&= 20,83.
\end{aligned}$$

The calculation above shows that s_3 occurs after 20,83 hours $\approx$ 21 hours. For simplicity, assume that s_1 and s_2 occurs with 7 hours interval, so that $s_1 + s_2 + s_3 = 21$. By making this assumption, one is now able to estimate the Weibull shape and scale parameter.

$$\begin{aligned}
\hat{\beta} &= \frac{4}{4 * ln(30) - (ln(7) + ln(14) + ln(20,83) + ln(30))}, \\
&= 1,549, \\
\hat{\alpha} &= \frac{4}{30^{1,549}}, \\
&= 0,0206.
\end{aligned}$$

Then, recall the failure rate function for the Weibull distribution, first presented in Section 3.3.

$$Z(t) = \alpha\beta t^{\beta-1}.$$

A Python script is then made to plot the failure rate function as a function of time t (in hours). The script is shown in Figure C.1

```python
import matplotlib.pyplot as plt
import numpy as np

alpha = 0.0206  # scale
beta = 1.549  # shape

t = np.linspace(0, 500, 1000)

failure_function = alpha * beta * t**(beta-1)

plt.plot(t, failure_function)
plt.minorticks_on()
plt.xlabel("Time t")
plt.ylabel("Failure rate function \u03BB(t)=\u03B1*\u03B2*t^(\u03B2-1)")
plt.title("Weibull Failure Rate Function: Cleaning Pig Wear\n(Scale \u03B1=0.0206, Shape \u03B2=1.549)")
plt.grid(linestyle="--", linewidth=0.5)
plt.show()
```

Figure C.1: The script made for plotting the failure rate function as a function of time t

The script shown in Figure C.1 produces the plot shown in Figure C.2. Here, the failure rate function is plotted as a function of time. Recall that failure in this example is referred to as mass reduction due to wear.

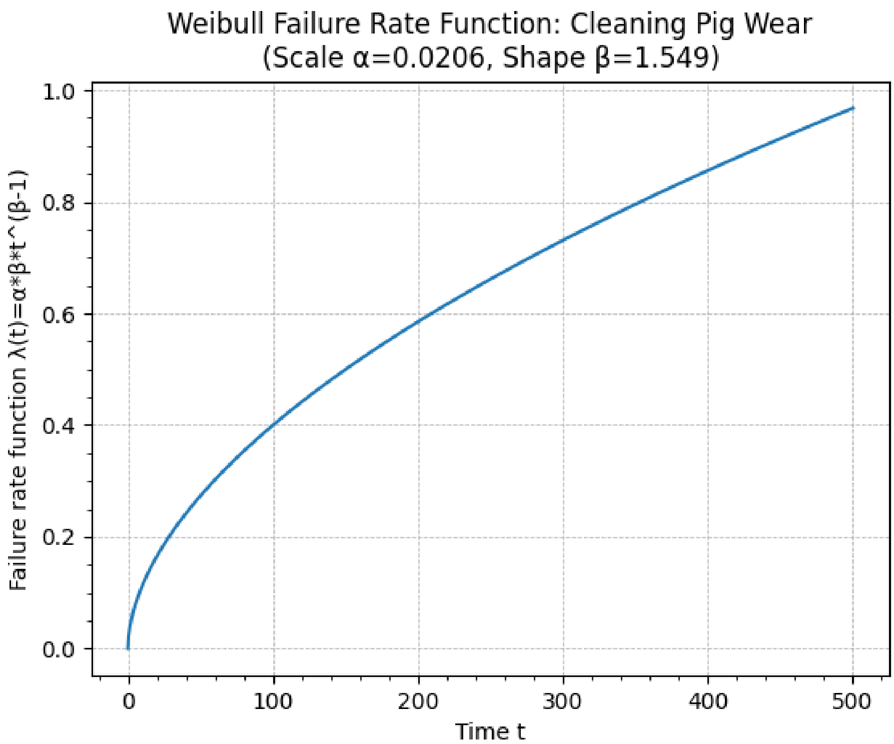


Figure C.2: The Weibull failure rate function plotted as a function of time

From the plot in Figure C.2, it seems obvious that there is an error with the assumptions that were made, the interpretation of the parameters used to find the estimates or both. According to the plot, the probability of failure (i.e. mass reduction) after around 500 hours of cleaning is close to 1. This doesn't seem right. It was found through testing that failure (mass reduction) had occurred 4 times in 30 hours.

The function used in this plot ($\lambda(t)$) is in [6, p. 284] referred to as a Weibull process. Whereas the Poisson process, which is used throughout this thesis, is homogeneous, the Weibull process is non-homogeneous. The fact that the process is non-homogeneous means that the rate at which events occur can vary over time. A consequence of this is that the theory presented in this thesis, which is based on the assumption that the random variables are independent and identically distributed, no longer holds. [6, p. 278]. Based on the definition of a non-homogeneous Poisson process, it seemed like a suitable tool to describe the cleaning pig's wear. It is reasonable to think that the rate between percentage mass reduced will vary over time.

Appendix D

Derivation of the Limiting Availability Formula

Imagine a component with two states. In state 1, the component is functioning. When in state 0, the component has failed. The transition graph for such a component is presented in Figure D.1.

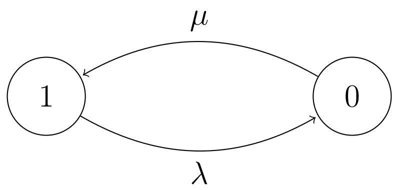


Figure D.1: Transition graph for a component with two states

Figure D.1 is interpreted as follows:

- At t=*0*, the component is in state 1.
- At a time t, the component fails (transition from state 1->0 with probability=λ).
- At a time t, the component is reinstalled and repaired (transition from state 0->1 with probability=μ).

Based on the transition graph, the following transition probability matrix can be made:

$$P_s = \begin{pmatrix} -\mu & \mu \\ \lambda & -\lambda \end{pmatrix}. \tag{D.1}$$

To find an expression for the probability that the system works at time t(i.e. $p_{11}(t)$), one can utilise the Kolmogorov forward equation. Recall the Kolmogorov forward equation first presented in equation 3.23.

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{ij}(t) = \sum_{k \neq j} q_{kj} p_{ik}(t) - \nu_j p_{ij}(t),\ i, j \in \mathcal{S}.$$

From the transition probability matrix in D.1, it can be found that $q_{01} = \mu$ and $\nu_1 = \lambda$. Inserting this in the Kolmogorov forward equation p_{11}(t) gives that

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{11}(t) = \mu p_{10}(t) - \lambda p_{11}(t)$$

It is also known that $p_{10}(t) = 1 - p_{11}(t)$. inserting this for $p_{10}(t)$ yields

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{11}(t) = \mu(1 - p_{11}) - \lambda p_{11}(t)$$

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{11}(t) = \mu - \mu p_{11}(t) - \lambda p_{11}(t)$$

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{11}(t) = \mu - (\mu + \lambda) p_{11}(t)$$

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{11}(t) + (\mu + \lambda) p_{11}(t) = \mu$$

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{11}(t) e^{(\mu+\lambda)t} + (\mu + \lambda) p_{11}(t) e^{(\mu+\lambda)t} = \mu e^{(\mu+\lambda)t}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(e^{(\mu+\lambda)t} p_{11}(t)) = \mu e^{(\mu+\lambda)t}$$

$$e^{(\mu+\lambda)t} p_{11}(t) = \int \mu e^{(\mu+\lambda)t}$$

$$e^{(\mu+\lambda)t} p_{11}(t) = \frac{\mu}{\mu + \lambda} e^{(\mu+\lambda)t} + C$$

$$p_{11}(t) = \frac{\mu}{\mu + \lambda} + C * e^{-(\mu+\lambda)t}.$$

The constant C can be found by using the initial conditions that at time t=0, the is in state 1 (i.e. $p_{11}(0) = 1$).

$$p_{11}(0) = \frac{\mu}{\mu+\lambda} + C * e^{-(\mu+\lambda)*0} = 1$$

$$\frac{\mu}{\mu+\lambda} + C = 1$$

$$\mu + C * (\mu+\lambda) = (\mu+\lambda)$$

$$C * (\mu+\lambda) = (\mu+\lambda) - \mu$$

$$C = \frac{(\mu+\lambda)}{(\mu+\lambda)} - \frac{\mu}{(\mu+\lambda)}$$

$$C = \frac{\mu+\lambda-\mu}{(\mu+\lambda)}$$

$$C = \frac{\lambda}{\mu+\lambda}.$$

Inserting the expression for C gives that

$$p_{11}(t) = \frac{\mu}{\mu+\lambda} + \frac{\lambda}{\mu+\lambda} e^{-(\mu+\lambda)t}.$$

Further, if one let $t \to \infty$, $e^{-(\mu+\lambda)t} \to 0$. As a consequence, one is left with a simpler expression.

$$p_{11}(t) = \frac{\mu}{\mu+\lambda}.$$

Further, it is known that $\mu = \frac{1}{MTTR}$ and that $\lambda = \frac{1}{MTTF}$. Inserting this for μ and λ in the expression yields the limiting availability formula.

$$p_{11}(t) = \frac{\frac{1}{MTTR}}{\frac{1}{MTTR} + \frac{1}{MTTF}}$$

$$p_{11}(t) = \frac{MTTF}{MTTF + MTTR}.$$