

Master's Programme in Advanced Energy Solutions

Reliability and remaining useful life estimation of power plant components

Markus Kolehmainen

Master's Thesis
2021

Copyright ©2021 Markus Kolehmainen

Author Markus Kolehmainen

Title of thesis Reliability and remaining useful life estimation of power plant components

Programme Advanced Energy Solutions

Major Sustainable Energy Systems and Markets

Thesis supervisor Prof. Matti Lehtonen

Thesis advisor M.Sc. (Tech) Marko Mäkinen

Collaborative partner Fortum Power and Heat Oy

Date 31.8.2021	Number of pages 121 + 5	Language English
-----------------------	--------------------------------	-------------------------

Abstract

Reliabilities and remaining useful lives (RULs) are estimated to determine optimal intervention strategies and thus minimize total cost of ownership for physical assets. Reliability centered maintenance is a structured analysis method which can be used to identify the optimal strategies. Object of the literature review was to study failure modes of various power plant components and review how reliability and RUL can be estimated.

Reliability and RUL models can be divided to four different categories: physics of failure models, statistical models, hybrid models and artificial intelligence models. Data availability largely determines the choice of model. Different types of input data are event data, continuous condition monitoring and periodical inspections. Physics of failure models can be built even without data but it requires deep expertise of the specific device.

Object of the empirical part was to develop a reliability model for programmable logic controller and human machine interface. Firstly, failure mode and effect analysis was performed to identify all potential failure modes. 27 different failure modes were recognized, and comprehensive event tree models were outlined based on those. Realistically achievable level of detail in the models depends on the failure data's level of detail.

Event data for the models was acquired from work orders. Permanent failures could not be distinguished from intermittent failures, which caused difficulties to the modeling. Different modeling approaches were tested and compared. Conventional event tree approach is sufficient for typical failure rates, but stochastic modeling may be useful for devices suffering from frequent failures. Initiating event frequency can be estimated conditioned on the device age or time from the last failure. Time from the last failure is more informative as many of the failures are recurrent. Furthermore, the effect of manufacturer was studied by regression analysis. The ideal model should consider both age and past failures as well as the manufacturer.

Keywords Reliability Centered Maintenance, Failure Mode and Effect Analysis, Remaining Useful Life, Programmable Logic Controller, Human Machine Interface

Tekijä Markus Kolehmainen

Työn nimi Voimalaitoskomponenttien luotettavuuden ja jäljellä olevan käyttöiän estimointi

Koulutusohjelma Advanced Energy Solutions

Pääaine Sustainable Energy Systems and Markets

Vastuuopettaja/valvoja Prof. Matti Lehtonen

Työn ohjaaja DI Marko Mäkinen

Yhteistyötaho Fortum Power and Heat Oy

Päivämäärä 31.8.2021 **Sivumäärä** 121 + 5 **Kieli** Englanti

Tiivistelmä

Luotettavuuksia ja jäljellä olevia käyttöiä arvioidaan optimaalisten interventiostrategioiden määrittämiseksi ja siten kuluvan käyttöomaisuuden kokonaiskustannusten minimoimiseksi. Luotettavuuskeskeinen kunnossapito on jäsenelty analyysimenetelmä, jota voidaan käyttää optimaalisten strategioiden tunnistamiseksi. Kirjallisuuskatsauksen tavoitteena oli tutkia eri voimalaitoskomponenttien vikaantumistapoja ja tarkastella kuinka luotettavuutta ja jäljellä olevaa käyttöikää voidaan arvioida.

Luotettavuus- ja elinikämallit voidaan jakaa neljään kategoriaan: vikaantumisen fysiikkaan perustuvat mallit, tilastolliset mallit, hybridimallit ja tekoälymallit. Käytettävissä oleva data määrittää suurelta osin mallin valinnan. Syötedatan eri tyypit ovat tapahtumadata, jatkuva kunnonvalvonta ja määrääikaistarkastukset. Vikaantumisen fysiikkaan perustuvia malleja voidaan rakentaa jopa ilman dataa, mutta se vaatii syvällistä asiantuntemusta kyseisestä laitteesta.

Työn empiirisen osan tavoitteena oli kehittää luotettavuusmalli ohjelmoitavalle logiikalle sekä ihmisen ja koneen väliselle käyttöliittymälle. Ensimmäinen tehtiin vikatila- ja vaikutusanalyysi kaikkien potentiaalisten vikaantumistapojen tunnistamiseksi. 27 eri vikaantumistapaa tunnistettiin, ja niiden perusteella luonnosteltiin kattavat tapahtumapuumallit. Realistisesti saavutettavissa oleva mallien yksityiskohtaisuus riippuu vikatioiden yksityiskohtaisuudesta.

Tapahtumadata mallinnusta varten saatiin työmääräimistä. Pysyviä vikoja ei voitu erottaa ajoittaisista vioista, mikä aiheutti vaikeuksia mallinnuksessa. Työssä kokeiltiin ja vertailtiin erilaisia mallinnustapoja. Tavanomainen tapahtumapuun mallinnustapa on riittävä tyypillisille vikataajuuksille, mutta stokastinen mallinnustapa voi olla hyödyllinen usein vikaantuville laitteille. Alkutapahtumataajuuksia voidaan arvioida joko laitteen iän tai ajan viimeisimmästä viasta perusteella. Aika viimeisimmästä viasta on informatiivisempi, koska monet vioista ovat toistuvia. Lisäksi valmistajan vaikutusta tutkittiin regressioanalyysillä. Ideaalisesti mallin tulisi ottaa huomioon sekä ikä ja aiemmat vikaantumiset että valmistaja.

Avainsanat Luotettavuuskeskeinen kunnossapito, vika- ja vaikutusanalyysi, jäljellä oleva käyttöikä, ohjelmoitava logiikka, ihmisen ja koneen välinen käyttöliittymä

Table of Contents

1	Introduction.....	11
2	Power plant components.....	14
2.1	Thermal power components	14
2.1.1	Fuel handling.....	15
2.1.2	Water & steam system	16
2.1.3	Boiler.....	17
2.1.4	Flue gas cleaning.....	19
2.2	Turbine	22
2.2.1	Steam turbine	23
2.2.2	Gas turbine.....	24
2.2.3	Hydro turbine	25
2.2.4	Wind turbine.....	26
2.3	Gearbox.....	29
2.4	Generator.....	30
2.4.1	Turbogenerator.....	30
2.4.2	Salient pole generator.....	31
2.4.3	Wind power generators	32
2.5	Balance of plant	35
2.5.1	Transformer.....	35
2.5.2	Switchgear.....	37
2.5.3	Automation	38
3	Reliability centered maintenance.....	43
3.1	Reliability theory	44
3.1.1	Reliability metrics and distributions.....	45
3.1.2	Failure patterns	47
3.2	Failure mode and effect analysis.....	51
3.2.1	Functions	51
3.2.2	Failure modes	52
3.2.3	Failure effects.....	53
3.3	Criticality analysis	53
3.4	Maintenance strategies.....	55
3.4.1	Corrective maintenance.....	58
3.4.2	Preventive maintenance	59

3.5	Default strategies.....	61
4	Reliability and remaining useful life estimation.....	62
4.1	Physics of failure models.....	64
4.2	Statistical models.....	65
4.2.1	Non-parametric models.....	67
4.2.2	Parametric models.....	68
4.2.3	Regression models.....	73
4.2.4	Stochastic models.....	74
4.3	Hybrid models.....	77
4.4	Artificial intelligence models.....	78
5	Case Study: Programmable logic controller and human machine interface.....	80
5.1	Failure mode and effect analysis.....	80
5.1.1	Software.....	85
5.1.2	Manufacturer.....	87
5.1.3	Power supply module and chassis.....	88
5.1.4	Input and output modules.....	90
5.1.5	Processor module and human machine interface.....	92
5.2	Literature and manufacturer failure data.....	95
5.3	Data acquisition.....	99
5.4	Survival analysis.....	100
5.5	Survival regression.....	101
5.6	Event tree models.....	102
5.6.1	Conventional event tree model.....	102
5.6.2	Stochastic event tree model.....	103
5.6.3	Initiating event frequency estimation.....	104
5.6.4	Model comparison.....	105
6	Results, findings and discussion.....	108
7	Conclusion.....	111
A.	Appendix – Digital instrumentation and control failure modes.....	122
B.	Appendix – Comprehensive models.....	124

Preface

I hope that this thesis helps to take a small step forward on the long journey towards more objective data-based decision making and ultimately for a cleaner world.

Thanks for this opportunity and thanks to all who have supported me, especially to my advisor Marko Mäkinen and supervisor Prof. Matti Lehtonen.

Helsinki, 31 August 2021

Markus Kolehmainen

Symbols, operators and abbreviations

Symbols

\bar{A}	Column vector consisting of regression parameters
\bar{X}	Row vector consisting of explanatory variables
d_i	Number of events at time t_i
n_i	Number of units at risk at time t_i
T	Time to an event (random variable)
T_{50}	Median time to an event
t_i	Time point i
β	Shape parameter
γ	Location parameter
Δt	Time interval
η	Scale parameter
λ	Rate parameter
μ	Mean
σ	Standard deviation

Operators

$B(t)$	Standard Brownian motion
$E(X Y)$	Expected value of X given Y
$f(x)$	Probability density function of random variable X
$F(x)$	Cumulative distribution function of random variable X
$h(t)$	Hazard function of random variable T
$H(t)$	Cumulative hazard function of random variable T
$L(t)$	Expected time to event given that event has not occurred prior to time t
$P(X Y)$	Probability of X given Y
$S(t)$	Survival function of random variable T
$X(t)$	Degradation level at time t
$\Phi(x)$	Standard normal cumulative density function
$\prod_{i \leq t} x_i$	Product of an above bounded sequence
$\sum_{k=1}^n t_k$	Sum of t_k , from $k = 1$ to n

Abbreviations

AC	Alternating Current
AFR	Average Failure Rate
AI	Artificial Intelligence
ALARP	As Low As Reasonably Practicable
ARIMAX	Autoregressive Integrated Moving Average model with an Exogenous variable
BoP	Balance of Plant
CCGT	Combined Cycle Gas Turbine
cdf	cumulative density function
CPU	Central Processing Unit
DC	Direct Current
DFIG	Doubly-Fed Induction Generator
EESG	Electrically Excited Synchronous Generator
EHS	Environment, Health and Safety
F(I)MEA	Failure (Intrusion) Modes and Effects Analysis
FIT	Failures Per 10 ⁹ Hours
FMEA	Failure Mode and Effect Analysis
FMEDA	Failure Modes, Effects and Diagnostic Analysis
GSU	Generator Step-Up Transformer
HAWT	Horizontal-Axis Wind Turbine
HGS	Hydroelectric Generator System
HMI	Human Machine Interface
HRSG	Heat Recovery Steam Generator
HV	High Voltage
I&C	Instrumentation and Control
I/O	Input/Output
IC	Integrated Circuit
LED	Light Emitting Diode
MLE	Maximum Likelihood Estimation
MTBF	Mean Time Between Failures
MTTF	Mean Time To Failure
MTTR	Mean Time To Repair
NASA	National Aeronautics and Space Administration
O&M	Operations and Maintenance
PC	Personal Computer

PCA	Principal Component Analysis
pdf	probability density function
P-F	Potential-to-Functional Failure
PLC	Programmable Logic Controller
PM	Particulate Matter
PMSG	Permanent Magnet Synchronous Generator
PoF	Probability of Failure
PRA	Probabilistic Risk Assessment
RAM	Random-Access Memory
RCM	Reliability Centered Maintenance
ROM	Read-Only Memory
RUL	Remaining Useful Life
SAE	Society of Automotive Engineers
SCADA	Supervisory Control And Data Acquisition
SCR	Selective Catalytic Reduction
SNCR	Selective Non-Catalytic Reduction
U.S. MSDP	U.S. Maintenance System Development Program
U.S. SSMD	U.S. SIMA Submarine Maintenance Division
U.S. SUBMEPP	U.S. Submarine Maintenance Engineering, Planning and Procurement
U.S. DoD	U.S. Department of Defense
U.S. NIST	U.S. National Institute of Standards and Technology
U.S. NWSC	U.S. Naval Surface Warfare Center
UK HSE	UK Health and Safety Executive
UK OFGEM	UK Office of Gas and Electricity Markets
USBR	U.S. Bureau of Reclamation
WTG	Wind Turbine Generator

1 Introduction

Energy industry is a capital intensive industry involving extensive amount of physical assets subject to degradation. As for every capital intensive industry, asset reliability is a vital part of cost-efficient operation also for power plants. In the pre-liberalized electricity markets the monopoly position ensured profitability. Afterwards, the market has been liberalized to introduce competition and improve efficiency. A price-taker must ensure competitiveness by reducing production costs. In addition to direct costs resulting from spare parts and repairs; low reliability increases also lost production as well as insurance and financing expenses due to increased uncertainties.

Electricity is vital for modern societies, and thus reliable power production is desirable also from societal perspective, although transmission and distribution networks are more critical regarding to blackouts. Customer interruption cost can be 10-100 times higher than the price of delivered electricity when all direct and indirect consequences are considered (Gündüz et al. 2018). Reliability improvements can be pursued from two different viewpoints: manufacturer can affect reliability through design, while owner or operator can improve the reliability through asset management practices. This thesis approaches the topic from owner's and operator's perspective.

Well planned and implemented asset management can extend asset's lifetime and reduce downtime. In order to enable asset's life cycle cost optimization, reliability and remaining useful lifetime (RUL) must be estimated. By knowing reliability and RUL estimates, asset interventions such as maintenance and replacements can be optimized. Recently, with the increasing availability of data and computational power, there has been increasing academic and industrial interest towards the reliability and RUL estimation (Lei et al. 2018).

This thesis consists of two distinguishable parts which have their own research objectives and methods. Chapters from 2 to 4 comprise theoretical part of the study. Main objective of the theoretical part is to provide comprehensive yet concise overview of power plant components' reliability and RUL estimation. Interim objectives of the theoretical part are giving adequate background knowledge to support the main objective and presenting how the main objective is related to other topics in the field. Research method used in the theoretical part is literature study.

Chapters 5 and 6 comprise empirical part of the study. Empirical part consist of a case study of programmable logic controller (PLC) and human machine interface (HMI). Objective of the empirical part is to implement reliability model by applying theories presented in the literature while using empirical data. Used data is primary and descriptive in the sense that it has not been earlier used for research purposes and it is collected from actual operating environments without intervening.

Another way to distinguish the thesis is type of sources used and its relation to knowledge. Well established basic theories are presented for each subject. To cover the latest advances in the field, more recent yet possibly controversial theories are studied. Also commercial sources e.g. component manufacturers are utilized. Commercial sources can be considered less trustworthy than peer reviewed academic journals or books since information obtained from commercial parties may be influenced by financial interests. However, manufacturers' information is useful when studying specific components, and using different sources allows for comparison. Empirical part of the study contributes to knowledge by applying existing scientific theories to a new case and new data.

The rest of this thesis is organized as follows. Chapter 2 provides basic knowledge of selected power plant components including working principles, expected lifetimes and the most common failure mechanisms. Selected components are main components of steam, gas, hydro and wind power plants. The types of power plants covered are chosen based on collaborative partner's interests and power generation portfolio. In Chapter 3 reliability centered maintenance (RCM) framework is studied. By introducing RCM, concept of reliability estimation is placed in wider context of maintenance optimization. Maintenance optimization, in turn, relates to even higher level concept of asset management. Since business decisions are evaluated in monetary terms, reliability metrics must be converted to asset's cash flow generating capacity to support the decision making. However, actual investment analysis is out of the scope of this thesis. Aforementioned chapters give background knowledge of the main topic and its relations to closely linked subjects. Chapter 4 is a literature review of reliability and RUL estimation. All of the most used estimation methods are presented, and also some recent advances are introduced. Chapter 4 provides theoretical framework to following empirical part of the thesis.

Chapter 5 includes a case study of PLC and HMI. Reliability estimates are obtained from the literature and data collected from collaborative partner's power plants. Case study includes modelling process from data acquisition and preparation to model construction and testing. Chapter 6 discusses the results and findings observed during the study. Also perceived limitations and possibilities for further studies are discussed. Chapter 7 provides a conclusion. The thesis is structured in a logical sequence such that earlier chapter provides prerequisites and context for the following chapter.

2 Power plant components

This chapter introduces main components of steam, gas, hydro and wind power plants to provide necessary prerequisite knowledge required for reliability and RUL modelling. Considered topics include operating principles, expected lifetimes and typical failure modes. In current operating environment also effect of load cycling is an important point to consider. Cyclic use exposes components to unavoidable thermal and pressure stresses. This causes damage especially for high temperature components through creep-fatigue interaction. Hence the cost of cycling should be known when optimizing power plant operation as it is recognized to increase maintenance costs and forced outages. (Kumar et al. 2012) Since load cycling related failures do not occur immediately, real consequences will be known only after enough experience is gained from grids with large share of variable renewable generation.

The chapter is structured by component groups rather than power plants types. Therefore, components common to all can be reviewed at once. It is noteworthy that all main components of modern power plants include automation. However, in this chapter automation is discussed in a separate subchapter dedicated to it.

2.1 Thermal power components

Thermal power production comprises technologies in which heat energy is converted to electricity. Source of the heat can be combustion of a fuel; or nuclear, solar or geothermal energy. However, only solid biomass, municipal waste, coal and gas fired plants are included in this thesis. Solid biomass can be e.g. wood chips, pellets or agricultural wastes like straw. In addition to heat source, another way to distinguish thermal power production technologies is the utilized thermodynamic cycle. Usually gas fired plants try to follow the Brayton cycle, while plants powered by other fuels utilize the Rankine cycle. Combined cycle gas turbine (CCGT) power plants utilize both cycles to achieve maximal electrical efficiency. Working fluid of typical open-loop Brayton cycle is air, and working fluid of the Rankine cycle is water.

This subchapter introduces main components of thermal power production. Main components of the Rankine cycle power plant are feedwater pump, boiler, steam turbine and condenser. Other essential components are those related to transferring the heat into the

working cycle e.g. fuel handling. Main components of the open-cycle Brayton cycle power plant are compressor, burner and gas turbine, which are discussed in 2.2.2. Flue gas cleaning is discussed jointly for both cycles in this subchapter. Also reciprocating engines can be used in power production. However, they are much less common and are therefore not discussed in this work.

Creep-fatigue damage is the dominant failure mechanism for many thermal power production components (Kumar et al. 2012). Creep is slowly occurring permanent deformation of a solid material under persistent mechanical stress, and it accelerates with temperature. Creep occurs also under constant stress and temperature (Zhang et al. 2017). Creep-fatigue is a failure mechanism caused by interaction of creep and transients resulting from cyclic use. Damage from it typically occurs first at constrained locations exposed to thermal transients. Examples of constrained locations are joining of thick to thin section and joining of materials with different expansion coefficients. (Kumar et al. 2012)

2.1.1 Fuel handling

Fuel handling refers to all equipment required to make the fuel available for combustion. The equipment needed depends on the type of fuel used. Solid fuels require conveyors, liquid fuels use pumps and gaseous fuels require compressors. In addition to fuel transportation, fuel handling can include fuel comminution, storages and also fuel heating in case of viscous liquid fuels like heavy fuel oil. Comminution is a process where particle size of a solid material is reduced. Decreased particle size leads to improved ignition and combustion properties. Hence coal is typically pulverized before combustion, and biomass can be milled. Wood pellets can be milled in coal mills, though grinding energy consumption is increased (Masche et al. 2017). Cyclic operation increases iron wear rates of mills as a result of occasional low fuel flows (Kumar et al. 2012).

Unlike coal, wood pellets must be stored inside large halls to prevent them from absorbing moisture. Torrefaction is a relatively new process in which combustion and storability properties of biomass are improved. Torrefied biomass has higher energy density, more homogenous composition, better grindability and hydrophobic behavior. It does not absorb moisture like traditional solid biofuels, and it can hence be stored under open sky. (Mandø 2013)

Solid fuel conveying equipment includes belt, chain, screw and pneumatic conveyors. Belt conveyors are an economical choice for longer distances, while screw conveyors are more suitable for shorter distances. Pneumatic transportation is used for pulverized fuels. Also cranes are used especially in waste incineration plants. Waste is the most challenging fuel as it is not homogenous and can contain unwanted or even dangerous objects. (Mikus et al. 2016)

2.1.2 Water & steam system

Water & steam system of a Rankine cycle power plant includes its main components feedwater pump, boiler, steam turbine, condenser and piping connecting the components as shown in Figure 1. This section includes feedwater pump, condenser and piping, while rest of the main components are discussed in own sections.

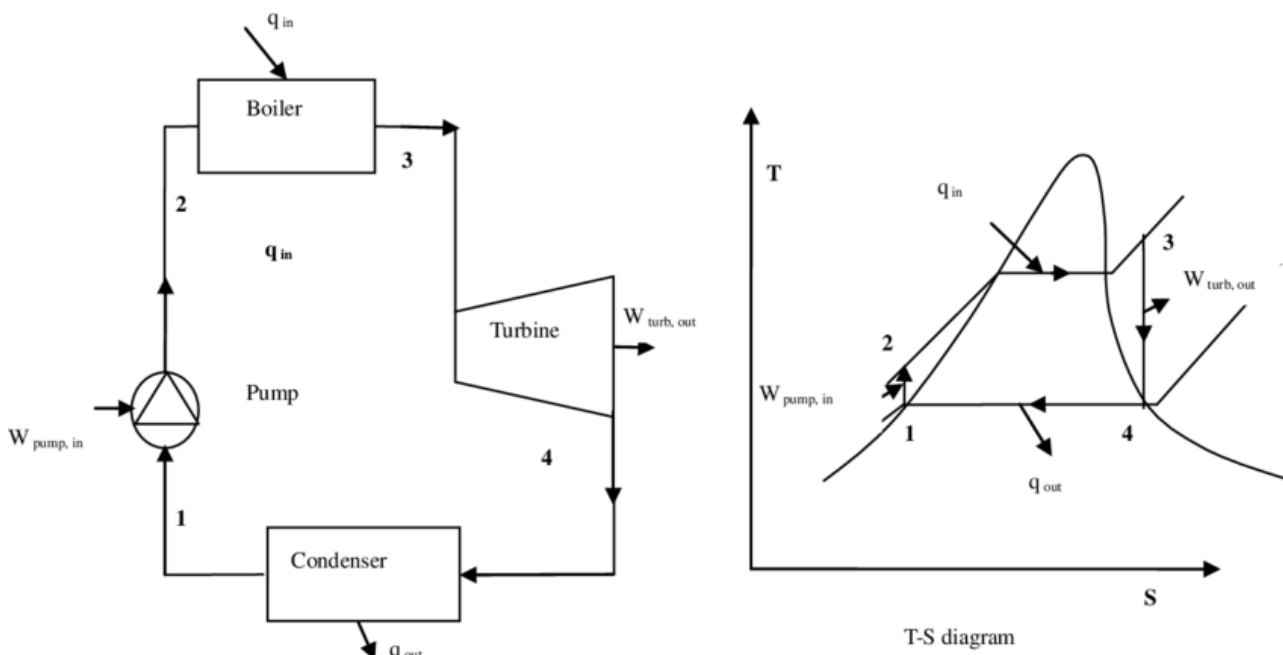


Figure 1. Main components and idealized T-S diagram of a simple Rankine cycle. (Vundela et al. 2010)

Feedwater pump supplies high-pressure water to the boiler (transition from point 1 to point 2 in Figure 1). In the boiler water turns into steam and gets superheated (2→3). Then the superheated steam expands in the turbine and work is extracted (3→4). After the turbine steam goes to a condenser, where it condenses back to liquid water (4→1). In reheating cycles there is many turbines, and the steam is reheated between the turbines. Reheating can be used to improve efficiency of the plant.

Feedwater pumps are typically centrifugal pumps, where the fluid enters the pump near to the rotating axis and is accelerated by the impeller as it flows radially to the volute. A centrifugal pump consists of an impeller, shaft, volute casing, stuffing box and bearings. The stuffing box prevents the pump from leaking at the location where the shaft goes through the casing. Three main factors affecting reliability of a centrifugal pump are operating speed, impeller diameter and flow rate. Typical failure mechanisms are cavitation, vortexing, interference, corrosion, erosion, fatigue and bearing failure. (U.S. NWSC 2011)

In combined heat and power plants the latent heat of condensation is transferred to the district heating water, and in condensing power plants the heat is rejected to a body of water or atmosphere. Since condensers of power only plants operate below atmospheric pressure, leakage will result in ingress of contaminants. Many of the corrosion problems in boilers and low pressure turbines have been traced to cooling water ingress occurring in the condenser. (Sinha 2010) Another possible failure is clogging, which is mainly caused by fouling of calcium compounds. Initially, there is a trade-off relation between fouling and flow induced erosion and corrosion. Low flow velocities cause fouling, but high velocities accelerate erosion. However, after fouling has occurred, it can cause local high velocities and induce erosion. (Ranjbar 2010) Other failure mechanisms of heat exchangers are fatigue and vibration (Addepalli et al. 2015).

Function of the piping is to contain and transfer the working fluid between components. Critical part of the piping is the section between the boiler and turbine which contains superheated steam. It is exposed to creep-fatigue comparable to the superheater header and the turbine casing.

2.1.3 Boiler

Boiler is a device used to produce and superheat steam. Modern industrial scale boilers are typically high pressure water-tube boilers, in which water-filled tubes make up the walls of the furnace. Different combustion technologies used for solid fuels are grate combustion, pulverized combustion and fluidized-bed combustion. Fluidized-bed consists typically of sand particles, and the bed is brought into a fluid-like state by blowing air from below the bed. Grate combustion is the most common technology in municipal waste incineration, and pulverized combustion dominates coal-fired plants. All of the three technologies are utilized

in combustion of biomass. (Mandø 2013) Also other techniques, like rotary kilns used for hazardous waste incineration, exist but they are not further discussed in this thesis.

Pulverized combustion has the best efficiency but it requires high-quality dry fuel to work. Grate-fired boilers are capable of burning almost any type of solid fuel at the expense of lower efficiency. Fluidized-bed is the middle option regarding fuel flexibility and efficiency. The problems of fluidized-bed combustion are erosion of surfaces in contact with the bed material and agglomeration. Biomass is more challenging fuel for the boiler compared to coal. It is more corrosive, it causes more fouling to the furnace surfaces and the ash from biomass is more prone to agglomeration. The problems are exacerbated when using herbaceous biomasses i.e. plants that have a nonwoody stem. (Mandø 2013)

Boiler degradation mechanisms are creep, thermal fatigue, fireside corrosion and erosion, and steam side oxidation and spallation (Buhre et al. 2002). Thermal efficiency of Rankine cycle improves with increasing turbine inlet steam temperatures, but degradation accelerates with temperature. Fireside corrosion is the main life-limiting factor in pulverized combustion (Dudziak et al. 2016). Namely, corrosion rate of ferritic steels increases linearly with temperature and becomes unacceptable at about 550 °C. Austenitic steels are suitable for temperatures up to about 630 °C. Superheater and reheater mid-wall temperatures are 50–60 °C higher than the steam temperature. (Wheeldon & Shingledecker 2013) Figure 2 presents the relation between fireside corrosion rate and steam temperature.

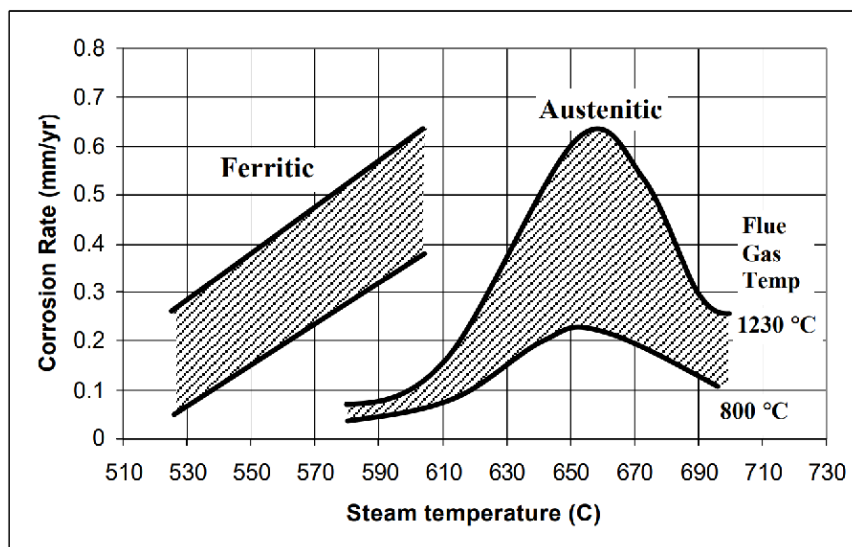


Figure 2. Fireside corrosion rates of ferritic and austenitic steels as a function of steam temperature. (Buhre et al. 2002)

The bell-shaped curve of austenitic steels results from behavior of the ash deposit on the tube surface. High corrosion rates occur when the ash deposit turns molten and layer of molten alkali-metal-trisulfates accelerates corrosion. At lower temperatures the deposit is dry, and at higher temperatures the molten sulfates vaporize. (Buhre et al. 2002) Even though the fireside corrosion of austenitic steels decelerates at temperatures above 700 °C, other properties like high thermal expansion makes them unsuitable for such temperatures. Hence, nickel-based alloys familiar from gas turbines must be used in superheaters and reheaters of advanced ultra-supercritical boilers. Costs of austenitic stainless steels and nickel-based alloys are 3–10 and 40–50 times higher than the price of low-alloy ferritic steel. Different parts of the boiler are exposed to different temperatures and can be made of different materials given that thermal expansion coefficients match. (Wheeldon & Shingledecker 2013) That is to say, water walls can be made of lower grade material than superheat and reheat tubing.

Superheater header is recognized to be one of the most sensitive parts of a thermal power plant as it is exposed to high temperatures and pressures. Cracks may form especially to connections between the superheater tubes and the reservoir. Torshizi & Jahangiri (2018) studied creep-fatigue crack growth of a superheater header. According to their study remaining lifetime of a header with 1 mm crack is about 13 years including 420 cycles. With this cycling rate fatigue is negligible compared to creep.

2.1.4 Flue gas cleaning

Flue gas cleaning comprises technologies used to remove pollutants from the flue gas. Although biomass can be seen as a carbon neutral fuel, the primary pollutants resulting from biomass combustion are worse than those produced by combustion of gas and comparable to those produced by combustion of coal (Mandø 2013). The main primary pollutants are particulate matter (PM), products of incomplete combustion (CO and hydrocarbons), nitrogen oxides (NO_x), sulfur oxides (SO_x) and heavy metals (Hg, Cd, Tl, As, Ni, Pb). Secondary pollutants, such as ground-level ozone and acid rain, are formed when primary pollutants react in the atmosphere. Fine particulate matter (PM_{2.5}) is the primary air pollutant that causes the most mortality (Singh & Shukla 2014). There has been a trend towards more and more strict regulations regarding flue gas pollutants. According to Mikus et al. (2016), a good method to predict upcoming EU regulations is to observe changes in German laws, which are often adopted by the EU as standards for the member countries.

In waste incineration, the emission control should begin already before the combustion by removing dangerous waste. Characteristics of the combustion process affect especially NO_x formation regardless of the fuel type. After the combustion there is the actual flue gas cleaning systems. This section introduces common methods used for PM, NO_x and SO_x reduction.

Reduction of particulate matter and particle-bounded heavy metals

Different methods for removal of particles are cyclones, electrostatic precipitators and filters. Optimal choice of technique depends on variety of issues including particle load in the flue gas, average particle size, size distribution, flow-rate, temperature, compatibility with surrounding equipment and required limits. Cyclones utilize centrifugal force to remove particles and are suitable only for larger particle sizes. Electrostatic precipitators ionize particles and then attract them to collector plates with an opposite electric charge. Voltage levels of electrostatic precipitators are 20-100 kV. (Mikus et al. 2016) They can reach 99 % efficiency for particles 1 to 10 µm in size (Miller, 2005). This reduction may not be enough and fabric filters are thus used in series with electrostatic precipitators. Fabric filters are very efficient but precleaning the flue gas before the fabric filters prolongs their lifetime. Also surface filters can be used to filter coarse particles before the depth-loaded filter. Depth-loaded filter is capable of filtering fine particles by stopping them inside the medium. Coarse particles can be cleaned off but pressure drop resulting from accumulation of fine particles can be managed only by replacing the filters. (Mikus et al. 2016)

Reduction of nitrogen oxides

NO_x are formed in two different ways during combustion. Thermal NO_x can be formed from nitrogen present in the combustion air when temperature is above 1000 °C. Fuel NO_x can be formed when nitrogen in the fuel oxidizes via radical reaction. Primary technique to control NO_x emissions is preventing the formation. This can be done by e.g. controlling temperature and air supply, flue gas recirculation, oxygen injection, natural gas injection or water injection. (Mikus et al. 2016)

If the primary techniques are not sufficient, secondary techniques must be used. Secondary flue gas NO_x reduction methods are selective catalytic reduction (SCR) and selective non-catalytic reduction (SNCR). SNCR processes reduce NO_x by injecting ammonia water or urea into the flue gas. Reduction efficiency of the SNCR methods is 60-80 %, and it is limited by

ammonia slip emissions. (Mikus et al. 2016) Expected lifetime of a SNCR equipment is 20 years (Sorrels et al. 2019). SCR is the most efficient NO_x reduction method with efficiencies up to 95 %. In SCR the flue gas passes over a catalyst after the ammonia water or urea injection. However, SCR system is 3-5 times more expensive than SNCR, and it requires more space. (Mikus et al. 2016) Expected lifetime of a SCR equipment is 30 years. Vendor-guaranteed catalyst lifetime is typically 3 years and actual lifetimes are often in 5-7 years range. (Sorrels et al. 2019)

Lifetime of a NO_x reduction system is affected by the flue gas temperature and pollutant content. SNCR must operate at temperatures between 900-1100 °C to work optimally, and it has thus lower expected lifetime. SCR operates at temperatures between 180-450 °C. Figure 3 presents different alternatives in which the SCR can be placed in a flue gas cleaning system. In location (a) the SCR is exposed to high dust and high SO_x flue gas which deteriorates the SCR equipment and catalyst. On the contrary, in location (d) the flue gas has low particle and sulfur content but is too cool for the catalytic reduction and requires reheating. However, location (d) is the most common location in waste incineration. (Mikus et al. 2016)

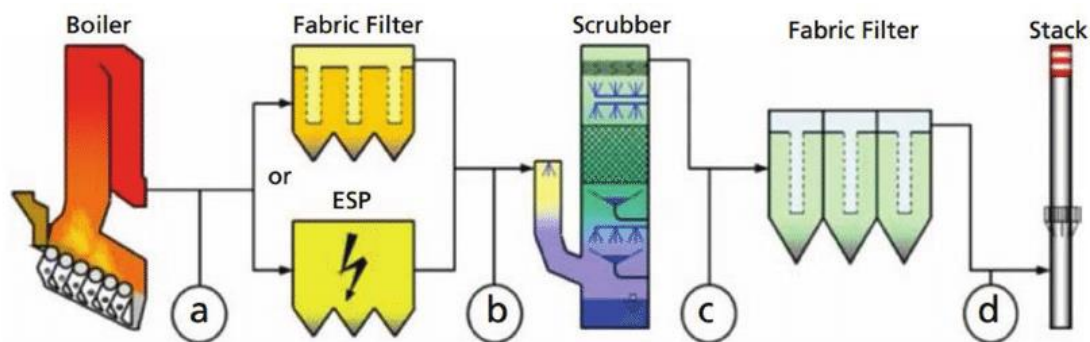


Figure 3. Four optional locations for SCR. Tail end is the best location regarding SCR lifetime but requires flue gas reheating. (Mikus et al. 2016)

Reduction of sulfur oxides

Desulphurization techniques can be divided to wet, semi-dry and dry methods. Wet calcium method is the most common method. It is also an efficient method reaching desulphurization efficiencies of over 90 %. Semi-dry and dry methods are less common and less efficient. In wet calcium method SO_x is reduced in a spraying counter-current absorber reactor, where the flue gas is sprayed with scrubbing liquid. The scrubbing liquid is water containing finely ground calcium compounds like limestone. The reactor must be made of

materials resistant to corrosion and abrasion. Different reactor types are spray towers, venturi scrubbers, plate towers and mobile packed beds. Simple design such as spray tower is best in avoiding problems due to scale buildup, plugging and erosion. The process generates slurry which must be disposed properly. However, when using limestone, the process can be implemented such that it produces gypsum, which is a saleable product. (Zagala & Abdelaal 2017)

2.2 Turbine

Turbine is a rotating machine that extracts energy from a fluid and converts it to rotational energy of a shaft. In power generation the shaft is connected to a generator, which is discussed in next subsection. The working fluid can be compressible or incompressible, and it contains energy as potential and kinetic energies. Nozzles are used to convert potential energy to kinetic energy and to direct the fluid. (Chaplin 2009) The turbine is designed based on characteristics of the working fluid. However, all turbines share certain common characteristics. All turbines have rotor, which comprises a shaft with blades attached to it. There can be one or multiple stages of blades. Compressible fluids require multiple turbine stages to extract maximal amount of useful work. Hence, steam and gas turbines have many stages. On the contrary, water is almost incompressible, and hydro turbines have only one turbine stage. Bearings support the rotor and allow it to rotate with low friction. Significant failure mechanisms of a journal bearing are abrasive and adhesive wear, plastic deformation and indentation, and non-uniform wear causing geometrical variations. (Muzakkir et al. 2015) Most turbines have a casing to contain the fluid and fixed stator blades attached to it acting as nozzles. In contrast to other turbines, wind turbine operates at ambient pressure and does not have casing or nozzles.

Turbines can be divided to two different types depending on the physical principle utilized to convert energy. The principles and thus turbines can be distinguished by location where the pressure drop occurs. In impulse turbine pressure drop occurs only during stator stage and pressure stays constant when it passes through the rotor stage. (Chaplin 2009) Stator converts fluid's potential energy to kinetic energy. Accelerated fluid then decelerates and changes direction when it impinges the rotor blades. Kinetic energy of the fluid is converted to rotational energy of the rotor based on *impulse* principle.

In reaction turbines pressure drop occurs during both stator and rotor stages. Rotor stage acts as a nozzle which accelerates the fluid. Accelerating fluid causes force in the opposite direction and this jet *reaction* drives the turbine. However, in practice there is always also impulse effect occurring. (Chaplin 2009) Nonetheless, this classification is useful to distinguish pure impulse turbines from reaction turbines, where both effects occur.

Modern multistage turbines often utilize both principles. For example, high pressure stages can be impulse type and low pressure stages reaction type. Degree of reaction can also vary within single blade e.g. tip of the blade is reaction type while root of the blade is more of an impulse type. (Chaplin 2009) Wind turbines use airfoils to create aerodynamic lift. Depending on angle of attack there is also impulse effect occurring. Hydro turbines can be most easily distinguished between the two types. Francis and Kaplan turbines are reaction type while Pelton turbine is impulse type.

2.2.1 Steam turbine

Steam turbines extract work from high-pressure superheated steam. Work is extracted as steam travels through multiple stages and expands to a low pressure. Most of the modern utility-scale steam turbines are multistage axial-turbines. Main types of steam turbines are back pressure turbines, which are used combined heat and power production, and condensing turbines which are used in power only production. A steam turbine consists of a rotor, stator, casing, bearings, valves, sealings and turning gear. Turning gear is needed during shutdown to prevent the shaft from bending. Efficiency of a steam turbine can be improved by increasing temperature and pressure difference over the turbine. However, more demanding conditions can have adverse effect on reliability. Inlet steam temperature has been typically about 540 °C as it is maximum temperature for conventional ferritic alloys. Austenitic alloys are problematic for rotating machines due to high thermal expansion. Best available ferritic alloys designed for ultra-supercritical steam cycles can be used at temperatures up to 620 °C. Higher temperatures require expensive nickel-based alloys. (Arrell 2006)

Common deterioration mechanisms of a steam turbine are fatigue, erosion, corrosion and creep. Blade root and disc joints are critical locations as they experience the most severe creep-fatigue. (Sakurai & Isobe 2010) Though, according to Ziegler et al. (2013) fatigue, stress corrosion cracking and corrosion fatigue of low pressure turbine blades are the most

significant failure mechanisms. First stages of high pressure turbine can suffer from solid particle erosion, and final stages of low pressure turbine suffer from water-droplet erosion. The solid particle erosion is typically caused by exfoliation of superheater or reheater tube material. (Nomoto 2017) The water-droplet formation could be prevented by keeping the steam still superheated in final stages, but it would decrease efficiency.

2.2.2 Gas turbine

Open-loop gas turbine is a turbine that uses atmospheric air as the working fluid. Also closed-loop gas turbine exist, but they are less common and are not discussed in this thesis. Main components of a gas turbine are a compressor, combustor and turbine which is connected to the same shaft as the compressor. In a simplified way, operation of a gas turbine consists of four steps. First, air flows through the compressor that compresses it to a high pressure (transition from point 1 to point 2 in Figure 4). After the compressor fuel is combusted with the pressurized air in the combustor ($2 \rightarrow 3$). Then resulting high-temperature pressurized gas flows through the turbine producing shaft work output ($3 \rightarrow 4$). As the turbine produces more power than the compressor consumes, rest can be used for external work such as driving a generator. The cycle gets closed when exhaust gases return to the ambient conditions ($4 \rightarrow 1$).

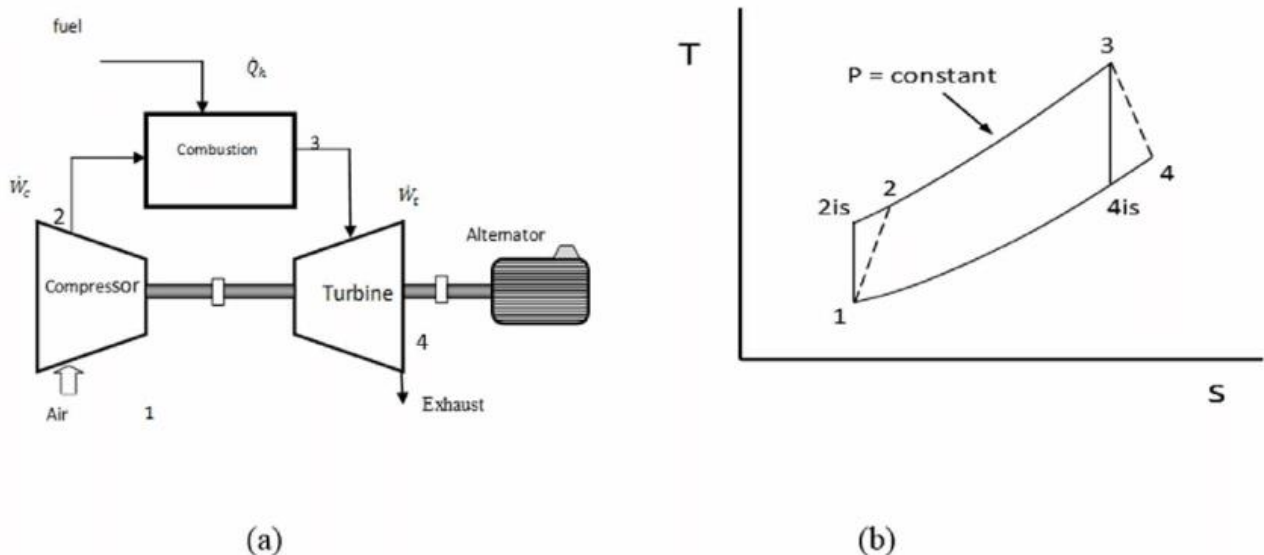


Figure 4. Simple Brayton gas cycle (a) components (b) T-S diagram of idealized and real cycles. Solid lines illustrate the ideal isentropic compression and expansion, while dashed lines describe the real irreversible process. (Abed et al. 2016)

In open-cycle systems the working fluid is rejected to ambient after the turbine. However, the flue gas has still considerable heat content after the turbine. Some of the heat can be

recovered with a combustion air preheater or a heat recovery steam generator (HRSG). In CCGT power plants HRSG is used to power a Rankine cycle with exhaust gases of a Brayton cycle.

Gas turbines operate in more hostile conditions compared to steam turbines since the temperature is higher and composition of the working fluid is less controllable. Efficiency of a gas turbine can be improved by increasing turbine inlet temperature, which is limited by properties of the materials exposed to the hot gas. Increased temperature leads to accelerated creep-fatigue and hot corrosion. As gas turbine works with ambient air, filtration is important to prevent foreign objects and impurities from entering the turbine. Foreign object erosion is a significant degradation mechanism for compressors in addition to creep-fatigue. (Yang & Xu 2011) Even though gas turbines are often meant to be load following plants, cyclic operation can cause excessive thermal fatigue damage especially to hot gas path components (Kumar et al. 2012).

2.2.3 Hydro turbine

Hydro turbines are devices that convert kinetic and potential energy of water into mechanical work. Three main types of hydro turbines are Kaplan, Francis and Pelton turbine. Competing designs exist since pressure and flow rate of the working fluid cannot be freely determined like in thermal power plants. Generally, reaction type Kaplan and Francis turbines are used for low head plants, and impulse type Pelton turbines are preferred in high head sites. Figure 5 illustrates how the water flow behaves in different turbines. Pelton turbine is a tangential flow turbine, and Kaplan turbine is an axial turbine, while Francis turbine combines radial and axial flows. Kaplan turbines have adjustable blade angles to operate efficiently at different heads. (Dorji & Ghomaschchi 2014) Kaplan turbine is the most common type in Finland as most of the sites have low heads.

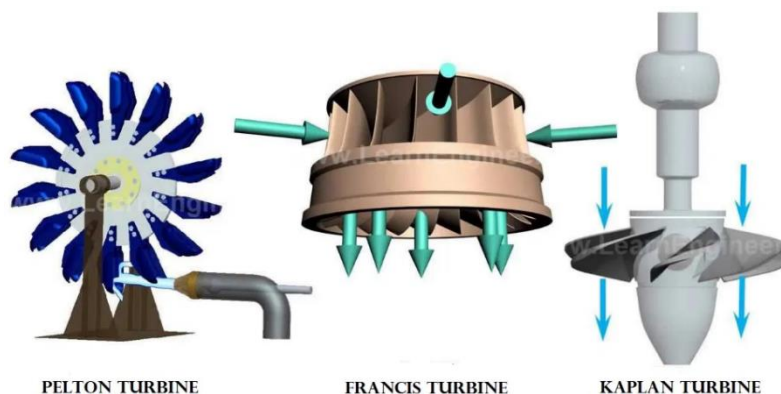


Figure 5. Main types of hydro turbines and their flow directions. (Filios 2013)

According to a large hydro power operator U.S. Bureau of Reclamation (USBR 2017), average service life of a hydro turbine runner is 50 years while dams are expected to have service lives of 100 years. However, runner wearing rings must be replaced every 20 years. Dorji & Ghomaschchi (2014) gave an overview of hydro turbine failure mechanisms. Identified three main failure mechanisms were cavitation, erosion and fatigue. A reaction turbine is most likely to fail due to cavitation while impulse turbine will most likely fail due to erosion. Cavitation refers to formation and collapse of vapour bubbles in liquids, which can cause extremely high local pressures up to 70 MPa. It occurs when local static pressure decreases below vapour pressure, and it first happens in locations where the flow velocity is highest. Typical locations for a Kaplan turbine are blades and guide vanes. Cavitation can cause surface penetration damage of up to 10 mm per year. In addition to design changes, methods to mitigate cavitation are online vibration monitoring and injection of air into the draft tube.

Erosion of the hydro power components occurs as a result of high velocity flow and abrasive sediments. Abrasive wear damages the flow guiding surfaces, which affects flow profiles and can initiate cavitation. Deteriorated flow profiles will also decrease plant efficiency. Leading and trailing edge of the blades are exposed to the maximum erosion. Methods to prevent the sediment erosion are erosion resistant coatings, de-silting chambers and ceasing operation during high sediment concentration. Fatigue failures occur mainly due to vibration and can therefore be prevented by monitoring the vibration level. The vibration can be hydraulic or caused by any component in the turbine-generator assembly. (Dorji & Ghomaschchi 2014) The hydraulic vibration can occur as a result of deteriorated flow surfaces, which means that the failure mechanisms are interrelated and can have accelerating effects on each other.

2.2.4 Wind turbine

Wind turbine is a device that converts kinetic energy of wind to rotational energy. Differently from other types of turbines, the term wind turbine often refers to whole construction needed to convert energy of the working fluid to electricity; comprising rotor, generator, surrounding structure and balance of plant. The surrounding structure includes nacelle, tower and foundation. Design lifetime of a wind turbine is typically 20-25 years, while structural components normally last way longer. (Letcher 2017, p. 299) This subchapter considers only the rotor and surrounding structures. Generator and balance of plant are discussed in separate sections.

Although many different designs exist, three-bladed horizontal-axis wind turbine (HAWT) is currently the only economically viable design. Alternative wind turbine designs are studied by e.g. Dilimulati et al. (2018) and are omitted in this thesis. At least until now, size and capacity of wind turbines has been increasing from kilowatt-scale of first commercial turbines to multi-megawatt-scale of current turbines.

However, at some point the square-cube law will limit the growth like has happened in the aircraft industry (Veers et al. 2019). When size of the turbine increases, rotor swept area increases quadratically but mass increases cubically. It is worth noting that in thermal production square-cube law advantages large units as there surface area is related to heat losses, and surface area to volume ratio is minimized with increasing unit size. Nonetheless, there are other factors favoring the growth like better wind conditions at higher elevations. Also advances in material technology may offset the increase in mass.

Wind turbine rotor consist of a hub and blades attached to it. The hub is connected to a drive train. The blades are made off structural composite materials, which are light yet have decent extreme strength and fatigue durability. The blades must withstand extreme weather events, fatigue loads from 20-25 years of operation as well as lightnings. (Letcher 2017, p. 299)

Modern wind turbines' blades can be rotated around their longitudinal axis. This feature is called blade pitch angle control, and it is used to control power. It is also used as a safety function to stop the turbine. Thus, there must be an energy storage to be able to stop the turbine during outage. The pitch control can be implemented with electric drives or hydraulics, and hence the energy storage can be a battery or a hydraulic accumulator. (Letcher 2017, p. 152-153) Electric drives are nowadays used more since they are more cost efficient. Downsides of the electric pitch are backlashes between gears and wear of the gears. Hydraulic pitch contains fewer components and damps structural vibrations. However, there is risk of leakage, and energy losses of hydraulic accumulator are larger than battery's. Pressure of the accumulator must be kept up to 200 bars to enable full-power emergency stop. (Böhmeke 2020)

Another important control system of a wind turbine is yaw system. The yaw system aligns the hub of the turbine against the wind and prevents it from moving. The system consists of

azimuth bearing, driving mechanism and yaw brakes. The azimuth bearing is typically double row sealed ball bearing, although also friction bearings can be used. The driving mechanism is similar to pitch drives. Electrical drives are favored while hydraulic implementation is also possible. Yaw brakes maintain the alignment and protect the drives from absorbing yaw loads. Gears of the drives should be pretensioned against each other after operation. If there is clearance between the teeth, the drives are exposed to harmful backlashes. (Böhmeke 2020)

Early wind turbines suffered from reliability issues. However, reliability has since been improving, and according to Letcher (2017) average availability of onshore turbines has already reached 98 % with MTBF being 7000 hours. The majority of wind turbine components have no redundancy, meaning that failure of a single component leads to a functional failure. Offshore turbines are about 7 % less reliable. Offshore wind turbines are characterized by remote locations, difficult access and being unmanned, which leads to increased downtimes and maintenance costs. One important factor affecting downtime resulting from a failure is whether a crane is needed for the repair. Consequently, O&M costs of offshore plants are twice the cost of onshore. (Letcher 2017, p. 299)

Numerous studies have been conducted on reliability of wind turbines. Pfaffel et al. (2017) made a review covering 15 of these studies. Figure 6 presents failure rates and mean down times obtained in 7 different studies considering onshore turbines. The review concluded that drive train failures are the largest contributor to down time due to long repair times. Electrical components cause more failures but less down time.

Large variations between different studies were found. Highest estimated total failure rate was 46.9 failure per year, which is over 100 times higher than lowest obtained estimate of 0.4 failures per year. The most important identified factor explaining this difference is significantly different definitions for what is considered a failure. University Nanjing's study counted remote resets as failures, which led to the highest failure rate yet lowest mean down time of 0.18 days per failure. The other extreme was to consider only events resulting to at least three days long outage, which led to mean down time of 5.42 days per failure. However, total yearly downtimes, which are obtained as the product of failure rate and mean down time, varied between 2.2 and 10.6 days per year. (Pfaffel et al. 2017) This is reasonable range and also in line with Letcher's (2017) availability estimate 98%.

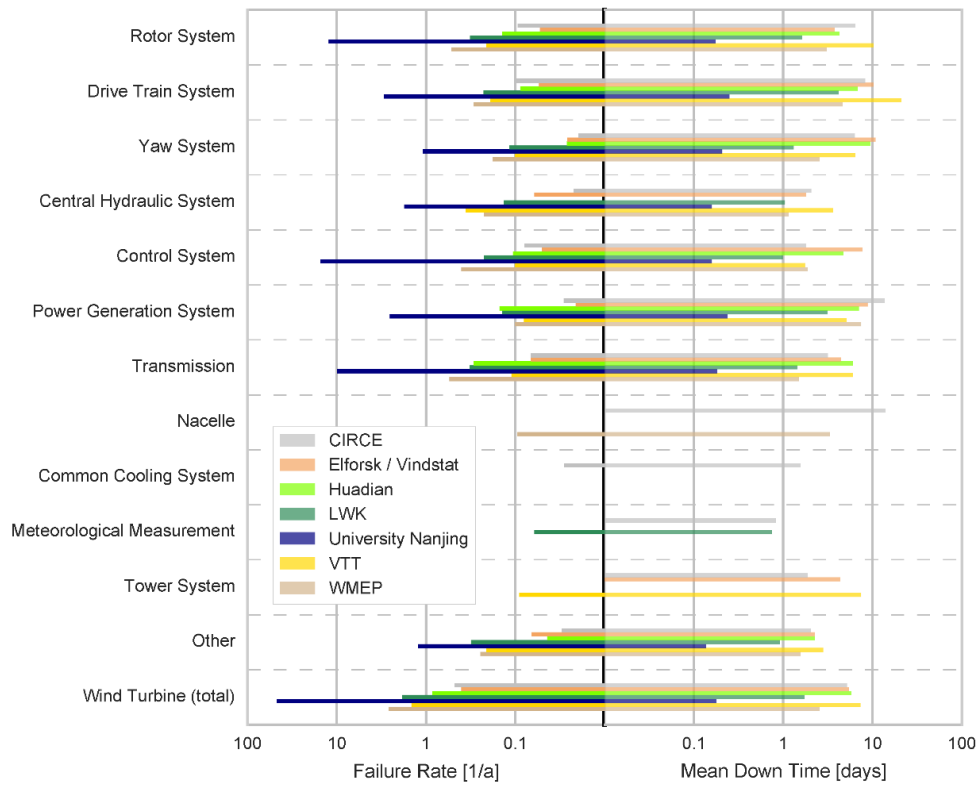


Figure 6. Wind turbine failure rates and mean down times from different studies. (Pfaffel et al. 2017)

2.3 Gearbox

A gear is a rotating device used to transmit torque and manipulate speed-torque relation. Gears are used in power generation to allow the turbine to rotate at optimal speed while the generator is tied to the grid frequency. Small steam turbines use reduction gears since they have optimal rotating speeds above the synchronous speed, and wind turbines use step-up gears as they have optimal rotating speeds below the synchronous speed.

According to Liu (2014), gear failure is one of the most expensive failures in power generation. The most common failure modes are tooth breakage and macropitting due to high cycle fatigue. Failure root causes include inadequate lubrication and material defects. Predicting wind turbine gearbox failures has been a topic of interest recently. Predictions have been made based on e.g. vibration signals and oil temperature. Wang et al. (2016) proposed a deep neural network method to predict gearbox failures from lubricant pressure SCADA data. They stated that lubricant pressure is a superior prognostic indicator as it is less sensitive to external conditions than the oil temperature, and commercial wind turbines

are not equipped with gearbox vibration sensors. The proposed method was successful in detecting impending failures two to three days ahead.

2.4 Generator

Generator is a device that converts mechanical energy to electrical energy by applying Faraday's law of induction. The law of induction states that change of magnetic flux through a loop induces a voltage. Rotating part of a generator is called rotor and stationary part of a generator is called stator. Depending on the design, one of these generates the magnetic field and the other has a winding in which an electric current is induced by the changing field. Power plant generators can be divided to two categories based on how the excitation is implemented. The magnetic field of a *synchronous generator* can be generated electrically by field winding or by permanent magnets. Electrically excited synchronous generators require an excitation system. Practically all large generators are electrically excited synchronous generators, as they have the best controllability features. *Induction generators* draw excitation power from the grid. Thus, they consume reactive power and are less controllable. Induction generators are called also asynchronous generators as they must rotate faster than the synchronous speed to produce active power. Turbogenerators and salient pole generators are synchronous generators while both types are used in wind turbines.

2.4.1 Turbogenerator

Generators used in thermal power production are called turbogenerators. Turbogenerators are characterized by high rotation speed and low number of poles. Normal operation speed of a turbogenerator is 3000 rpm for two poles and 1500 rpm for four poles at 50 Hz. The rotor is cylindrical, and it has long axial length and small diameter to withstand the high centrifugal stresses resulting from the high speed.

Turbogenerators can be considered the most critical type of generators. They have largest capacities, and thermal power plant unit has typically only one generator as opposed to hydro power plants or wind farms which have multiple parallel generators. As a result, failure of a turbogenerator leads to considerable lost production. Babin et al. (2020) made a statistical analysis of turbogenerator failures based mainly on Russian operating experience. Most of the generator in the data were manufactured in 1960s and 1970s. It was identified that failure rates of the generators have decreased by 1.5–10 times during their lifetimes. The

improvement was assumed to be due to the fact that many parts have since been replaced with modern ones. Table 1 presents failure distribution by main components.

Table 1. Turbogenerator failure distribution by components according to Babin et al. (2020)

Component	Share of failures
Stator	23 %
Exciter	20 %
Bearings and sealings	19 %
Brushes and slip rings	16 %
Rotor	14 %
Air and gas cooling system	8 %

It was recognized that most of the failures happened in the rotor-stator system and exciter. Also distribution of failure causes was analyzed in the study, and the following distribution was obtained from insurance data: 40 % operation, 43 % design, 14 % ageing and 3 % other. (Babin et al. 2020) The operation related failures includes failures resulting from non-compliance with the maintenance regulations.

2.4.2 Salient pole generator

Salient pole generator is a generator type commonly used in hydropower plants. It is characterized by low rotating speed and large number of projected poles. The number of poles can vary between 4 to 60, which corresponds to operating speeds of 1500 to 100 rpm at 50 Hz. The rotor has short axial length and large diameter. Hence, it can be mounted vertically.

USBR (2017) expects 40 years of service life from its generators. Failures of a hydroelectric generator systems (HGS) can be divided to three categories based on the root cause: hydraulic, mechanical and electric. Xu et al. (2019) conducted a study on fault diagnosis of a HGS. 15 different failure modes were identified, and the fault revealing frequency characteristics were studied. Figure 7 presents the failure modes with related diagnosis indicators.

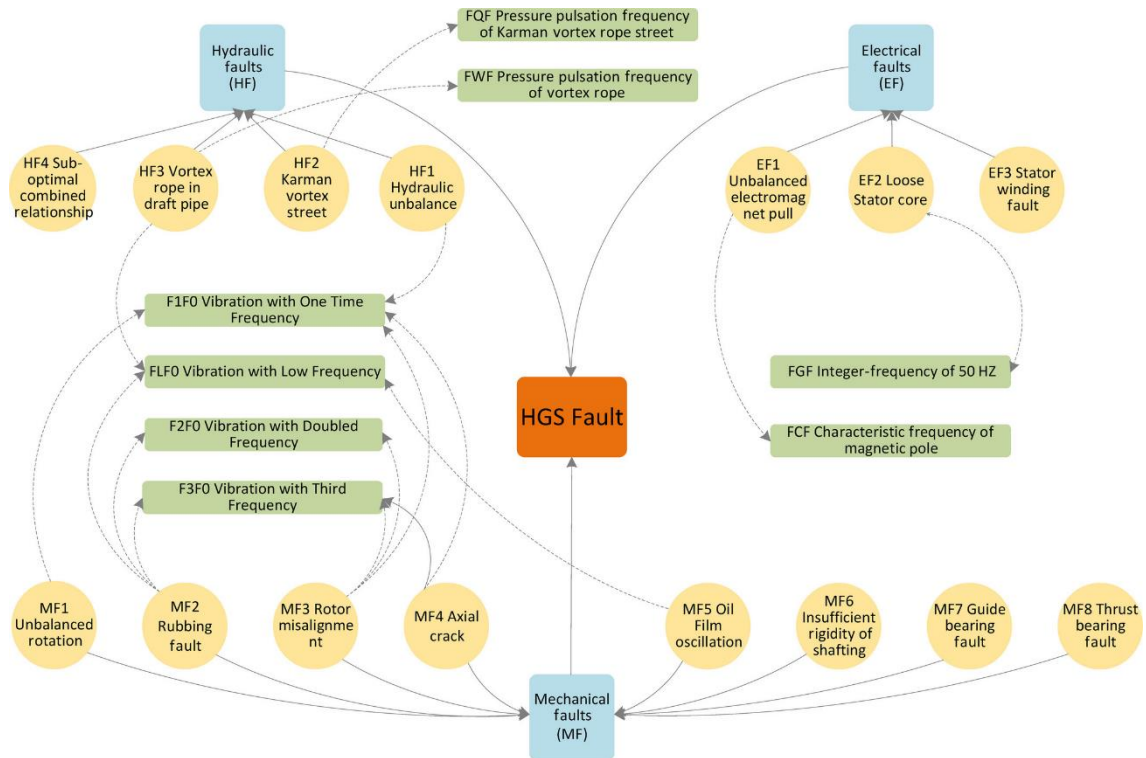


Figure 7. HGS failure modes and associated frequency characteristics indicating a fault. (Xu et al. 2019)

Vibration is a suitable fault indicator as 80 % of HGS' failures are caused by vibration of hydraulic, mechanical or electrical components. Generally, most of the vibration failures result from unbalanced rotating bodies or pressure pulsation of flow passage components. (Xu et al. 2019) However, there are still failures which cannot be detected from vibration. For example, insulation of generator stator windings deteriorates with age, and stator winding is exposed to more voltage stresses than the rotor winding. (USBR 2017) USBR suspects that generator stator winding could be a major area of concern when starts and stops of a hydropower plant are increased. The concern arises from an assumption that thermal cycling resulting from cyclic operation could damage the insulation of stator coils. However, there is yet to be little evidence of this relation. (USBR 2014)

2.4.3 Wind power generators

Wind is a challenging prime mover for a generator. Due to its intermittent nature input torque fluctuates substantially. When discussing wind power generators, it is necessary to consider the whole drive train architecture. Rotational speed of a wind turbine is far slower than frequency of the grid, meaning that there is need for gear box, power converter,

generator with many pole pairs or some combination of these. Figure 8 presents different wind turbine generator (WTG) designs used in commercial operation.

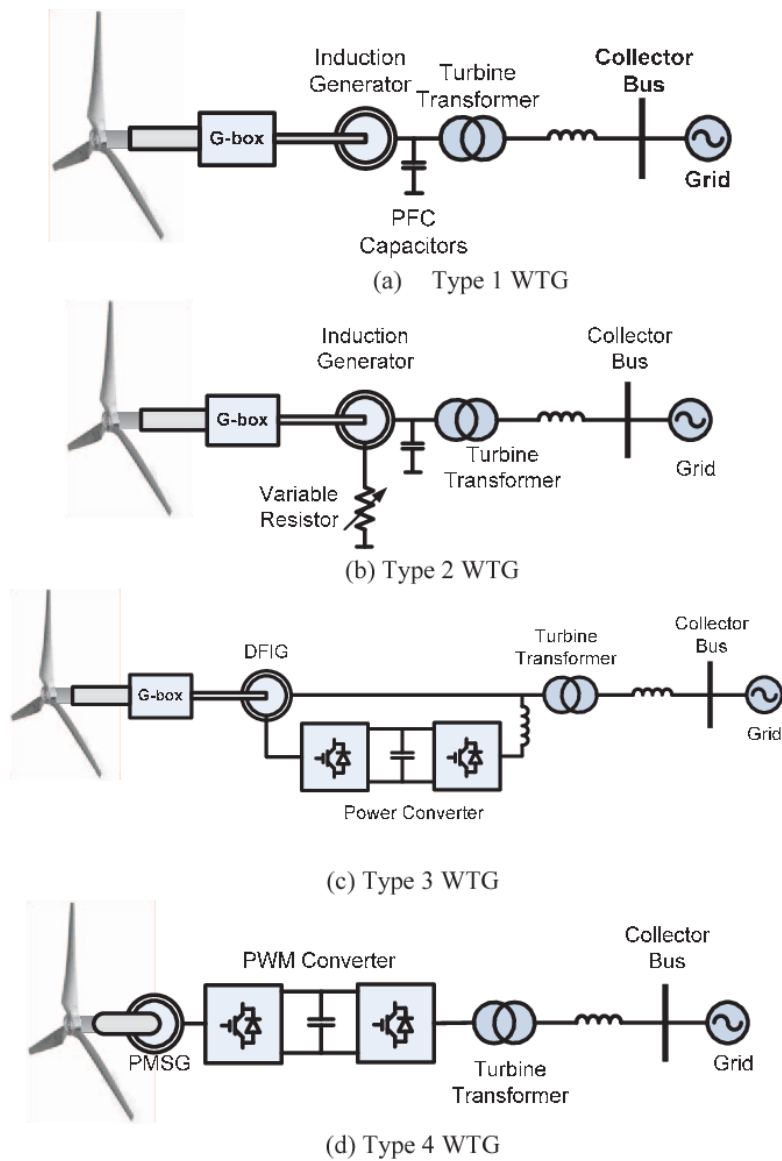


Figure 8. Different types of wind turbine generators.
(Girsang et al. 2014)

While HAWT seems to have reached technology lock-in, there is still competing design alternatives for drive train architectures. The original design was fixed-speed squirrel cage induction generator (Type 1 WTG), which has largely become obsolete due to tightened grid compliance regulation and alternatives with better efficiencies. One approach to overcome issues related to fixed-speed operation was rotor resistance controlled limited variable-speed design (Type 2 WTG). However, this design was not a success, and the manufacturer Vestas has discontinued its production. (Letcher 2017, p. 155-156) Nowadays, variable-speed

designs with improved power quality are preferred. Variable-speed design allows turbines' rotational speed to increase when a gust of wind hits the turbine. This reduces exposure to large forces resulting from gusts. Variable speed design also enables the turbine to operate at optimal tip speed ratio at all wind speeds under rated wind speed. (Letcher, 2017, p. 154) At rated wind speed generator has reached its maximum capacity, and after that increase in wind speed does not increase electricity production. Wind turbine generators are generally dimensioned considerably smaller than maximum mechanical power output of the turbine to ensure sufficient full load hours needed for cost-efficiency.

Variable speed drive train architectures can be subdivided to two types: partial-scale (Type 3 WTG) and full-scale (Type 4 WTG) power converter architectures. Partial-scale converter designs use high speed geared doubly-fed induction generators (DFIG). DFIG architecture is currently the most used option in onshore turbines. Full-scale converter designs use typically synchronous generators, which can be either direct drive or include a gear box. (Böhmeke 2020)

Doubly-fed induction generator

High-speed DFIG concept, originally proposed by Pena et al. (1996), aims to enable variable speed operation at minimal cost. In this design the turbine rotor is connected to a step-up gearbox. Rotor of the generator is connected to the grid through a power converter and the stator is connected to the grid directly. This configuration allows generator speed to fluctuate. Capacity of the converter defines generator's speed range, and it is typically around $\pm 30\%$ of the synchronous speed. Main advantage of DFIG is that required power converter capacity is only about 30 % of the generator's rated capacity. (Letcher 2017, p. 157) This results in considerable savings compared to full capacity converter required by low-speed direct drive train architecture. However, DFIG drive train requires a gearbox and slip rings, which require regular maintenance. According to Ragheb & Ragheb (2010) gearbox has been weakest link of wind turbines with failures being common within 5 years of operation. This is an expensive failure as replacement cost of the gearbox is approximately 10 % of the total cost of a wind turbine.

Full-scale converter synchronous generator

Full-scale power converter drivetrain can be implemented with or without a gearbox. In low-speed direct drive train architecture the turbine is directly coupled to the generator i.e. there

is no step-up gear to increase rotational speed of the generator's rotor. Generator is connected to a full capacity converter, which decouples mechanical drivetrain from the grid. Generator can be either electrically excited (EESG) or use permanent magnets (PMSG). PMSG has better efficiency as it does not require electricity for excitation. However, PMSG requires rare-earth metals which are associated with volatile prices and political risks since China possesses majority of the proven reserves (Chen & Zheng 2019).

This drive train design has certain advantages as well as disadvantages. Absence of the gearbox improves drive train's reliability and simplifies maintenance. Also efficiency is increased when there is no mechanical losses due to gearbox. Further, slow speed decelerates bearings' wear off. Generator can operate at any speed and turbine is free to accelerate when gust hits it. However, full capacity converter is expensive and low speed multipole generator is larger than high speed ones, and thus more expensive. Therefore, gearboxes are used also with full-scale converters. While generator is not exposed to disturbances coming from the grid, it is directly exposed to rotor loads. Gearless direct drive wind turbines suffer from more frequent electrical and electronics failures compared to ones with a gear box but have higher overall availability than the geared ones. (Böhmeke 2020)

2.5 Balance of plant

Balance of plant (BoP) refers to supporting components and auxiliary systems of a power plant. BoP systems are needed to ensure stable, efficient and safe operation as well as grid compliance. BoP can be divided to mechanical and electrical BoP. Depending on type of the power plant, mechanical BoP can include e.g. supporting structures, cooling and process water systems, air conditioning and fire protection devices. Electrical BoP is basically similar for all power plants. This subchapter introduces main components of the electrical BoP.

2.5.1 Transformer

Transformer is an electrical device that can transfer alternating current from a circuit to another without conductive connection and change its voltage. Both circuits are connected to the transformer's core by windings. A varying current in the primary winding generates a varying magnetic flux in the core, which then induces a current in the secondary winding. Turns ratio between the primary and secondary windings determines the voltage difference over the transformer. Main components of a transformer are core, windings, insulation, bushings, tap changer, cooling system and enclosure. Tap changers are used to adjust the

output voltage by changing the turns ratio. Power plants have generator step-up transformers (GSU) to increase the generator output voltage to the grid voltage.

Expected service life of a transformer is 45 years. (USB 2017) Primary life-limiting factor of a transformer is insulating paper ageing. Other possible failure mechanisms of a transformer are core insulation deterioration, thermal faults, winding movement, dielectric fault and corrosive oil. (UK OFGEM 2017) Antoun (2018) made a comprehensive statistical analysis of power transformer and high voltage circuit breaker failure modes. The data included 964 transformer failures from 167 459 transformer-years, corresponding to failure rate of 0.0058 failures per year. GSUs were less reliable than substation transformers with failure rates 0.0093 and 0.0053 per year, respectively. The failure modes were aggregated and the following distribution was obtained: 37 % dielectric, 20 % mechanical, 16 % electrical, 11 % thermal, 3 % chemical and 13 % remained unknown. Also contributions of different components were studied as shown in Figure 9.

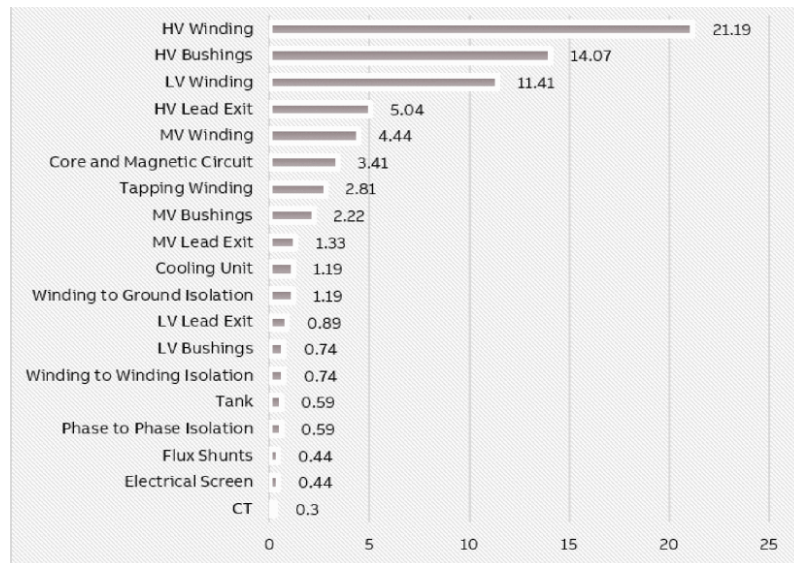


Figure 9. Power transformer failure distribution by components. (Antoun 2018)

Windings and high voltage (HV) bushings are the components that are most susceptible to failures, with dielectric failure of HV bushing being probably the most common failure. Most common identified failure root causes were: 12 % ageing, 12 % external short circuit, 10 % design, 10 % manufacturing and 6 % improper repair. Dissolved gas analysis combined with load current and voltage monitoring provides wide diagnostic coverage of oil immersed transformers' failures. (Antoun 2018) According to USBR, generator's starts and stops do not cause significant degradation to the GSU transformer as GSUs are typically kept

energized when generator is offline. As a result, thermal cycling resulting from the GSU being de-energized and re-energized is avoided. (USBR 2014)

2.5.2 Switchgear

Switchgear contains switches, circuit breakers and fuses. Function of the switchgear is to control, protect and isolate electrical equipment. Generally, switches are intended for normal operation while circuit breakers and fuses are used for protection and fault isolation. Circuit breakers and fuses are automatically operated protective devices used to protect electrical circuits from harmful overcurrent. Primary function of the both devices is to detect the overcurrent and open the circuit such that the overcurrent does not damage other devices connected to the circuit. Hence, the most critical failure mode of protective switchgear is failure to open on demand. (Antoun 2018) Main characteristics of the electrical protective devices are rated current, rated voltage, breaking capacity and operating time.

Fundamental difference between fuses and circuit breakers is that fuse can operate only once, while circuit breaker can operate multiple times. Fuse opens the circuit by melting, and thus it must be replaced after every operation. Circuit breaker has separate mechanisms for detecting and interrupting the overcurrent. The current is typically detected from its magnetic or heating effects. In low voltage applications the overcurrent detection is commonly built-in, while in high voltage applications there is a protective relay to detect the overcurrent. The contacts are opened by electromagnetic force resulting from the current itself or mechanical energy stored in spring or compressed air. (Garzon 2002)

USBR expects service life of 45 years from its air magnetic/air blast circuit breakers, and 50 years from oil tank, SF₆ and vacuum type circuit breakers. Life expectancy has increased since earlier report from 2005 estimated service life of only 35 years for all circuit breakers. (USBR 2017) However, for fault clearing devices number of operations has significant effect on the service life. Thus, especially for circuit breakers subject to high number of operations, end of life is likely to be defined by number of operations rather than age (UK OFGEM 2017).

Antoun (2018) studied HV circuit breaker failure modes and their root causes based on operating experience from 83 utilities. Obtained total failure rate was 0.003 failures per year. Circuit breakers implemented for shunt reactor or capacitor bank switching exhibited considerably higher failure rates of 0.025 and 0.01, respectively. The most significant failure

modes resulting to total failure were: 28 % does not close on demand, 25 % locked in position, 16 % does not open on demand, 11 % electrical breakdown in main circuit, 8 % loss of mechanical integrity, 5 % spurious opening and 0.2 % spurious closure. Locked in position refers to situations where the circuit breaker is unable to operate but it is discovered without an actual demand. Majority of the partial failures were medium leaks. Root cause was wear/ageing in about half of the failures. The next most common (20 %) root causes were those introduced before the equipment entered service e.g. manufacturing, design, transport and installation failures. The component which tends to fail is typically related to the operating mechanism e.g. compressor for pneumatic mechanism and mechanical transmission for spring mechanism.

2.5.3 Automation

Automation refers to technologies used to reduce need for human interventions. Power plant processes are continuous flow processes in which control systems are used to operate the plant more efficiently and safely. Control systems can be divided to two main types: open-loop and closed-loop systems. Open-loop control system operates independently from the process output i.e. it does not get feedback. Hence, the system should be modelled accurately so that the control would succeed without possibility to adjust based on feedback. In practice, open-loop control is used in simple applications where the result is known to be sufficient without feedback. On the contrary, closed-loop system measures the process outputs and adjust operation based on difference between the measured and desired value. As a result, the set point can be met even without accurate model of the system. Figure 10 presents general closed-loop control system scheme.

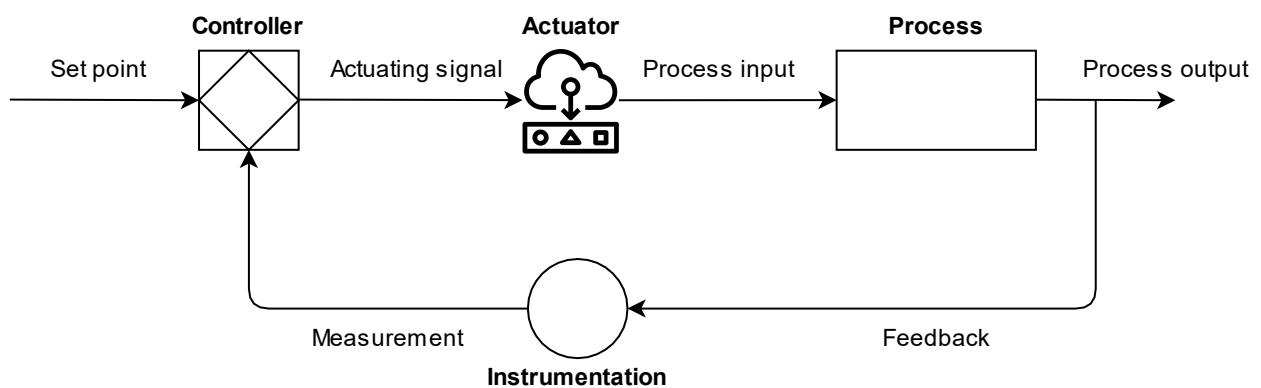


Figure 10. Closed-loop control system compares process output to the set point and adjusts operation based on the difference. Open-loop control systems do not have feedback loops.

Main components of a control system are instrumentation, controller, actuator and human machine interface. Instrumentation is the interface between physical process quantities and the control system. Instrumentation sends input signals to the controller, which processes them and sends actuating signals to the actuator. The actuator converts the signal back to physical process quantities. Human machine interface allows the operator to interact with the process. This section introduces the main components of a control system. Reliability of programmable logic controller and human machine interface is analyzed in detail in the case study.

Instrumentation

Instrumentation is a collective term referring to measuring instruments which are used for measuring, indicating and recording physical quantities. By definition, measuring means quantification of a physical quantity on some interpretable scale. In a case of process control, this scale must correspond to the input signals accepted by the controller e.g. 4-20 mA current. Basic measurements used in power plant processes are flow, level, pressure, temperature, vibration, voltage, current and frequency measurements. Besides from total failure where output generation ceases, instrumentation failure modes include bias, drifting and precision degradation. (Doymaz et al. 2001) Faulty outputs may seem plausible, and therefore it can be difficult to distinguish failed instrumentation from process upsets.

In addition to process control, another important application of instrumentation is fault diagnosis and prognosis. Instrumentation has long been used in applications such as bearing temperature monitoring. Current state-of-the-art in online condition monitoring is combining high frequency sensors with artificial intelligence to detect incipient failures before they can be detected with traditional methods. High frequency condition monitoring techniques have been used especially in bearing vibration analysis. (Malla & Panigrahi 2019)

Controller

Process controller is a device that controls process variables based on received inputs and its internal logic. There are many different types of controllers, spanning from the centrifugal controller to modern computer based systems. However, this study focuses on programmable logic controllers (PLC) as they are prevalent in modern industrial processes. PLC is a digital computer which was initially introduced to replace hard-wired relay logics earlier used in industry. Main advantages of a PLC compared to relay logics are easier

programmability and smaller size. Main differences between a PLC and general computers are that the PLC is optimized for process control and it is designed to withstand harsh industrial conditions e.g. vibration, dust, moisture and extreme temperatures. (Bolton 2015, p. 3) PLCs use real-time operating systems since there are constraints for response time as failure to operate within required time may result in unintended consequences for the controllable physical process.

Size of a PLC ranges from small compact device with tens of inputs and outputs to large rack-mounted modular devices with I/O counted in thousands. Despite the large differences in size, all PLCs generally consists of a power supply unit, central processing unit (CPU), input and output interfaces, and communications interface. (Bolton 2015, p. 4-5) Power supply unit converts mains AC voltage to low DC voltage used by the PLC. As a switching power supply it consists of a filter circuit, a rectifier circuit, a switching converter circuit and a pulse rectifier circuit. (Xu et al. 2020)

Processor unit interprets the input signals received from the input interface and executes control actions according to the program stored in its memory. Input interfaces use optoisolators to transmit signals without conductive connection. Optoisolator consist of a light-emitting diode and a photo-transistor, and it protects the processor from voltage spikes coming from the input devices. Processor unit includes control unit, arithmetic and logic unit, and memory. Control unit is used for directing and timing of the operations. Arithmetic and logic unit is responsible for arithmetic operations such as addition and subtraction, and logical operations like AND, OR, NOT and exclusive-OR. Memory architectures vary but typically there are read-only memory (ROM) for the operating system and random-access memory (RAM) for the user's program and data. There can be also erasable and programmable ROM or Secure Digital memory card. RAM is typically dynamic-RAM which stores each bit of data into a dedicated capacitor. Electric charge will leak from the capacitors, meaning that the memory must be frequently refreshed or it will be lost. (Bolton 2015, p. 4-10)

Output interface communicates the processor's commands to actuators. Different types of outputs are relay, transistor and triac outputs. Relays can switch large currents and are suitable for AC and DC switching. However, they have moving parts and are thus relatively slow and subject to wear-out. Transistor and triacs are fast to operate as solid-state devices,

but need optoisolators for protection. Transistors are used for DC and triacs for AC. Communications interface connects the PLC to the supervisory system and other PLCs. An external programming device is typically used to program the PLC. (Bolton 2015, p. 4-10)

PLC programming languages are designed to be intuitive such that no programming skills are necessarily needed to set up or change the control programs. Standard IEC 61131-3 defines five languages used for PLC programming. There are 3 graphical languages and 2 textual languages. The graphical languages are ladder diagram, function block diagram and sequential function chart. Ladder diagram resembles relay logics, which eased transition from hardwired relay logics to PLCs. The textual languages are instruction list and structured text. (Bolton 2015, p. 14-15) Despite the differences, all of the languages are based on logical connections. Even though the languages are standardized, manufacturers try to differentiate their products such that a customer would get locked-in to their products (Padhi & Lila 2018).

Actuator

An actuator is a device that converts input signal to some form of motion. Typical applications in flow processes are controlling valves and dampers which convert the controller's commands into changes of flow and pressure. The actuating signal is relatively low energy, and power is supplied separately. The actuator can be powered by e.g. electricity, hydraulic pressure or compressed air.

Pneumatic actuators are reliable, inexpensive and fast-operating, but they require compressors to operate. Hydraulic actuators are powerful, accurate and fast, yet they need pumps. Electric actuator is the most precise and energy efficient type of actuator, and it requires only an ordinary power supply. However, they tend to be more expensive compared to other types. (Lindsley et al. 2018, p. 217-223)

Human machine interface

Human machine interface (HMI) is a device that allows humans to interact with a machine. HMIs are used to supervise and control industrial processes. A HMI hardware generally consists of a display, CPU and I/O devices. (Zhang 2010) Input devices include e.g. buttons, switches, keyboards and mice. In addition to visual outputs, also sound and haptic outputs can be utilized. A touchscreen is simultaneously an input and output device.

Industrial HMIs can be divided to two main categories: supervisory and machine level. (Zhang 2008) Supervisory level HMIs are located in control rooms, and they are intended for system control and data acquisition (SCADA). Used hardware on the supervisory level is nowadays typically PCs and displays. Machine level HMIs are located in the process facilities and have more limited functionalities. They can be basic panels with buttons or touchscreens.

3 Reliability centered maintenance

Reliability centered maintenance (RCM) is a structured analysis method used to identify safe and cost effective failure management strategies. The term failure management strategies is used to emphasize that scope of the RCM extends beyond maintenance. RCM was initially intended for aircraft maintenance planning but to date the scope has expanded, and the method has proven track record from nearly all industries (Regan 2012, p. 1-5).

Main output of the RCM analysis is a cost effective and defensible maintenance plan. Other outputs may be improvements to operating procedures, redesigns of equipment, updates to technical publications, modified training programs, supply changes, enhanced troubleshooting procedures, and revised emergency procedures. (Regan 2012, p. 2) In addition to aforementioned outputs, one intangible benefit of RCM is greater understanding of risks that the organization is exposed to. After the initial RCM analysis, it should remain as a continuous process that is updated as new information becomes available.

There has been many kinds of service providers and methods which have declared to provide the RCM. Therefore, a standard has been created to avoid confusion about what is regarded as the RCM. According to the standard SAE JA1011 (2009), RCM analysis must answer to seven following questions:

1. What are the functions and associated desired standards of performance of the asset in its present operating context (functions)?
2. In what ways can it fail to fulfill its functions (functional failures)?
3. What causes each functional failure (failure modes)?
4. What happens when each failure occurs (failure effects)?
5. In what way does each failure matter (failure consequences)?
6. What should be done to predict or prevent each failure (proactive tasks and task intervals)?
7. What should be done if a suitable proactive task cannot be found (default actions)?

This chapter discusses all the necessary steps related to the RCM framework. To begin with, the relevant reliability theory is introduced as prerequisite information. Failure mode and effect analysis (FMEA) covers the first four questions. Criticality analysis subchapter

answers to the fifth question. The last two questions are addressed in subchapters Maintenance strategies and Default strategies.

3.1 Reliability theory

Reliability of a component can be defined as the ability to perform its intended functions satisfactorily under stated conditions for a specified period of time. It can be quantified as the probability of successful operation. Inability to perform the intended task when demanded is considered a failure. According to this definition state of the component is binary; it is either in operational state or failed state. Even though the RCM method distinguishes partial failures from total failures, most conventional reliability models handle state of a component as a binary variable. However, in reality state of a component is rather continuous than discrete or binary. Continuous nature of the state's development is difficult to represent in practical applications but discrete scale with more than two steps can be useful. Despite the discrete representation of the state's development, the underlying degradation process is continuous, and there must be a defined limit state between operational states and failed states. Defining this limit state depends on the context and is non-trivial in some cases as there may be partial failures. RUL models often assume a predefined and measurable failure threshold. The threshold is usually determined based on engineering domain expertise or industry standards. (Si et al. 2013) However, as the review by Pfaffel et al. (2017) revealed, definitions of a failure may vary significantly between different studies, which impairs comparability.

It is worth noting that reliability of a component is not only its inherent and independent feature, since it depends on conditions that the component is subjected to. To consider effect of external factors, Menčík (2016) has introduced two different terms to describe reliability of a component. Inherent reliability is a characteristic of the component, which is a result of design, material choices and manufacturing. Operational reliability is the actual reliability achieved during use. It is affected by way of use, conditions and maintenance. Operational reliability can be improved with O&M practices, but it can never exceed the inherent reliability.

This subchapter introduces the basic theories in reliability engineering. Foundation of the reliability theory is probability and statistics. At first, metrics used to measure reliability and

relevant distributions are presented. After that, failure patterns, i.e. how reliability changes as a function of age, are studied.

3.1.1 Reliability metrics and distributions

Reliability is measured with various different metrics, e.g. mean time between failures (MTBF), failure rate, mean time to repair (MTTR), availability and probability of failure. In the context of power plants also lost production and capacity factor can be used as reliability metrics for the whole plant. The term MTBF is used for repairable components, while corresponding term to non-repairable components is mean time to failure (MTTF). MTBF is the reliability metric usually provided by manufacturers, which is based on assumption of constant failure rate i.e. the flat middle section of the bathtub curve. (Kececioglu 2002)

Failure rate is the inverse of MTBF. The fact that MTBF is often preferred against the failure rate is due to more intuitive interpretation of MTBF. MTBF is typically measured in hours, which may mean time from deployment or operational hours. However, MTBF can be as well measured in e.g. distance travelled in kilometers or number of operational cycles. Generally, MTBF or failure rate measured in time is more suitable for components in continuous operation, and in demands approach is more suitable for components which are used infrequently and for a short period of time. An example of a case where failure rate per demand approach could be practical is power plant's lifting crane which is used only during overhauls.

MTBF and failure rate do not consider downtime resulting from failures. MTTR is the average time it takes to restore the failed component back to operational state. Availability is the ratio of component's operational time to total time. It can be calculated from MTBF and MTTR with the following equation:

$$Availability = \frac{MTBF}{MTBF + MTTR} \quad (1)$$

However, MTBF or availability do not consider age of the component, which can affect the reliability. Consequently, probability calculus is used to estimate component's reliability given it has survived to a certain age. In reliability theory, lifetime of a component is assumed to be a random variable and it is typically denoted with T . Functions commonly used in

reliability estimation are survival and hazard functions. (Holmberg 2020) These functions can be derived from the probability density function (pdf) and cumulative density function (cdf). All of the four aforementioned functions contain the same information i.e. if one of them is known, others can be derived from it.

The pdf represents how component failures are distributed in time, and the cdf shows the probability that a component will fail at or before a certain time. Survival function is the complement of the cdf:

$$S(t) = P(T > t) = 1 - F(t) \quad (2)$$

When survival function is evaluated at time t , the result can be interpreted as the probability that single random unit survives beyond time t or as the proportion of the entire population surviving beyond time t . If it is known that a certain unit has survived until time t , the conditional probability that the unit survives beyond $t + \Delta t$ can be calculated from conditional survival function:

$$S(t + \Delta t | t) = P(T > t + \Delta t | T > t) = \frac{S(t + \Delta t)}{S(t)}, \Delta t \geq 0 \quad (3)$$

Further, probability of failure (PoF) for any closed interval can also be calculated with the survival function. The closed interval can be e.g. next year. As equation 3 calculates the probability that a component survives beyond time $t + \Delta t$ given it has survived until time t , its complement event is that the component has failed by time $t + \Delta t$:

$$PoF = P(t < T \leq t + \Delta t | T > t) = 1 - \frac{S(t + \Delta t)}{S(t)}, \Delta t \geq 0 \quad (4)$$

Again, the result can be also interpreted as the share of population failing during the interval. Average failure rate (AFR) for an interval is obtained by dividing number of failures d occurring during the interval by the total time expended by the population n i.e. component-years:

$$AFR = \frac{d}{\sum_{k=1}^n t_k}, \quad (5)$$

where t_k is the operating time of component k during the interval. t_k can be less than the interval if the component fails during the interval, enters the experiment late or exits the experiment early. When the time interval Δt approaches zero, failure rate can be calculated by dividing conditional probability of failure by the time interval. Then average failure rate becomes instantaneous failure rate. Hazard function is defined as:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{S(t) - S(t + \Delta t)}{S(t)\Delta t} = \frac{f(t)}{S(t)} \quad (6)$$

Output of the hazard function can be interpreted as the failure rate during the next time instant for the survivors until time t . Expected time to failure for a component that has survived until time t can be obtained from the following equation (Holmberg 2020):

$$L(t) = E(T - t | T > t) = \int_t^{\infty} (\tau - t) \frac{f(\tau)}{S(t)} d\tau, \tau \geq t \quad (7)$$

In a special case of constant failure rate the expected time to a failure is equal to the MTBF. The only distribution to have a constant failure rate is the exponential distribution. With all other distributions the failure rate varies as a function of age. (U.S. NIST 2013)

3.1.2 Failure patterns

A failure pattern describes how failure rate, or probability of failure, develops as a function of age. Different generations of reliability engineers have had different beliefs regarding the patterns of failure. The first generation, which lasted until 1950's, believed that all equipment have a useful life, after which the failure rate increases. The second generation, which lasted until mid-seventies, began when it was identified that also many brand new equipment are more prone to failure. (Moubray 2001, p. 2-4) Nowadays there are six recognized failure patterns, and even more proposed patterns. However, earlier generations were not necessarily completely wrong as equipment was different back then. (McLeod et al. 2015)

Second generation: Bathtub curve

Awareness of infant mortality led to the invention of the so-called bathtub curve, which is still widely used today. The curve shown in Figure 11 proposes that initially failure rate decreases with time, then stays approximately constant and finally increases towards the end of life.

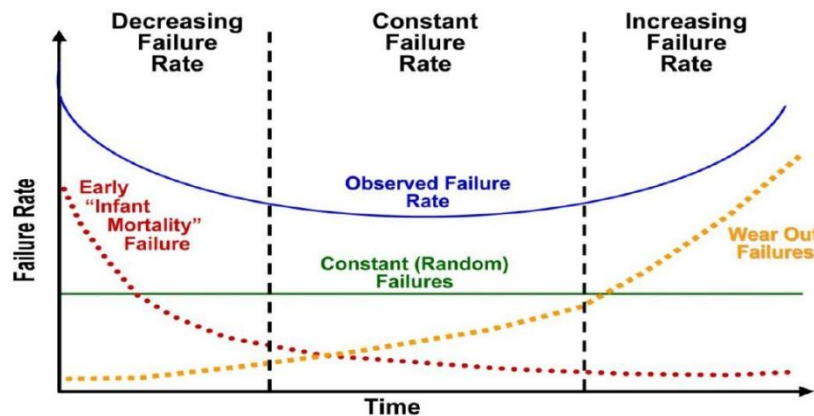


Figure 11. Superposition of three types of failure comprises the bathtub curve. (Maisonnier 2018)

The bathtub curve can be divided to three parts: infant mortality, random, and wear-out failures. The term infant mortality refers to early failures and it decreases with age. Infant mortality is caused by manufacturing defects, deployment mistakes and software bugs. (Kececiloglu 2002) These failures are problematic from owner's and operator's perspective. Owner can't affect or model manufacturer's processes. Deployment mistakes caused by human error are difficult to predict. Hence, infant mortality can be estimated only with statistical models utilizing historical data. Also failures resulting from operator errors can have decreasing rate since it is likely that susceptibility to error decreases with experience. It is worth noting that maintenance activities, especially ones requiring disassembly, may expose the component to infant mortality again, as shown in Figure 12.

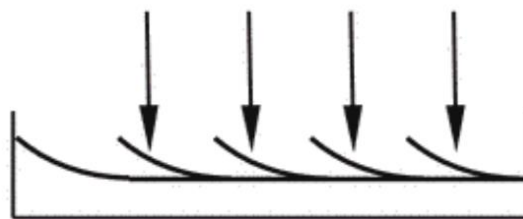


Figure 12. Time-based maintenance reintroducing infant mortality. Periodic maintenance is suitable only for assets for which wear-out is greater problem than infant mortality (Regan 2012)

Random failure rate does not depend on age of the component, and it stays constant throughout the lifetime. Random failures mostly result from external shocks. Response time from occurrence of the external factor to failure of the component is often immediate or short, meaning that oncoming random failure cannot be predicted by monitoring the component itself. Consequently, estimation of random failure rate initiates from estimating frequency of external shocks leading to the failure. (Kececioglu 2002) For electrical components such an external shock can be e.g. voltage spike coming from the grid, and for mechanical components the cause can be force which exceeds the design load. Other external events which may cause immediate failure include fire, lightning, earthquake and flood. Plant operator cannot prevent these kind of events from happening but can limit consequences of such events. Severity of potential consequences determines how much attention should be paid to these kind of rare yet disastrous events. Nuclear industry must consider extremely unlikely events since potential consequences of the failure are so disastrous. In most industries it would be prohibitively expensive to design components to withstand highly unlikely events, meaning that some level of random failures must be just accepted. However, risks can be mitigated also by insurance.

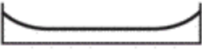
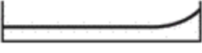



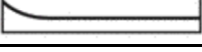
Wear-out failures are caused by accumulation of damage, and thus failure rate of wear-out failures increases with age. Degradation process, which eventually causes wear-out failure, advances gradually, and it can be typically detected before the failure. If physical properties of the process are known, time to failure can be estimated even without data. Wear-out depends on material properties and external conditions. Examples of wear-out processes are thermal degradation, corrosion, erosion, fatigue and crack growth. (Kececioglu 2002) In some sense, wear-out failure is the most preferable failure type as failure is inevitable at some point for components subject to wear. Theoretically, it would be optimal that components fail just after reaching the end of system's lifetime since overdesign is not a cost effective practice. However, in practice occurrence time of a failure is stochastic by nature, system's lifetime is difficult to predict and also residual value might be a point worth considering.

Third generation: Six patterns of failure

Investigations made in the aviation industry during 1960's and 70's led to the beginning of the third generation. The investigations were initiated after, by today's standards highly unacceptable, failure and crash rates were tried to be improved by shortening overhaul

intervals, but it did not help. (Regan 2012, p. 6-8) The revolutionary results challenged the prevalent maintenance philosophy that failures are related to age, and revealed that there are actually six patterns of failure, as shown in Table 2.

Table 2. Shares of failure patterns according to different studies. (McLeod et al. 2015)

Failure pattern	Aircrafts (Broberg 1973)	Aircrafts (Nowlan & Heap 1978)	Navy (U.S. MSDP 1982)	Electronics (U.S. SSMD 1993)	Navy (U.S. SUBMEPP 2001)
A 	3 %	4 %	3 %	6 %	2 %
B 	1 %	2 %	17 %	-	10 %
C 	4 %	5 %	3 %	-	17 %
D 	11 %	7 %	6 %	-	9 %
E 	15 %	14 %	42 %	60 %	56 %
F 	66 %	68 %	29 %	33 %	6 %

Failure patterns A, B and C exhibit age or wear-out related phenomenon, and patterns D, E and F exhibit randomness after initial settling period. All of the studies got similar results that most of the failures (71-93 %) are random rather than related to age. However, mechanical devices often follow pattern A or B, and pattern C can be observed when fatigue is the dominating failure mode. (McLeod et al. 2015) Generally, reliability decreases with age if there is a dominating age-related failure mode (Moubray 2001, p. 13). Complex devices typically have random failure patterns. Pattern D is related to complex devices with high stress working conditions. Complex devices or machinery with a well-balanced design follow pattern E. Pattern F is observed for electronic components and complex devices after corrective maintenance. Computers and PLCs follow pattern F. (McLeod et al. 2015)

Even though complex devices can have constant failure rates, most of the individual internal failure mechanisms are actually time dependent. (Troyer 2018) Failures caused by external events are typically independent of the device's age. A constant rate can be observed if failure modes with different patterns even out, the component is removed from use before the wear-out occurs, external events dominate or if the internal failure mechanisms have a constant failure rate. Figure 13 shows how failure modes exhibiting different patterns can even out to form a constant failure rate.

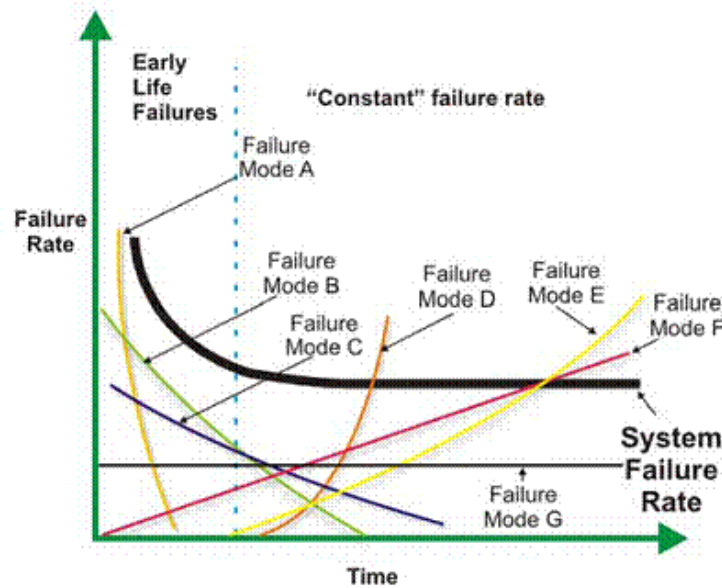


Figure 13. Observed failure rate is the sum of competing failure modes. (Troyer 2018)

Also dead on arrival can be included when discussing failure patterns. Dead on arrival means that asset is already in failed state when received, and it belongs to the manufacturer's side of the reliability assessment. Dead on arrival can be prevented with testing. (Hansen & Thyregod 1992) However, it is noteworthy that cost optimal defect rate from manufacturer's perspective is generally small although not zero. Investments to production process quality have diminishing marginal utility since additional investments reduce profit obtained from non-defect products. Optimal investments to production quality for reducing defect rate are studied by e.g. Dey & Giri (2014) and are not further considered in this thesis. From customer's perspective it is important that warranty covers early failures.

3.2 Failure mode and effect analysis

Failure mode and effect analysis (FMEA) is an engineering technique used to identify potential failure modes and their causes and effects. FMEA can be performed also independently, but is included in the RCM analysis, which has wider scope than the FMEA.

3.2.1 Functions

The first step in the RCM analysis is to define what is required of an asset. Only after that it can be understood whether the asset is capable of fulfilling those requirements. These requirements can be divided to the primary function and secondary functions. The primary function is the main reason why the asset exists, and secondary functions are other useful functions the asset can provide. (Regan 2012, p. 23-24) What the asset does is important,

not what it is. A distinction must be made between two important features of asset performance: *design capability* and *required performance*. Design capability means what an asset is originally designed to be capable of doing, and required performance means what is required of an asset in a specific operational context under review. Obviously, required performance can never exceed the design capability. Even though this is a very simple concept, there has been catastrophic failures resulting from required performance exceeding the design capability. (Regan 2012, p. 15-21)

The second step in the RCM method is identifying functional failures i.e. in what ways the asset can fail to fulfill its functions. Functional failures can be divided to total failures and partial failures. Total failure is a complete loss of a function, while partial failure means that the asset is not able function at the level of required performance. (Regan 2012, p. 24) In typical case partial failure occurs when performance decreases below the required performance as a result of gradual degradation. However, performance can be also limited by upper limit e.g. maximum allowed leak or both upper and lower limits e.g. signal.

3.2.2 Failure modes

Identifying failure modes for the functional failures is the third step in the RCM process. A failure mode is a cause of a functional failure. Failure modes should include age and usage related ones like those due to e.g. wear, corrosion and fatigue. However, it is equally important to consider also failure modes covering random causes like external shocks and human errors. (Regan 2012, p. 24-25) Nonetheless, not all possible failure modes should be included. According to the standard SAE JA1011 (2009), RCM analysis should include all failure modes that are *reasonably likely* to occur. What is considered as reasonably likely must be decided on a case-by-case basis, leaving room for subjectivity. Regan (2012, p. 25) has presented four questions to support this decision:

1. Has the Failure Mode happened before?
2. If the Failure Mode has not happened, is it a real possibility?
3. Is the Failure Mode unlikely to occur but the consequences are severe?
4. Is the Failure Mode currently managed via proactive maintenance?

If an affirmative answer is given to one or more of the above questions, the failure mode should be included in the analysis.

3.2.3 Failure effects

Step four of the RCM analysis is identifying failure effects for each failure mode. A failure effect describes what would happen if nothing were done to prevent or predict the failure mode, and the following step five discusses why does it matter. Description of the failure effects should include the following (Regan 2012, p. 93-94):

1. A Description of the Failure Process from the Occurrence of the Failure Mode to the Functional Failure
2. Physical Evidence that the Failure has Occurred
3. How it Adversely Affects Safety and/or the Environment
4. How It Affects Operational Capability
5. Specific Operating Restrictions as a Result of the Failure
6. Secondary Damage
7. What Must be Done and How Long It Takes to Repair the Failure

3.3 Criticality analysis

Criticality analysis is a method used to assess criticality of assets based on what can happen if they fail. Consequences of a failure can be divided to certain consequences and risks. Risk can be defined as exposure to some undesirable event. From the definition of risk it follows implicitly that the event is possible but not certain. Risk is quantified as the product of event probability and outcome. Outcome is a numerically expressed consequence which is faced if the risk realizes.

According to the RCM methodology, evaluation of failure consequences is a two-step process. First, failure modes are divided to two categories: evident and hidden. Evident failure mode is defined as a single failure mode which becomes evident to the operating crew without inspection under normal conditions. Evident failure mode must become evident on its own without other failures occurring. (Regan 2012, p. 100-103) If these conditions are not met, the failure mode is considered hidden.

After failure modes are classified as evident or hidden, failure consequences are assessed. The RCM methodology recognizes four types of consequences in the following order of priority: safety, environmental, operational, and non-operational. (Regan 2012, p. 111) If a failure can result in injury or death, it has safety consequences. Environmental consequences are the second most severe consequences. According to Regan (2012, p. 113),

only breaches of environmental laws, standards or regulations have environmental consequences, which is somewhat questionable statement as environment can be damaged without breaking the official orders. Economic consequences can be either operational or non-operational. Operational consequences affect operational capability e.g. result in lost production or increased operating costs. Non-operational consequences involve only the cost of repair. (Regan 2012, p. 115-118)

In order to assess total consequences, different types of consequences must be made comparable. One method to enable comparison is converting different types of consequences to the same unit. Conversion factors used for this include e.g. value of statistical life. Comparing human lives to money lays ground for ethical discussions which, however, are omitted in this thesis. Nonetheless, such assessments are necessary for that very reason that limited resources can be allocated in the best way to minimize total environment, health and safety (EHS) consequences. Therefore, it may be more pleasant to assess EHS risks implicitly than it is to price something that is considered priceless in the society. Non-monetary criticality assessment methods include e.g. risk matrices and risk priority numbers (Basu 2017).

Another question worth discussing is an acceptable level of EHS risk. If the decision is left to a company, it follows from the fundamental purpose of a company that the question reduces to a financial optimization problem, where cost of risk is optimized against cost of risk reduction. Importantly, in addition to direct financial consequences, cost of risk must include EHS and also reputation consequences. According to International Organization for Standardization (2014) guidelines, acceptability of a risk should be assessed relatively to the current values of society i.e. what was acceptable risk 50 years ago may not be acceptable today. If liabilities are assigned correctly, company's financial optimization should lead to socially optimal outcome. However, liability as a mean of controlling risk is based on assumption that the company is willing and able to do the risk assessment correctly. (Shavell 2008) If it is suspected that the assumption does not hold, there might be a need for regulation.

One general risk regulation principle is "as low as reasonably practicable" (ALARP). The ALARP principle arises from an assumption that reducing risk to zero would require infinite time, effort and money. Thus, risk is considered ALARP when it can be demonstrated that

the cost of reducing the risk further would be grossly disproportionate to the benefit gained. (UK HSE 2021) Defining what is ALARP requires judgement and the decision must be made on a case-by-case basis. While general risk regulation principles tend to be somewhat subjective, in some cases the regulation can be also well defined e.g. it can be required that safety related equipment have redundancy.

Reliability maximation is often not reasonable since increased reliability typically comes with a cost. Figure 14 presents a trade-off relation between outage and construction costs. Similar relation often applies in general between risk cost and risk reduction cost.

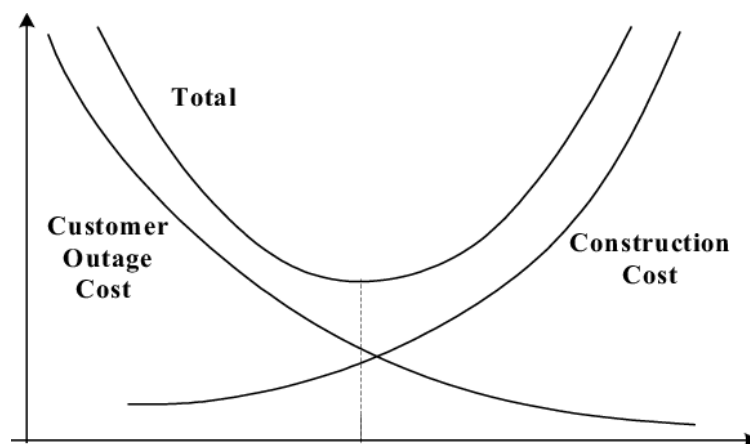


Figure 14. Outage, construction and total cost as a function of reliability. (Tinh et al. 2021)

3.4 Maintenance strategies

Maintenance is an entirety of actions performed to maintain an asset's functionality or to restore an asset to functional state. Maintenance can be divided to two main types: *preventive maintenance* comprises means to preserve asset's functions, while *corrective maintenance* recovers asset's functions after a functional failure. Objective of the maintenance is not to keep condition of the asset as good as new. Instead, the objective is asset's life cycle cost optimization. Optimization problem is often assumed to be a trade-off between cost of maintenance and reliability. However, relation between preventive maintenance activities and reliability may be more complicated than sometimes thought. In addition to increased maintenance costs, unnecessary maintenance can even increase probability of failure, since there is always risk of human error present. Thus, maintenance activities can expose asset to infant mortality and unnecessary maintenance activities should be avoided.

In order to thoroughly comprehend the current state of maintenance strategies, it must be understood how the current state has been reached. Past affects the present since it is always easier to maintain the status quo than it is to make changes and progress. Importance of maintenance has been increasing together with the degree of mechanization.

Evolution of maintenance can be traced to three generations. The first generation lasted until World War II. Back then industry was not highly mechanized, meaning that downtime was not an issue comparable to the current extent. Equipment was also simple and easy to repair. Therefore, *corrective maintenance* was the most prevalent maintenance strategy during the first generation. (Moubray 2001, p. 2)

The WWII marked a change also for the maintenance strategies and many advances in maintenance have been made in the military. Degree of mechanization and complexity of equipment increased, which led to downtime being more of an issue. An important milestone regarding asset criticality is whether the asset can be temporarily substituted by labour or other means. When consequences of a failure became more severe, idea of *preventive maintenance* gained attraction. Back then it comprised overhauls done at fixed intervals i.e. *time-based maintenance* was the first preventive maintenance strategy to be implemented. Consequently, cost of maintenance began to rise, which led to increased interest in maintenance planning. (Moubray 2001, p. 2)

Reliability centered maintenance is the basis of the third generation. To overcome the bias towards continuing what is currently being done, the RCM analysis is *zero-based*. Zero-based means that the RCM analysis starts from an assumption that nothing is done to prevent or predict failures. (Moubray 2001, p. 14) Default is to do nothing and preventive interventions must be justified to be *technically appropriate* and *worth doing* (Regan 2012, p. 123). In this way earlier unnecessary measures are tried to be avoided.

An intervention is technically appropriate if it reduces conditional probability of failure. Worth doing has different meaning for different consequences. In regulated operating environment EHS consequences must be reduced to an acceptable level by any means necessary or the operating permit will be lost. Thus, worthiness of such actions must be assessed against losing the operating permit. An intervention to reduce economic

consequences is worth doing if the intervention costs less than bearing the risk. (Regan 2012, p. 123-134) Same principles apply to default strategies discussed in the next subchapter.

Figure 15 illustrates electric motor deterioration as a function of time and the concept of potential-to-functional failure (P-F) interval. It is important to notice that only certain failure modes exhibit wear-out patterns, and an electric motor can fail also due to an unpredictable random failure. Failure initiated is the theoretical upper limit for detection of a potential failure. Advanced condition monitoring methods can detect the potential failures earlier than human. However, order of different methods varies from case to case and is subject to debate. For example, Bengtsson et al. (2018) stated that potential failure of a ball bearing can be detected from changes in vibrations 1-9 months prior to the functional failure and from particles detected by oil analysis 1-6 months before the functional failure.

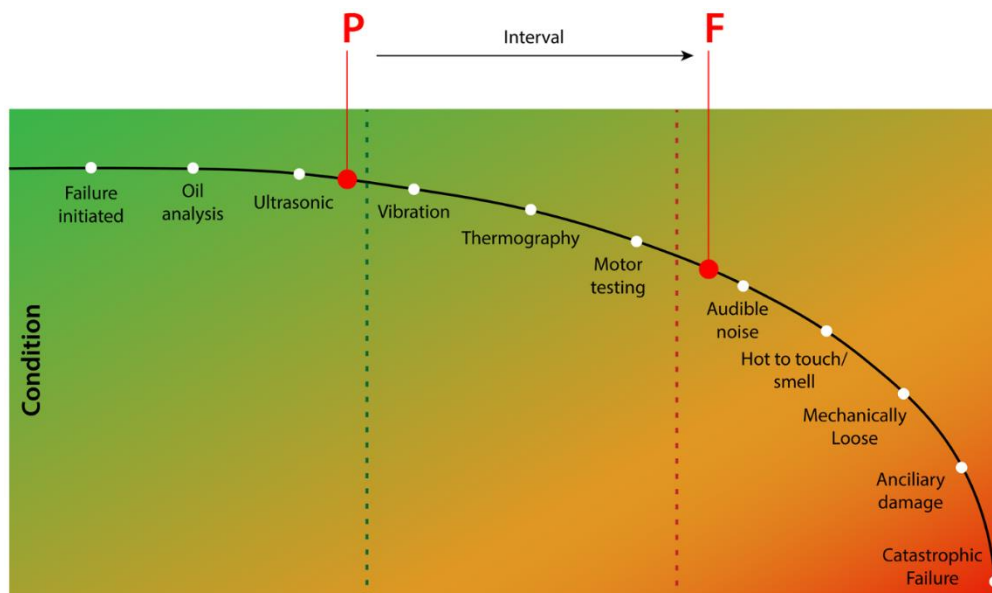


Figure 15. Conceptual P-F curve for wear-out of electric motors. P-F interval depends on used condition monitoring method and required level of performance. (Clark 2019)

A functional failure occurs when the asset performance is below the required performance level determined during the first step of the RCM. How much the asset can degrade before it is considered a functional failure depends on margin between the design capability and the required performance. A long P-F interval is desirable as it leaves time to plan and perform the preventive intervention. Service providers and spare part suppliers may also charge premium for short notice. If the P-F interval is too short, predictive or condition-based maintenance is impossible to implement. Rate of deterioration defines what condition

monitoring methods are needed to implement predictive or condition-based maintenance program. In this particular example the motor is considered failed before audible noise occurs, which is quite high performance requirement. In some other case failure threshold can be e.g. after hot to touch. For instance, in the study by Bengtsson et al. (2018) the functional failure was assumed to occur 1-4 weeks after audible noise and 1-5 days after hot to touch.

3.4.1 Corrective maintenance

Under corrective maintenance policy the component is repaired or replaced only after it has failed, which is the most primitive maintenance strategy. Benefit of corrective maintenance is that useful life is not wasted and unnecessary tasks are not performed. Usually this strategy is used for low importance components or components with backup. (Agustiady & Cudney 2016) It is an unacceptable strategy for failures with potential EHS consequences. If there is no unacceptable consequences, corrective maintenance is the optimal maintenance strategy for unpredictable random failures. Unpredictable random failures with EHS consequences must be managed with other means than maintenance, such as redesigns. Corrective maintenance is also optimal maintenance strategy for components with wear-out failures if expected financial consequences of a failure are less than the cost of implementing a preventive maintenance strategy. Corrective maintenance can be divided to immediate and deferred types.

Immediate corrective maintenance

Immediate corrective maintenance is performed as soon as possible after the failure has occurred. Failures of assets causing downtime are typically wanted to be corrected as soon as possible. Even though the asset would be desired to be corrected immediately, there might be unavoidable delays due to e.g. spare parts delivery and availability of labor. Also, suppliers may charge premium for short notice, which must be compared to consequences of prolonged outage.

Deferred corrective maintenance

In deferred corrective maintenance restoration or replacement of a failed asset is postponed to a better time. Failures with only non-operational consequences most often belong to this strategy. Deferring can be beneficial e.g. to prioritize more important cases, enable better

planning, ease budget issues and avoid short notice premiums. However, if maintenance task are deferred too much, backlog will begin to be an issue at some point.

3.4.2 Preventive maintenance

Preventive maintenance includes strategies that attempt to perform the maintenance tasks before a fault occurs to prevent it. Preventive maintenance is the only acceptable type of maintenance for functions which can cause EHS consequences in a case of failure. Preventive maintenance or at least scheduled inspections are needed also for functions whose fault can remain unnoticed during operation. In situations where consequences of a failure are acceptable, decision between corrective and preventive maintenance strategies can be made on purely economic basis.

Preventive maintenance strategies are further divided to three types based on how the maintenance activities are scheduled. The strategies are: time-based, condition-based and predictive maintenance, listed from the least developed to the most advanced.

Time-based maintenance

Time-based maintenance, or scheduled maintenance, is a maintenance policy in which tasks are performed at predetermined times without considering condition of the asset at the time of the task. Scheduled task can be restoration or replacement task. (Regan 2012, p. 122) In order to justify a time-based maintenance policy, it should be ensured that conditional probability of failure increases with age and that the scheduled maintenance improves reliability (SAE 2009). Time-based interventions are technically appropriate for assets which have recognizable useful life and wear-out failure modes dominate. However, in practice it is often difficult to define the maintenance interval and scheduled replacements can waste useful life. (Wang et al. 2007)

Condition-based maintenance

Condition-based maintenance is a maintenance strategy that is based on inspections or measurements that are performed to monitor how condition of the asset develops over time. Information of the asset's condition is obtained from periodical inspections or continuous condition monitoring measurements. (Regan 2012, p. 136) After inspection decision must be made whether to perform the preventive maintenance or postpone it. If the condition can

be quantified, e.g. by condition monitoring measurements, there can be a predefined preventive maintenance limit.

In order to implement condition-based maintenance, there must be a definable and detectable potential failure. Luckily, most failure modes give premonitory signs of impending failure. (Regan 2012, p. 134) However, P-F interval should be long enough such that there is time for inspection and intervention. (SAE 2009) Determining the inspection interval is a trade-off between the cost of inspection and the value of information gained. Long inspection interval increases the risk that potential failure is detected too late. One general rule of thumb is to use inspection interval that is half of the P-F interval. (Regan 2012, p. 138-141) Figure 16 presents the worst case scenario in which the inspection is performed just before the potential failure can be detected. V-belt is a good example of a component that fails predominantly due to wear.

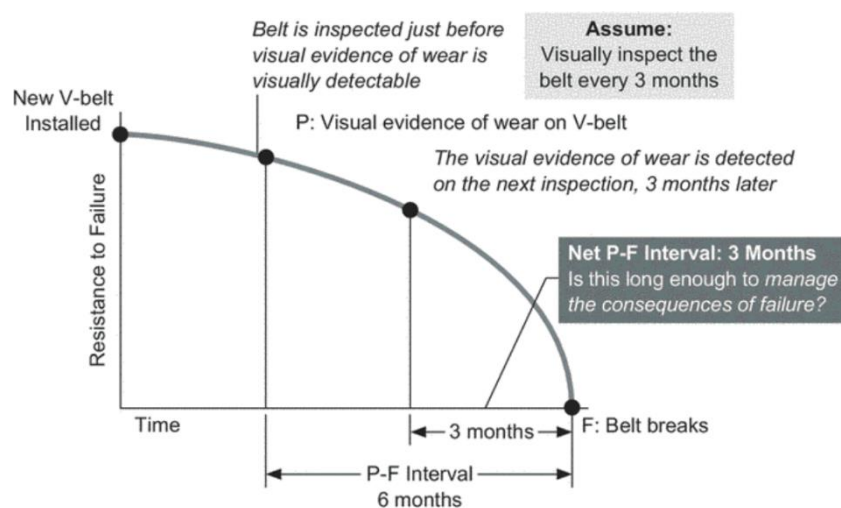


Figure 16. Illustration of the worst case scenario in periodic condition inspections. (Regan 2012)

Predictive maintenance

Predictive maintenance can be seen as a more advanced version of the condition-based maintenance. It is also based on monitoring the state of the asset, but instead of making decisions based on current state of the asset, prediction of the condition's development is used for decision making. In other words, condition-based maintenance uses diagnostics while predictive maintenance uses prognostics. Objective of the strategy is to detect impending failures early, and it is typically used with advanced sensing technologies. Advantages of the predictive approach are maximized asset uptime and delaying/reducing maintenance activities. Used predictive tools can be based on historical data like machine

learning techniques, integrity factors e.g. wear and visual aspects, statistical inference methods or engineering approaches. (Carvalho et al. 2019) Reliability and RUL estimation is an important part of predictive maintenance strategies as RUL model is a prerequisite for them. The recurring predictive maintenance process consist of consecutive steps: observation, analysis, and action. In optimal case these can be performed without human intervention, which enables short update interval. The ultimate goal in predictive maintenance, as described in Industry 4.0 concept, is components that are self-aware and self-predictive; machines that are able to self-compare in addition to being self-aware and self-predictive; and production systems that self-configure, self-maintain and self-organize. (Zonta et al. 2020)

3.5 Default strategies

There exist also other means to manage failures than maintenance. These other means are called default strategies in the RCM methodology. Default strategies can include improvements to operating procedures, redesigns of equipment, updates to technical publications, modified training programs, supply changes, enhanced troubleshooting procedures, and revised emergency procedures. (Regan 2012, p. 2-5)

Maintenance is an useful failure management strategy only for assets which exhibit wear-out failure patterns. As section 3.1.2 demonstrated, wear-out failures comprise only a minor part of the total failures. Since random failures cannot be effectively managed with maintenance, majority of the failure management must be performed by other means.

4 Reliability and remaining useful life estimation

Reliability and RUL models can be divided to four different categories: physics of failure models, statistical models, hybrid models and artificial intelligence (AI) models (Lei et al. 2018). Physics of failure models are used to model how specific failure mode develops over time. As these models are for one specific failure mode, complete reliability model for an asset with many failure modes would require multiple models considering interrelations of the failure modes. From this it can be reasoned that physics of failure models are most useful for assets which have single or at most a few dominating failure modes. The models require knowledge about physical characteristics of the degradation process. However, physics of failure models don't necessarily need data from the asset's end of life, which is advantageous especially for new products since data from end of the life doesn't exist. Physics of failure models can be considered white-box models since cause-and-affect relations are expressed analytically. (Thaduri et al. 2013) This is advantageous as the obtained knowledge can be used to prevent future failures.

Data-driven statistical models make predictions based on data, which can be event data, continuous condition monitoring or periodical inspections. Statistical models can be more general than failure mode specific physics of failure models, and they don't necessarily require expertise about physics of degradation. Thus, statistical models can be considered as black-box models since cause-and-affect relations leading to a failure are typically not included. (Thaduri et al. 2013) However, a FMEA can be performed as a complement. Further, if there is enough data with enough details, separate statistical models could be built for each failure mode. This way shares and failure patterns of different failure modes could be studied. If this modelling approach is used, it is important to notice that failure modes may be correlated i.e. one failure mode can accelerate other failure modes.

Copula theory can be used to consider the correlations among competing failure modes. According to the theory, a joint probability distribution can be expressed with a copula function which is a function of marginal probability distributions i.e. the copula function connects marginal distributions to the joint distribution. Copula functions can express nonlinear correlations and asymmetric tail dependencies. Also time-varying copulas can be implemented since correlation characteristics can vary over time. For example, Sun et al. (2018) implemented a time-varying copula to study crack propagation. Dynamic modelling

is advantageous in this case as there is two competing failure modes: fatigue failure and static strength failure, and the correlation characteristics between the failure modes change as the crack grows.

Hybrid models combine properties of both statistical and physics of failure models. The most recently founded category is AI models. AI models are trained with large datasets to recognize degradation patterns from condition monitoring data. They are attractive since they are capable of dealing with complex systems whose degradation processes can have difficult interrelations. However, the result of AI models may be hard to explain as they lack transparency. (Lei et al. 2018) Hence AI models can be considered black-box models.

Nonetheless, more advanced modelling requires more time, effort and money. Thaduri et al. (2013) studied life cycle costs of different reliability prediction methods regarding electronic components. The least demanding estimation method is using existing frameworks and statistics. For example, MIL-HDBK-217F Military Handbook: Reliability Prediction of Electronic Equipment (U.S. DoD 1995) has been widely used yet nowadays it is largely obsolete. General failure rate statistics are typically based on assuming a constant failure rate. For mechanical equipment, there is for example Handbook of Reliability Prediction Procedures for Mechanical Equipment (U.S. NSWC 2011). Both of the handbooks include means to consider external factors e.g. environment and way of use. These external factors typically affect reliability by accelerating wear-out, which contradicts the constant failure rate assumption. Hence, the U.S. Department of Defense (DoD) and National Aeronautics and Space Administration (NASA) have abandoned handbook statistics and recommend the physics of failure approach. Handbook statistics are best suited for estimating the number of failures in populations that consist of equipment at different stages of life. (U.S. National Research Council) (Pecht et al. 2020) The U.S. DoD has had a significant contribution to reliability engineering as a discipline. Empirical tests with real equipment require substantial resources which only a few possess.

Conventional statistical modelling per se is relatively straightforward, but the data acquisition and preparation is often labor intensive. Stochastic and AI modelling are more demanding in terms of computing power. Depending on complexity of the equipment and number of relevant failure modes, physics of failure is typically the most laborious approach. (Thaduri et al. 2013) However, it leads to a deeper understanding of the assets and can

provide valuable insights. Physics of failure is also the only method which can predict new phenomena as the other methods rely on past observations.

Dependence on the past observations can be tried to overcome by testing. In accelerated life testing the component is subjected to more demanding conditions than during normal usage. Examples of these conditions are higher than usual stress, strain, pressure, vibration, temperature, voltage or some combination of them. It is important to notice that the dominant failure mode may change with the conditions. A model is fitted to the data to describe relation between subjected conditions and lifetime. The model can then be used to estimate lifetimes under normal conditions. There are established acceleration models for some conditions e.g. high temperature fatigue and temperature cycling. Accelerated life testing has been studied by, for example, Nelson (1980, 2009) and it is not further studied in this thesis. Accelerated life testing belongs more to the manufacturers side of reliability assessment, and costs associated to it are very high.

As a conclusion, the effort put into reliability estimation should be assessed relatively to the value of information gained. High-effort estimates may not be justified for non-critical assets, but some rough yet rational estimate can give valuable information with low-effort. In turn, if it is recognized that failure consequences may be severe, it can be sensible or even mandatory to put considerable amount of effort into the reliability estimation. This chapter introduces main types of models in aforementioned categories.

4.1 Physics of failure models

Physics of failure is a reliability modelling approach, which is based on the knowledge of root cause failure mechanisms. It is a up-front approach that can be used also during design phase to improve inherent reliability. History of the approach begins from 1962, when U.S. Air Force was unsatisfied with low reliability of electronic systems, and initiated investigations of degradation mechanisms. (Sadiku et al. 2016) Advantages of physics of failure approach are that operating experience is not required, and that the obtained knowledge can be useful for preventing future failures. It is especially useful for an equipment manufacturer designing a new product.

However, physics of failure is essentially a bottom-up approach. It implies that every relevant failure mode should be recognized and modelled separately to comprise the total

failure rate. Consequently, required effort of this approach increases with complexity of equipment. When number of failure modes increases, it gets more difficult to adhere to the mutually exclusive and completely exhaustive principle while also considering interrelations between the failure modes. Hence, physics of failure models are most suitable for relatively simple equipment.

Physics of failure models typically model wear-out failures, which can be caused by e.g. creep, fatigue, thermal degradation, corrosion, erosion or cavitation. A review by Lei et al. (2018) studied where physics based approaches are applied in the RUL prediction of machinery. One of the most used models was Paris-Erdogan model, which was used to model crack growth in e.g. gears and bearings. Other examples are Forman crack growth law to predict the RUL of cracked rotor shafts, and Norton law to estimate creep evolution of turbines. Recognized physical contributors to failures of electronics are thermal stress, thermal cycling, mechanical stress and humidity. Common laws used in RUL prediction of electronics are Arrhenius equation for thermal degradation, Norris-Landzberg equation for solder joint fatigue, Basquin's equation for vibration and Peck's equation for humidity (FIDES 2009).

4.2 Statistical models

Statistical data-driven models can be divided to different categories based on different characteristics. These different ways to categorize model types include covariates, condition state space, input data, and stochasticity. Basic non-parametric and parametric models, i.e. survival analysis, estimate reliability only as a function of time, and some models like discrete Markov chain model can use condition as the only input. On the contrary, it is also possible to include multiple covariates to consider effect of operational and environmental factors. For example, basic survival analysis can be expanded with Cox proportional hazards regression analysis. In survival analysis the asset's condition is binary: it is either alive or failed. Stochastic models have different approach: they assume that degradation progresses through small random increments which accumulate over time. Hence, the condition state space is continuous, or it can be discretized for simplicity.

Input data for statistical models can be divided to three categories: event data, continuous condition monitoring and periodical inspections. Event data based survival analysis is the traditional approach to reliability estimation. However, downside of the approach is

availability of such data. Continuous condition monitoring data and periodical inspections can be further divided into two sub-categories depending on if the measurements give direct or indirect information about the ongoing degradation process (Si et al. 2011). Direct condition monitoring measurement gives explicit information about the degradation level of the asset. An example of direct condition monitoring is crack size measurement. Indirect condition monitoring measurement does not explicitly indicate certain degradation level, and consequently failure threshold is more difficult to determine for indirect condition monitoring (Si et al. 2013). However, relation between the measurement and actual degradation level can be modeled. Indirect measurements are used since degradation can be hard to monitor directly, especially during use. Examples of indirect condition monitoring are oil and vibration monitoring. In case of a journal bearing, surface wear is the actual degradation process, of which oil impurities and vibration level may indicate. It is important to notice the difference between correlation and causality when using indirect condition monitoring or regression modelling. For example, metal concentration of engine oil is a good indicator of an engine's wear, but some other contaminants in the oil may affect wear rather than indicate wear. (Si et al. 2011)

Degradation process and failure occurrence is stochastic by nature. Thus, the more advanced models introduce random variables to consider inherent stochasticity of the underlying degradation processes. Degradation process of a population has two types of uncertainties: sampling and temporal. Sampling uncertainty refers to variation in degradation level between samples within the population. Temporal uncertainty refers to inherent uncertainty related to degradation process over time. Sampling uncertainty can be managed by increasing inspections, but temporal uncertainty cannot be eliminated as degradation process is aleatory by nature. Stochastic degradation models can be divided to two broad categories: random variable models and stochastic process models. Random variable models consider sampling uncertainty but cannot consider temporal uncertainty as specific sample path is deterministic in the random variable model. Stochastic process models incorporate both sources of uncertainty. (Pandey et al. 2009)

Downside of statistical models is data availability. Critical components are rarely allowed run to failure so fault data can be scarce (Si et al. 2011). For new components data does not exist at all and for components with short lifespans phase out may already be happening when there is enough data acquired for reliable model. This is challenging especially for

rapidly developing fields of technology, where certain type of a component is manufactured only for a short period of time until a new model is released. This issue can be tried to overcome with accelerated life testing or utilizing data from closely related products.

This subchapter presents the most common statistical reliability models in order of increasing complexity starting from the established basic models. The models are briefly introduced, and their suitability to different cases is discussed.

4.2.1 Non-parametric models

Non-parametric models are methods used to construct a survival curve. Contradictory to parametric models, non-parametric models do not make assumptions of the shape of the survival curve. As a result, survival curves obtained from non-parametric models are step functions. Non-parametric and parametric modelling both belong to survival analysis, and they are often performed sequentially. Basis of the non-parametric reliability modelling is to have event data from a population of assets. The event data should contain commissioning and decommissioning times as well as failure times. Times to failure or censoring can then be calculated from the event data. A data point is right censored if the asset is alive when the review period ends or if it is decommissioned alive during the review. An important condition is that censoring should be independent of the probability of failure. (Klein et al. 2014, p. 5-6) For example, if an asset is decommissioned alive, but the reason for decommissioning is upcoming failure, the censoring is not independent of the probability of failure. If this happens systematically, it will lead to a biased estimate.

There are two methods to construct a survival curve from event data: life table method and Kaplan-Meier estimator. As mentioned, non-parametric survival curve is a step function. Hence the points where the function is evaluated will define shape of the curve. Main difference between the methods is evaluation points. Life table uses equally spaced intervals, while the Kaplan-Meier uses event and censoring times. Kaplan-Meier estimator is better, or at least less subjective, in the sense that the result does not depend on chosen intervals. (Sullivan 2016) Kaplan-Meier estimator is preferred in reliability engineering, and life tables are used in actuarial sciences. Kaplan-Meier estimator of the survival function is given:

$$\hat{S}_{KM}(t) = \prod_{t_i \leq t} \left(1 - \frac{d_i}{n_i}\right), \quad (8)$$

where t_i is a time when at least one event occurred, d_i is the number of events that occurred at time t_i and n_i is the number of units at risk at time t_i .

4.2.2 Parametric models

Parametric model is often fitted to a non-parametric model, but it can be also built without event data, based on e.g. expert judgement. Parametric model is needed to obtain a proper estimate of the hazard function since the hazard function is not well defined for non-parametric models. The hazard function is well defined only for differentiable survival curves, and the non-parametric survival curve is non-differentiable as a step function. (Klein et al. 2014, p. 7) Parameter values can also give insights of failure patterns.

Parametric models make assumptions about shape of the survival curve. Thus, justification of the chosen model must be considered. Sometimes there are theoretical arguments to support the choice of model, but often the model is chosen solely based on its ability to fit into the data. The U.S National Institute of Standards and Technology (U.S. NIST 2013) has given three rules to justify the choice of a model:

1. There is a physical/statistical argument that theoretically matches a failure mechanism to a life distribution model
2. A particular model has previously been used successfully for the same or a similar failure mechanism
3. A convenient model provides a good empirical fit to all the failure data

Physical/statistical arguments include extreme value argument, multiplicative degradation argument and fatigue life argument. Lifetime distribution models that can be derived from aforementioned arguments are respectively Weibull distribution, lognormal distribution and Birnbaum-Saunders distribution. However, these arguments are typically valid only for individual failure modes, which complicates use of the first rule of justification when modelling complex equipment. Hence, U.S. NIST (2013) recommends modelling failure modes separately, and using bottom-up approach to model the system. The third rule requires considerable amount of data. Regardless of the rules of justification, the model should pass visual and statistical tests, and make sense in general. However, visually reasonably similar cdfs can have completely different hazard functions as shown in Figure 17. Therefore, it is desirable to have some knowledge regarding physics of failure to

understand e.g. whether the failure rate is monotonically increasing or unimodal. This is especially important when extrapolating beyond range of the data.

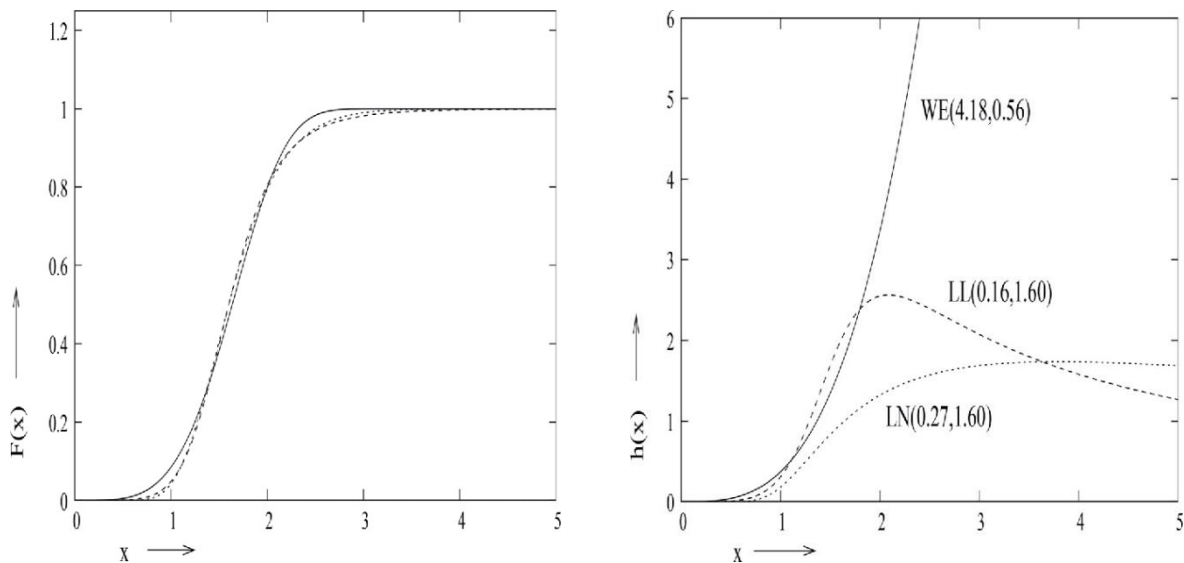


Figure 17. Cdfs and hazard functions of a Weibull, log-normal and log-logistic distributions. Visually quite similar cdfs can have completely different hazard functions, which emphasizes the importance of choosing the right model. (Raqab et al. 2018)

Two main methods for parameter estimation are maximum likelihood estimation (MLE) and rank regression, also known as the least squares method. Earlier, parameters have been estimated graphically, but nowadays numerical methods provided by computer software are preferred. MLE finds the distribution that is most likely to produce the data, and rank regression minimizes the squared distances between the data points and the fitted function. (Letcher 2017, p. 317-318) MLE is better in estimating from censored data, and it approaches the unbiased minimum variance estimator as the sample size increases. However, with small numbers of failures (less than 10) MLE can be heavily biased. Hence, rank regression may be better with small complete datasets. (U.S. NIST 2013) As event data typically contains censoring, MLE is therefore generally preferred.

Exponential distribution

Exponential model is the simplest life distribution model, which can lead to its use in also inappropriate situations. It has memoryless property i.e. failure rate is constant and does not depend on age. Consequently, pdf of the exponential function has no shape parameter. (Kececioglu 2002) It is thus suitable distribution only for situations where observed failure

rate equals random failure rate, and there is no infant mortality or wear-out failures. Survival function of a 2-parameter exponential distribution is given:

$$S(t) = e^{-\lambda(t-\gamma)}, \quad (9)$$

where λ is the rate parameter and γ is the location parameter (or failure free life). Rate parameter defines the failure rate which is the inverse of the MTBF. Location parameter can be used to move the distribution so that probability of failure is zero before earliest possible time of failure. The location parameter has the same unit as the x-axis e.g. hours, cycles or kilometers. In practice there is often no such a thing as failure free life i.e. $\gamma = 0$, and 2-parameter distribution reduces to 1-parameter distribution. Also all other parametric models can be expanded with location parameter by replacing t with $(t - \gamma)$.

Weibull distribution

Weibull distribution is one of the most used distributions in reliability engineering since it is very versatile (Kececioglu 2002). By changing parameter values Weibull distribution can be used to model all three sections of the bathtub curve i.e. it can model decreasing, constant and increasing failure rates. A single Weibull distribution is monotonic, but mixture Weibull distributions can be used to model the bathtub curve or other non-monotonic hazard rates.

Weibull (1951) himself stated that the function has no theoretical basis but empirical tests have proved it to work well in many cases. However, theoretical justification for Weibull distribution was later found from Gumbel's (1958) work with theory of extreme value distributions. Extreme value distributions are limiting distributions for minimums or maximums of a sample of independent and identically distributed random variables. Applying theory of extreme values to reliability engineering can be justified through the assumption that modelled system consists of n independent components in series i.e. the system fails when the first component fails. Then system failure time is the minimum of n component failure times, and distribution of system failure times approaches Weibull distribution when n approaches infinity. The same rationalizing can be applied to a single component level, if it assumed that component failure occurs when the first of many potential degradation processes develops into a failure. This may explain why the distribution has been successful in applications like ball bearing, relay, capacitor and material strength failures. (U.S. NIST 2013) However, independency condition may not be

met since failure modes can be correlated as discussed earlier. Survival function of a Weibull distribution is given:

$$S(t) = e^{-\left(\frac{t}{\eta}\right)^{\beta}}, \quad (10)$$

where η is the scale parameter (or characteristic life) and β is the shape parameter (or slope). Scale parameter scales the function i.e. when it increases the pdf becomes wider. It has the same unit as the x-axis. Scale parameter is inverse of the rate parameter used in exponential distribution. It can be interpreted as the approximate 0.368 quantile of the survival function i.e. the time at which 63.2 % of the population has failed.

In Bayesian framework it is desirable that parameter values can be intuitively interpreted such that experts of the field in question can provide the priors without possessing expertise on probability calculus. Often in reliability applications share of early failures e.g. 0.1 quantile of the cdf is the quantity of interest. In such cases reparameterization can be used to ease the prior setting and interpretation of results. (Li & Meeker 2014) In contrast to exponential distribution, scale parameter does not solely determine failure rate of a Weibull model as also the shape parameter affects it.

The shape parameter defines trend of the failure rate. $\beta < 1$ corresponds to decreasing failure rate, while $\beta = 1$ means that failure rate is constant, and $\beta > 1$ means that failure rate increases with age. When $\beta = 1$ Weibull distribution becomes exponential distribution i.e. exponential distribution is a special case of the Weibull distribution. The shape parameter has no unit.

Lognormal distribution

Lognormal distribution is also a very flexible distribution that can be empirically fitted to many types of failure data. Lognormal is theoretically justified model choice for failures caused by degradation that progresses at a rate proportional to the total amount of existing degradation. Justification arises from multiplicative degradation argument. The argument is based on assumption that degradation progresses in small random increments which increase the amount of degradation in multiplicative manner. If this assumption is accepted, lognormal distribution can be derived using central limit theorem. (U.S. NIST 2013)

Failure mechanisms which can exhibit this kind of behavior include corrosion, diffusion, crack growth, migration, electromigration, and generally failures resulting from chemical reactions. These degradation processes are recognized to be common to semiconductors. (U.S. NIST 2013) As name of the distribution tells, it is closely related to the normal distribution: a random variable is lognormally distributed if its logarithm is normally distributed. Lognormal survival function does not have a closed-form solution. The survival function can be presented in the following form:

$$S(t) = 1 - \Phi\left(\frac{\ln t - \ln T_{50}}{\sigma}\right), \quad (11)$$

where $\Phi(x)$ denotes the standard normal cdf, T_{50} is the median (or scale parameter) and σ is the shape parameter. If failure times are lognormally distributed, then natural logarithms of the failure times are normally distributed with mean $\mu = \ln T_{50}$ and standard deviation σ .

Birnbaum-Saunders distribution

Birnbaum-Saunders distribution is a life distribution model specifically developed to model physical fatigue processes where crack growth causes the failure, and it is hence called also fatigue life model. Derivation of the model begins with the assumption that crack growth resulting from each stress cycle is a random amount independent of the past cycles. It is noteworthy that this assumption is very different from the multiplicative degradation argument. However, the Birnbaum-Saunders assumption is consistent with a deterministic materials physics model called Miner's rule. If the assumption is accepted, the Birnbaum-Saunders distribution can be derived using central limit theorem with an additional assumption that since stress cycles occur often and last for a short time, number of cycles can be replaced by time. (U.S. NIST 2013) Survival function of the Birnbaum-Saunders distribution can be given:

$$S(t) = 1 - \Phi\left(\frac{1}{\beta}\left(\sqrt{\frac{t}{\eta}} - \sqrt{\frac{\eta}{t}}\right)\right), \quad (12)$$

where $\Phi(x)$ denotes the standard normal cdf, η is the scale parameter and β is the shape parameter.

4.2.3 Regression models

While the basic survival models estimate reliability only as a function of time, regression models can use additional explanatory variables to consider also other factors that may affect reliability. Proportional hazards model is probably the most used regression model in survival analysis. Other approaches to model survival data with covariates are parametric regression, accelerated life and proportional odds models. (Rodriguez 2010) Regression models include also random coefficient regression and autoregressive models.

Cox proportional hazards model

The proportional hazards model, originally proposed by Cox (1972), is used to estimate effect of explanatory variables on survival. It is more of an acceleration model than a life distribution model (U.S. NIST 2013). The model is called semi-parametric since it does not make assumptions about the shape of the baseline hazard function. There are however, assumptions for appropriate use of the model. A fundamental assumption of the model is that changes in explanatory variables have multiplicative effect on the hazard rate. Consequently, hazard curves for different groups are proportional and cannot cross. (Sullivan 2016) In the standard model it is also assumed that the effect of the external variables does not depend on time. Under these assumptions the failure rate can be presented as a product of two functions:

$$h(t, X) = h_0(t)g(\bar{X}, \bar{A}), \quad (13)$$

where $h_0(t)$ is an arbitrary baseline hazard function, $g(\bar{X}, \bar{A})$ is a time-independent covariate function, $\bar{X} = (x_1, x_2, \dots, x_n)$ is a row vector consisting of the explanatory variables and $\bar{A} = (a_1, a_2, \dots, a_n)^T$ is a column vector consisting of the regression parameters. The baseline hazard is obtained when all of the explanatory variables are equal to zero. The explanatory variables can be either continuous like temperature or binary variables indicating if a condition is present or not. The most usual form of the model is log-linear i.e. $g(X, A) = e^{XA}$.

Random coefficient regression

In conventional regression models it is assumed that all individuals in the data set come from a population with a single slope and intercept parameters. In random coefficient regression this assumption is relaxed by allowing the parameters to vary between individuals and then to be predicted by other covariates. This type of models are called also multilevel

or hierarchical models. The varying parameters are called random coefficients, and they are unobserved continuous variables. (Muthen et al. 2015) In reliability analysis random coefficients are used to describe stochasticity of degradation processes.

Autoregressive models

Autoregressive models (AR), or more generally autoregressive integrated moving average models with an exogenous variable (ARIMAX), are widely used to model and predict time series data. ARIMA models make predictions based only on the past observations of the forecasted variable. This approach is based on assumption that future values are linear functions of past values and random errors. ARIMAX is an extension to ARIMA model to consider also effect of other predictor variables. ARIMAX models are typically effective for short term predictions, but less useful for longer timescales due to dynamic noise, sensitivity to initial system conditions and accumulation of systematic errors. (Sikorska et al. 2011)

According to review by Lei et al. (2018), AR models and its extensions have been used to model e.g. degradation of bearings, RUL of aluminum plates with fatigue crack and RUL of a conveyor belt system. In some of the studies AR models have been combined with particle filtering algorithm to predict RUL of machinery.

4.2.4 Stochastic models

Degradation can be characterized as a process that progresses with small random increments which accumulate over time. Stochastic RUL models take this approach and are thus called also cumulative damage models (Nakamura & Nakagawa 2010). This approach is closely related to a mathematical object called random walk and particle trajectory modelling in physics. The continuous degradation process may be discretized in time and/or state space to simplify the modelling.

Many conventional stochastic processes share a characteristic called Markov property. Markov property denotes that next state of the process depends only on the current state independently from the earlier states i.e. the process is memoryless. Figure 18 provides an example of two different degradation paths with same current degradation level.

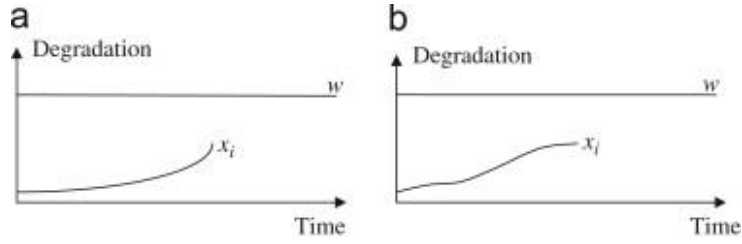


Figure 18. Two different degradation paths with same current degradation level. Models with Markov property estimate same RUL for both. (Si et al. 2013)

When the Markov property holds, a model gives same RUL estimate for both (a) and (b), even though intuitively (a) would be expected to degrade and fail faster. Markov property is a strong assumption for a physical process, and thus its justification must be considered carefully on a case-by-case basis. Often it is desirable to utilize the whole degradation data to date. Therefore, conventional stochastic processes have been further developed to condition the pdf of next state on earlier observations. One method to capture the entire history of observations to the estimate is Kalman filtering. However, Kalman filter is best suited for smooth curves and degradation process may have sudden changes or jumps. Thus, also Kalman filter has been further developed into recursive filters which are better in capturing characteristics of physical degradation processes. (Si et al. 2013)

Stochastic models typically utilize first hitting time approach to model the RUL. It is worth noting that theoretically the first hitting time approach is justified only for monotonic processes, and certain incipient failures have self-healing characteristics. For example bearings can have non-monotonic degradation paths due to temporarily self-healing nature of the degradation process. Initial surface damage causes increased vibration, then damage is smoothed by rolling movement, which decreases the acceleration for some time. However, damage will spread and eventually cause the bearing to fail. (Lei et al. 2018) Even though the degradation processes may not be purely monotonic, it is not a problem in practice since self-healing characteristics are temporary and the degradation will progress inevitably.

Markov models

Markov models are stochastic processes that possess the Markov property and in which each state corresponds to an observable event. Markov models can be divided to Markov chains, which have finite number of states, and Markov processes, which are continuous with infinite number of states. By defining probabilities associated with states and transitions from a state to another, probability of ending up to a failed state can be estimated. Markov

models can be further divided to discrete and continuous time processes. In discrete processes the waiting time between jumps is fixed, and in continuous processes the waiting time is exponentially distributed. Semi-Markov process is an extension to Markov process, which allows that the time spent in a certain state is not exponentially distributed. Hidden Markov models are an extension of Markov models to cases where all states are not directly observable. (Sikorska et al. 2011)

Wiener process models

Wiener process models are the most used stochastic process models in reliability estimation. A linear Wiener process is used to model degradation processes where the degradation progresses linearly with random noise. However, also nonlinear Wiener process based degradation models have been developed. The linear model consists of a drift term, which characterizes the degradation rate, and the Brownian motion noise term characterizing the stochasticity of the degradation process. The mean degradation path of a linear Wiener process is linear function of time. (Lei et al. 2018) The linear Wiener process can be represented as:

$$X(t) = \lambda t + \sigma B(t) , \quad (14)$$

where $X(t)$ is the degradation level at time t , λ is the drift coefficient, σ is the diffusion coefficient, and $B(t)$ is the standard Brownian motion.

In conventional Wiener process models the pdf of RUL depends only on the latest observed degradation level neglecting the earlier observations, which corresponds to the Markov property in Markov chain models. Si et al. (2013) proposed a new Wiener process based degradation model, which conditions the pdf of RUL on an entire history of observations, making the estimate path dependent.

Gamma process models

Gamma process is a continuous time Markov process with independent gamma distributed increments. Gamma process is a limiting compound Poisson process with Gamma distributed increments. The limit is reached when jump rate of the Poisson process approaches zero and size of the increments approaches zero. In other words, Gamma process models are based on an assumption that degradation progresses through very small but

frequent increments. Examples of such processes where damage accumulates gradually over time are creep, corrosion and wear. (Mahmoodian & Li 2016)

Inverse Gaussian process models

Inverse Gaussian process is a stochastic process, which progresses through inverse Gaussian distributed independent increments. It is less common than the Wiener and Gamma processes. Alike the Gamma process, also inverse Gaussian process is a limiting compound Poisson process, though with different jump size distribution. It has similar properties to the Gamma process, but is more flexible when it comes to incorporating additional random effects and explanatory variables. Use of the inverse Gaussian process can be justified through its relation to the Birnbaum–Saunders distribution. In particular, cdf of the first hitting time can be approximated as a Birnbaum–Saunders-type distribution, and with a certain parameter values the cdf reduces to a Birnbaum–Saunders distribution. (Ye & Nan 2014)

4.3 Hybrid models

Hybrid models, which are also called fusion models, combine different sources of information in order to give better predictions. Combined information can be e.g. statistics and physics of failure. One type of these models is health index models, in which the model is initially build based on statistics and physics of failure. Then asset reliability can be estimated based on its observed health condition. For example, Common Network Asset Indices Methodology is an industry standard health index model used by British distribution system operators. (UK OFGEM 2017)

Also Bayesian models utilizing prior expert knowledge can be seen as hybrid models. For example, Zhao et al. (2018) developed a multisource fusion Bayes model which utilizes expert knowledge, lifetime and degradation data. The degradation is modelled with linear Wiener process and inverse Gaussian distribution is used to describe the lifetime data. Expert knowledge provides priors for the models. Bayesian approach is efficient for small sample sizes. Even though it is desirable to utilize all available information, it must be noticed that prior information of a subjective nature can be biased and compromise the whole modelling effort. Sorrowfully yet realistically personal interests may override the desire to give impartial estimates. Importance of sensitivity analysis is emphasized when the model includes subjective estimates.

4.4 Artificial intelligence models

In the context of reliability and RUL estimation, artificial intelligence (AI) refers to learning algorithms which are trained with past data to predict the future failures. The algorithms can be divided to two main classes based on the training dataset: supervised and unsupervised algorithms. Unsupervised methods use process information without maintenance records and are mostly based on outlier detection algorithms. Methods used in unsupervised learning are auto-encoders, deep belief networks and statistical analysis. (Moleda et al. 2020)

In supervised approaches the training dataset includes information on the occurrence of failures in addition to process information. Supervised algorithm learns a function that maps inputs to outputs based on input-output pairs in the training dataset. Optimally the algorithm will map also unprecedented inputs correctly after the training. Supervised models can be further divided to two classes: classification and regression models. Classification models are used to predict categorial labels, and regression models are used if the result are continuous. Downside of classification is that it may require manual work and expert interpretation. AI techniques used in RUL prediction with supervised learning algorithms include artificial neural networks, k-nearest neighbors, support vector machines and Bayesian networks. (Moleda et al. 2020)

AI algorithms often use dimensionality reduction methods to transform the data from high-dimensional space to low-dimensional space while retaining maximal amount of meaningful information of the original data. Principal component analysis (PCA) is a common linear technique used for dimensionality reduction. Basic idea of the PCA is to find new variables that are linear functions of the original variables. New variables should be uncorrelated to maximize variance and preserved information. Finding such variables reduces to an eigenvalue/eigenvector problem. (Jolliffe & Cadima 2016)

Carvalho et al. (2019) made a review on applying machine learning methods to predictive maintenance. The study revealed that the most employed machine learning algorithm is Random Forest algorithm, which was used in 33 % of the analyzed research papers. The algorithm operates by constructing numerous decision trees during the training, which then work as an ensemble. It was used to model e.g. degradation of wind turbines from status and operational data, degradation of squirrel-cage induction motors from current and voltage

waveforms and degradation of industrial pumps from vibration data. Second most common was neural network based methods with 27 %. Neural networks are computational learning algorithms that use networks of functions to translate inputs to desired outputs. The algorithm tries to mimic the functioning of the human brain. Neural networks were used to model e.g. wind turbine failures from accelerometer data and failures of electrical power systems from electrical signals. Support vector machine was used in 25 % of the papers, and it was used to predict e.g. gearbox and bearing failures from vibration signals. K-means algorithm was the fourth most common with 13 %, and it was used for example to model oil immersed power transformer failures from dissolved gases concentrations. Vibration signal was the most used data for predictive maintenance.

5 Case Study: Programmable logic controller and human machine interface

This chapter consists of a qualitative FMEA and quantitative reliability models for PLC & HMI. The FMEA is performed based on literature sources. Also failure rates and shares of different failures are searched from the literature and manufacturers. After that, failure data is acquired from maintenance records, and survival analysis is performed. Lastly, event tree models are built and tested.

5.1 Failure mode and effect analysis

Primary function of a PLC is controlling process by sending actuating signals to an actuator. Failure modes depend on chosen level of abstraction. The level of abstraction defines the level of detail i.e. what is the smallest unit considered as a separate entity. The lowest level of abstraction could be e.g. analyzing each transistor separately, and the highest level would be the PLC. This study focuses mainly on the module level. However, regardless of the level of abstraction, failure modes of a PLC can be divided to five following categories: detectable/preventable failures, age-related failures, random failures, random sudden failures and intermittent failures (Korsah et al. 2010). Random sudden failures are categorized separately to make a distinction from random failures where gradual degradation can be observed.

Failures of electronic systems have sometimes been modelled in an oversimplified manner. Notable example is an assumption that electronic signal will fall to zero in case of a failure. For example, in common 4-20 mA range the signal may fall to zero e.g. if the connection breaks, but it may as well rise to 20 mA (or above), lock to intermediate value or slowly drift away from the correct value. (Lindsley et al. 2018, p. 237) If this is not properly taken into account, potentially dangerous control systems may exist.

Independent of the level of abstraction, functional failure of a PLC can have four different failure effects. Total failure means that generation of outputs ceases. Total failure can be haphazard, in which outputs freeze to unpredictable states, or ordered, where outputs are set to pre-defined supposedly safe values. Fail-safe designs try to achieve the ordered failure behavior. Partial failure of a PLC means that generation of outputs does not cease but generated outputs are incorrect. Partial failures can be distinguished to failures with

plausible and non-plausible behavior. Plausible behavior failure refers to a situation where generated outputs are wrong but external observer cannot detect it. Non-plausible behavior means that it is evident that the outputs are not correct. (Authén et al. 2015) Failure consequences cannot be assessed on general level since they depend on the controlled process.

HMI discussed in this study refers to a local HMI of a remote controlled control system. The HMI is less critical than the PLC as it is not necessarily needed during normal operation, if the remote control is not through the HMI. Primary function of the HMI is to allow the operator to supervise and control the process locally. The HMI can have three types of failures: failure to generate outputs to the user e.g. broken display, failure to receive user inputs e.g. broken keyboard and failure to communicate with the PLC. Apart from input and output devices, PLCs and HMIs suffer from similar failures as digital computers. However, PC-based HMI can have a fan and a hard drive, which are subject to wear-out as mechanical devices.

The FMEA is essentially a qualitative analysis method. In quantitative reliability analysis failure modes are often reduced to a generic "fails to function" failure mode. This simplifies the modelling, and it may be difficult or impossible to obtain separate failure rates for every failure mode. Generic failure databases and manufacturers typically give failure rates for the modules or components with no further specifications. An asset owner could of course collect data from its own operations with the desired level of detail. Nonetheless, the value created by a FMEA is not lost even though it may not be fully utilized in the quantitative part of the study as it can give valuable insights of root causes of the failures and knowledge of risks that the organization is exposed to.

An important part of FMEA is to consider how faults are detected. To underline this importance, Goble & Brombacher (1999) have presented an extended concept of Failure modes, effects and diagnostic analysis (FMEDA) to study diagnostic coverage of failures in programmable electronic systems. Diagnostics is an important part of reliability analysis especially when considering protective devices. Faults not covered by the online diagnostics can lead to a dangerous situation where hidden fault exists in the system. Different ways in which a fault can be detected are: online detection, offline detection, spurious action, latent revealed by demand and triggered by demand. (Authén et al. 2015) Online detection

contains various continuous measurements and offline detection includes e.g. periodic testing. In context of safety related systems failures are often divided to safe and dangerous failures. Failures detected by diagnostics should lead to a safe state when fail-safe design is used, meaning that undetected faults are the more dangerous ones.

This study applies methods from probabilistic risk assessment (PRA) of nuclear power plants as the nuclear industry has the most advanced PRA. However, nuclear safety related control systems have certain different characteristics compared to conventional systems. These characteristics include redundancies, voting logics and fail-safe designs. By taking these differences into account, research made for the nuclear industry can be utilized in reliability estimation of regular non-redundant PLCs.

The FMEA presented in this chapter is based on reports by Authén et al. (2015), Korsah et al. (2010) and Ostenso & May (1996). Authén et al. (2015) provide guidelines on how to analyze and model digital I&C systems in the context of PRA. The guidelines include failure mode taxonomy for both hardware and software failures. Chosen level of abstraction has significant effect on required effort. In the context of protection systems, following levels of abstraction are distinguished: system, division, I&C unit, module and basic component level. Module level is recommended in the report.

Main object of the study by Korsah et al. (2010) was to develop an unified framework for failure modes and mechanisms of digital I&C systems in nuclear power plants. The study was conducted by investigating operating experience databases. Poor quality of data in addition to funding and time constraints prevented achievement of the main objective. Despite the disappointment, some useful insights were still gained. Failure modes were divided to categories based on their characteristics, and shares of different categories were obtained. Table 3 presents shares of different failure mode categories. The categorization is based on 100 analyzed events acquired from Equipment Performance and Information Exchange Database. 35 % of the failures involved a PLC. Application-specific integrated circuits were involved in 8 % of the failures and field programmable gate arrays were involved in 3 %. Less than 10 % of the failures were software related. In many cases the cause of failure was not specified or could not be identified. As there were statistically insufficient number of events for each device type separately, the classification was made jointly.

Table 3. Shares of different failure mode categories in digital I&C systems according to Korsah et al. (2010).

Failure mode category	Share
Detectable/preventable failures	34 %
Age-related failures	23 %
Random failures	21 %
Random sudden failures	19 %
Intermittent failures	2 %

Detectable/preventable failures could have been possibly prevented with e.g. online monitoring, exhaustive testing prior to installation, adequate configuration control or verification & validation practices. For example, failures caused by “operating outside of specification” belong to this category. As this was the largest category, there is room for improvement. Majority of the age-related failures were caused by power supplies or components related to power supplies. Usual failure modes were degraded output voltage or complete power supply failure. 23 % is a non-negligible share of failures, which questions justification of the common constant failure rate assumption. Random failure categories include failure modes that did not seem to have any pattern or recurrence. Random sudden failures are separated to emphasize rapid occurrence of these failures in contrast to gradual degradation. Most of the causes of random failures remained unknown. The least significant failure category was intermittent failures, which seem to appear and disappear randomly. Only failure mode in this category was periodic processor hang-up, which was identified to be caused by inadequate environmental control. Appendix A presents all digital I&C failure modes identified by Korsah et al. (2010). However, the listing is based on only 100 analyzed events and is therefore not exhaustive.

The report by Ostenso & May (1996) defines the requirements for qualifying a commercially available PLC for safety-related operations in U.S. nuclear power plants. A manufacturer must perform a FMEA for its product in order to apply for the qualification. FMEA of a TRICON PLC is utilized in this study (Sinocruz 2007). Module and major component level is used also in these analyses. Module level is logical choice from operator's perspective as failed modules are replaced rather than repaired. FMEAs in the applications for the qualification are made for a certain product. However, this study considers generally all modular PLCs. Figure 19 present typical configuration of a modular PLC.

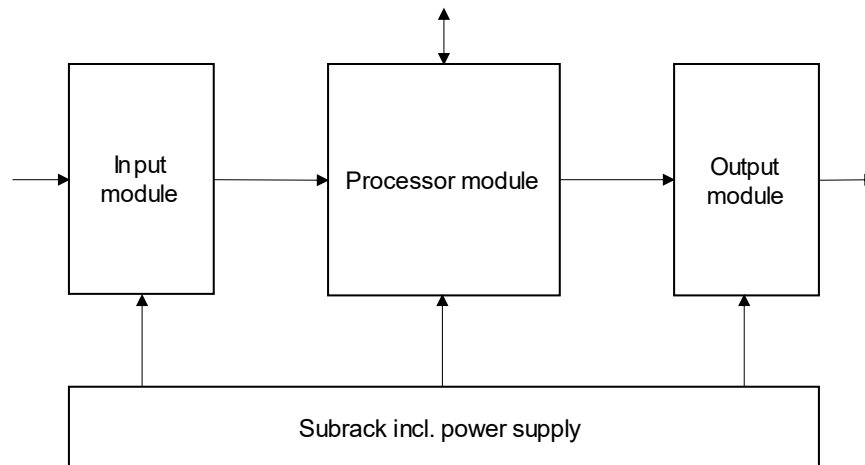


Figure 19. Block diagram of a typical PLC hardware. Modular PLCs can be expanded with additional modules. Arrows represent signal or power flows.

The following subchapters present FMEA for PLC and HMI in tabulated form, after software and manufacturer related causes are discussed. Internal electronic component failures are not distinguished in the tables as modern electronic devices are rather replaced than repaired, meaning that any component failure will lead to the same outcome. Nonetheless, special emphasis is given to age-related failure mechanisms as they can give insights related to replacement interval in critical applications. Age-related failure mechanisms of semiconductor devices are for example electromigration, time-dependent dielectric breakdown, negative bias temperature instability and hot carrier injection, which can all cause intermittent or permanent failures. A permanent failure at the basic component level can cause intermittent failures at the system level as a single component may not be needed continuously. Even though all of the mechanisms involve gradual accumulation of damage, electromigration and time-dependent dielectric breakdown appear random as the damage is not apparent until failure e.g. open circuit. Negative bias temperature instability and hot carrier injection cause failures when electrical characteristics gradually drift away from the desired values. (Vega et al. 2017)

PLC modules have often internal fuses, which are not user replaceable. Hence blown internal fuse leads to module replacement. This can be avoided by using external fuses or circuit breakers with lower current rating or faster operating time than the internal fuse. Required minimum intervention refers to the minimum action after which the system is restored to operational state, although the fault or error may still be present in the system. For example,

system can be restored to functional state after crashing by reboot, but the root cause will probably remain and cause more crashes in the future.

5.1.1 Software

If soft errors are neglected, software failures result from human errors made in the programming. An error becomes a failure when it causes incorrect output. Typical software failures are deterministic in the sense that failure occurs certainly when the defect part of the program is executed. However, these failures are stochastic in the sense that it cannot be known when the input that will cause the failure is fed to the system, if ever. These kind of failures have typically decreasing rate. (Jalote et al. 2008)

In addition to immediate failures, software errors can cause also gradual performance degradation and eventually failure by accumulated errors. Examples of such processes are numerical error accumulation, memory leakage, fragmentation, unterminated processes and data corruption. Accumulated errors usually generate a failure when resources, like memory or computing power, are exhausted. Time to failure depends on resource overhead, operational profile and resource loss per execution. Numerical error can cause failure e.g. when using variables stored as floating point data since resolution of floating point variable decreases as its value increases. These software aging mechanisms can be divided to volatile and non-volatile ones. Volatile effects can be removed by system reboot, meaning that occasional reboot, preferably during idle, may be advisable for continuous processes. Examples of non-volatile aging processes are file system and database metadata fragmentation. Table 4 shows classification of age-related software errors as presented by Grottke et al. (2008).

Table 4. Classes of age-related software errors. (Grottke et al. 2008)

Class	Extension	Examples
Resource Leakage	(1) OS-specific (2) App-specific	-Unreleased memory (1,2) File handlers (1) sockets (1) -Unterminated processes (1) Threads (1,2)
Fragmentation	(1) OS-specific (2) App-specific	-Physical memory (1) -File system (1) -Database files (2)
Numerical error accrual	(1) OS-specific (2) App-specific	-Round-off (1,2)
Data corruption accrual	(1) OS-specific (2) App-specific	-File system (1) -Database files (2)

According to Authén et al. (2015) software reliability modelling still has some unsolved issues. Software systems cannot be easily decomposed into components, and the interdependencies between the components are difficult to identify and model. Despite the unsolved issues, Authén et al. (2015) state that there seems to be philosophical consensus that software failures can be treated probabilistically, and it is meaningful to use software failure rates in reliability modelling. However, there has been also opposing opinions stating that software failure rates do not make sense (Singpurwalla 1995).

One approach to estimate reliability of software is errors per lines of code metric. Software industry average is 15-50 errors per 1000 lines of delivered code. (Sandu et al. 2018) Other approach is statistical analysis. For example, Chillarege et al. (1995) made an empirical study of an undisclosed yet widely distributed commercial software with several million lines of code. Obtained MTBF values were 2-4 years, and reliability increased with age.

The study by Korsah et al. (2010) tried to identify software related failures from operating experience data. Under 10 % of the records were attributed to software. It turned out to be difficult to distinguish software related failures as software is integral part of digital control systems and descriptions of the records did not have enough information. For example, "loss of communication" can be caused by software or hardware, and it is impossible to know the cause without further specifications. Table 5 presents identified causes of software-related failures. Most of the software failures lead to system crash and reboot (Siewiorek & Swarz 1998, p. 26).

Table 5. Identified causes of software-related failures. (Korsah et al. 2010)

Incomplete description of requirements
Incorrect firmware coding
Faulty calculation in program
Requirements error
Incorrect interpretation of requirements
Task/Application crash
Inadequate software version control
Software update incompatible with the Plant Process Computer design basis

Babeshko et al. (2008) proposed that FMEA of industrial control systems should be extended to F(I)MEA where "I" stands for intrusion. Recognized intrusion modes are sniffing, system remote control, Open Platform Communications buffer overflow and

denial-of-service attacks. Also prevention of software related failures should be seen as a regular maintenance activity. However, developing software of an existing system can be difficult or even impossible as e.g. operating system of a PLC can be burned to ROM. Examples of fault prevention actions include: checking of input parameters, minimization of corporate and industrial network interconnections, implementation of security checking on all levels of industrial control systems, data encryption, and removing all unused ports and services. Checking of input parameters refers to ensuring that incorrect inputs generate alarms rather than incorrect outputs.

5.1.2 Manufacturer

One reason why lifetime of a PLC can come to an end is if the manufacturer discontinues supporting the device. If a PLC is used after spare part availability has ended, failure can lead to a prolonged unplanned outage. Manufacturers have developed lifecycle models to provide information of their products lifecycle statuses. Figure 20 presents ABB's four-phase model.

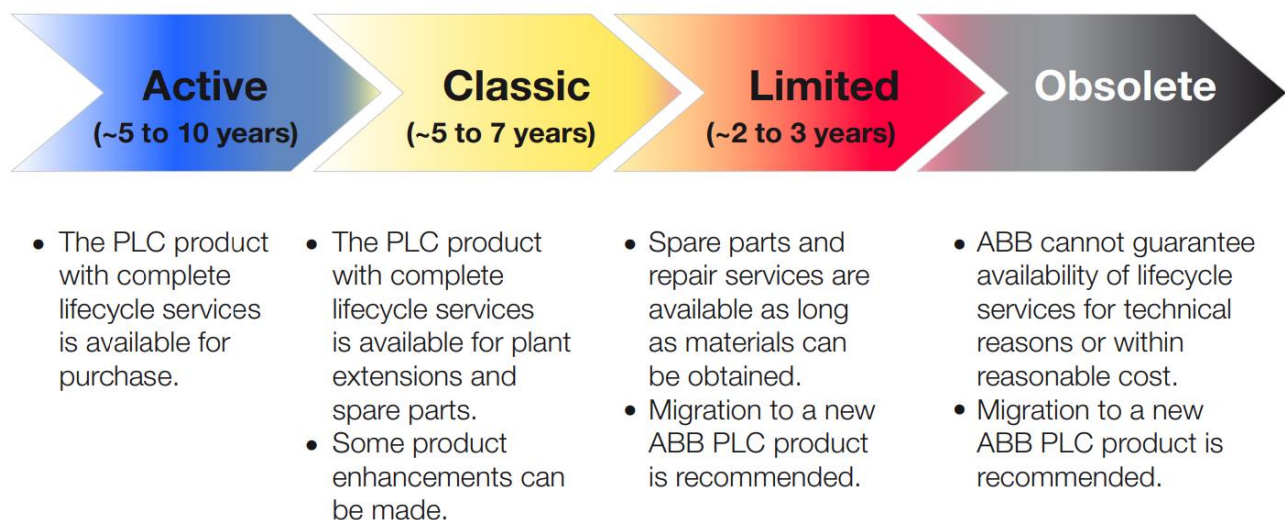


Figure 20. PLC lifecycle model. (ABB 2018)

In active and classic phases full support is guaranteed, and after that spare parts and services may be available but it cannot be guaranteed. If a situation where spare parts and support are not available is considered unacceptable, the PLC should be renewed every 10-17 years depending on the phase-out announcements. However, ABB states that the PLCs are typically supported for more than 20 years (ABB 2018). Figure 21 shows Siemens HMI lifecycle model.

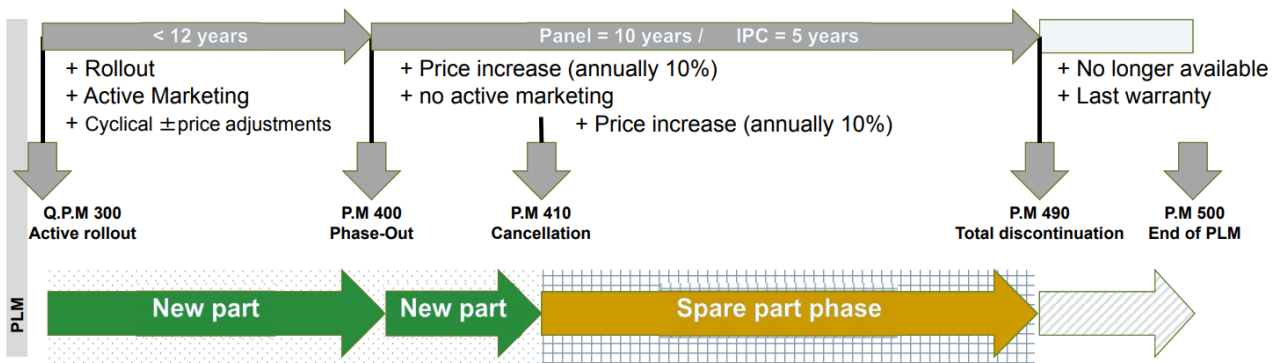


Figure 21. Siemens HMI lifecycle model. (Siemens 2014)

Another renowned PLC and HMI manufacturer, Siemens, states that from the start of the product delivery, HMIs are available to be ordered for up to 12 years. After the phase-out is announced, spare parts will be available for 10 years for PLCs and panels, and 5 years for industrial PCs. However, price of the spare parts will increase 10 % annually after the phaseout announcement. (Siemens 2014)

5.1.3 Power supply module and chassis

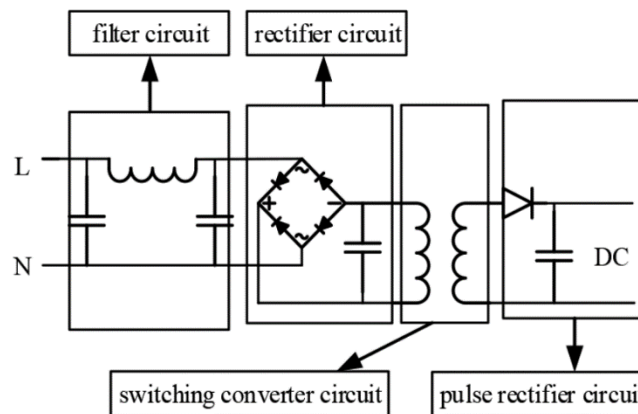


Figure 22. Schematic of a switching power supply. (Xu et al. 2020)

Power supply module consist of a filter circuit, a rectifier circuit, a switching converter circuit and a pulse rectifier circuit as shown in Figure 22. Electrolytic capacitors have been identified to be the life limiting factor of power supplies as the electrolyte gradually evaporates and diffuses out, which leads to decreased capacitance. Capacitors are used in power supplies for voltage smoothing and as energy storages. Degraded capacitors can impair low voltage ride through capability of the system, meaning that the PLC will crash from voltage dips it would have survived with intact power supply module. (Peck 2012) A manufacturer Omron estimates lifetime of at least 8-10 years at ambient temperature of 40

°C for its power supplies. The estimate is obtained by accelerated life testing of electrolytic capacitors. Lifetimes are extrapolated to normal operating conditions assuming that life expectancy doubles for every 10 °C decrease in temperature. (Omron 2021) Table 6 presents FMEA for power supply module and chassis.

Table 6. Power supply module and chassis FMEA. (Sinocruz 2007 (1), Korsah et al. 2010 (2), Authén et al. 2015 (3), Peck 2012 (4))

Failure mode	Failure mechanism	Failure effect (2)	Failure category (3)	Effect on PLC operability	Required minimum intervention
Loss of input power (1)	Facility blackout (1)	Total	Random sudden	Input signals will not be read. Analog and digital outputs fail low. (1)	Reboot
Loss of power supply module (1)	Electronic component failure, electrical power transient, insulation failure, fire, condensed moisture, mechanical shock (1)	Total	Several	Input signals will not be read. Analog and digital outputs fail low. (1)	Module replacement
Open circuit (1)	Blown external fuse, loose contact, cable cut (1)	Total	Several	Input signals will not be read. Analog and digital outputs fail low. (1)	Repair
Impaired low voltage ride through capability (4)	Degraded capacitors (4)	Partial, plausible	Age-related	PLC crashes from voltage dips it would survive with intact power supply. (4)	Reboot
Power supply output fails low (1)	Electronic component failure (1)	Partial, plausible	Several	Affected I/O fails low, affected field loads from relay outputs will fail to the de-energized state. (1)	Module replacement
Power supply output fails high (1)	Electronic component failure (1)	Partial, plausible	Several	Fed modules may be damaged, relay contacts may flash over, I/O may fail low if voltage is high enough to burn out affected I/O points. (1)	Module replacement
Battery output voltage fails low (1)	Battery aging or short circuit (1)	Partial, plausible	Several	Battery failure concurrent with power supply failure can result in loss of memory. (1)	Battery replacement

5.1.4 Input and output modules

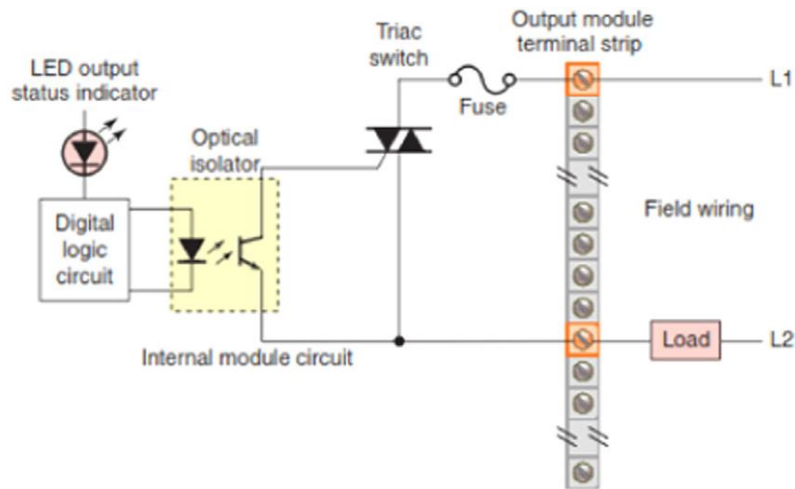


Figure 23. Simplified circuit diagram of a triac switch output. (Alphonsus & Abdullah 2016)

Figure 23 shows simplified circuit of a triac output. Optoisolators and relays are recognized to suffer from wear-out. Optoisolator consist of a light emitting diode (LED) and a phototransistor. Slama et al. (2007) studied aging of optoisolators by accelerated life testing. Degradation of the LED was identified to be the main life limiting factor of optoisolators. Operating current has high effect on the LED lifetime as shown in Figure 24. Relay outputs will eventually fail due to contact wear since electric arc occurs during every operation. Panasonic estimates lifetime of more than 100 000 operations for its relay outputs (Panasonic 2014). Based on Rizwan (2006) it is assumed that failure of a single I/O port can be repaired but failure of multiple ports leads to module replacement. Table 7 presents FMEA for input and output modules.

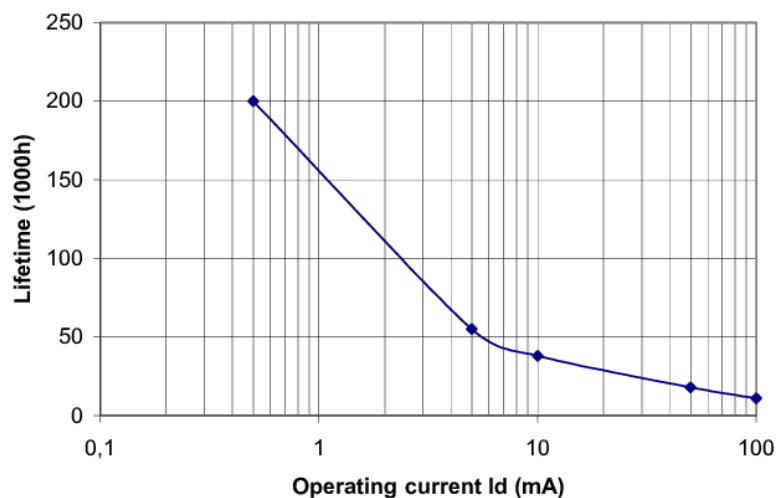


Figure 24. Optoisolator lifetime as a function of operating current. (Slama et al. 2007)

Table 7. Input and output module FMEA. (Sinocruz 2007 (1), Korsah et al. 2010 (2), Authén et al. 2015 (3))

Failure mode	Failure mechanism	Failure effect (2)	Failure category (3)	Effect on PLC operability	Required minimum intervention
Loss of I/O module(s) (1)	Electronic component failure, electrical power transient, fire, condensed moisture, mechanical shock (1)	Total	Several	Input signals will not be read. Analog and digital outputs fail low. (1)	Module replacement
Input(s) processing failure (1)	Electronic component or software failure (1)	Total	Several	Affected input(s) will not be read. Processor diagnostics typically detects the fault. (1)	Module replacement
Output(s) processing failure (1)	Electronic component or software failure (1)	Total	Several	PLC unable to control the affected output point(s). Processor diagnostics typically detects the fault. (1)	Module replacement
Digital input stuck to a value (1)	Electronic component failure (1)	Partial, plausible	Several	PLC unable to correctly determine the state of the affected point. Failure can remain undetected and lead to incorrect outputs. (1)	Repair
Analog input fails high or low (1)	Electronic component failure (1)	Partial, plausible	Several	PLC unable to correctly determine the state of the affected point. Diagnostics will detect the fault if values are out of range. (1)	Repair
Analog or digital output fails high or low (1)	Electronic component failure (1)	Partial, plausible	Several	PLC unable to control the affected output point. Depends on diagnostics and process output, whether the fault is detected or not. (1)	Repair
Relay output fails open or closed (1)	Electronic component failure (1)	Partial, plausible	Several	PLC unable to control the affected output point. Depends on diagnostics and process output, whether the fault is detected or not. (1)	Repair

5.1.5 Processor module and human machine interface

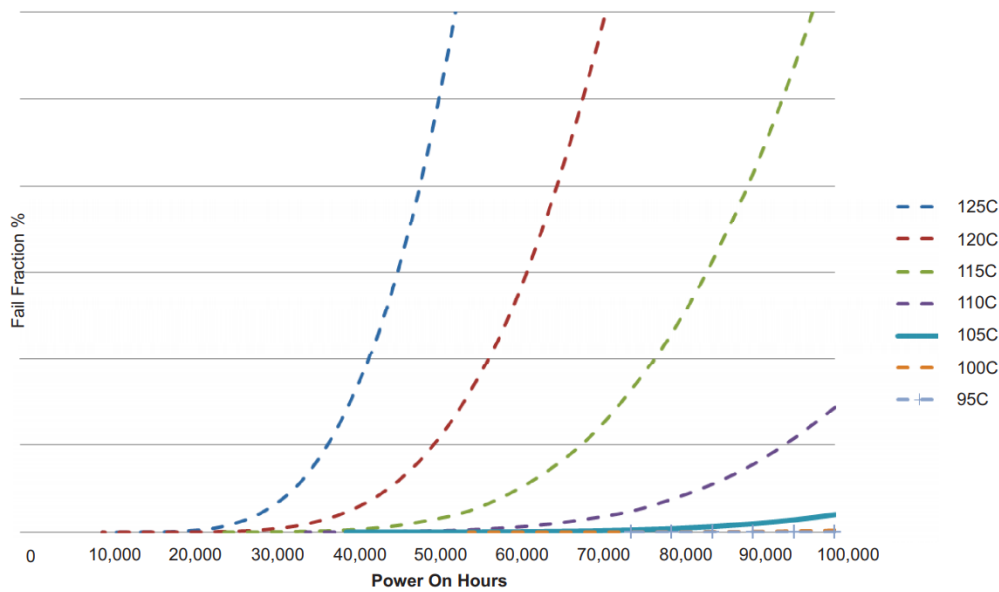


Figure 25. Embedded processors' failure fraction from electromigration as a function of power on hours for different junction temperatures. (Webber 2020)

Electromigration is said to be the critical wear-out mechanism for metal-oxide-semiconductor integrated circuits (Azidehak 2017) (Webber 2020). DC circuits with high current densities can suffer from electromigration. It occurs when conducting electrons transfer momentum to the metal atoms of the conductor. This causes net movement to the direction of electron flow, which leads to increased resistance and eventually to an open circuit. Locations most prone to electromigration are vias, which are conductive paths through the layers of a circuit board. (Shin 2008) Electromigration accelerates significantly when junction temperature increases as shown in Figure 25. Texas Instruments designs its embedded processors to have an useful life of 100 000 power-on hours at junction temperature of 105 °C, after which the failure rate increases (Eller 2015). Vega et al. (2017) stated that failure times due to electromigration follow lognormal distribution with $\sigma \approx 0.2$, which corresponds to eventually steeply increasing cdfs as depicted by Webber (2020) for high junction temperatures. MTTFs of semiconductor circuits due to electromigration can be estimated using empirically derived Black's equation. (Shin 2008) Table 8 presents FMEA for processor module and human machine interface. Ancillary damage resulting from a faulty control signal is assigned to processor module. From PLC's point of view it is an intermittent failure mode as it can happen multiple times if the root cause is not removed.

Table 8. Processor module and human machine interface FMEA. (Sinocruz 2007 (1), Korsah et al. 2010 (2), Authén et al. 2015 (3), Grottke et al. 2008 (4), Vega et al. 2017 (5), Schneider Electric 2016 (6))

Failure mode	Failure mechanism	Failure effect (2)	Failure category (3)	Effect on PLC operability	Required minimum intervention
Loss of processor module (1)	Electronic component failure, electrical power transient, fire, condensed moisture, mechanical shock (1)	Total	Several	Input signals will not be read. Analog and digital outputs fail low. (1)	Module replacement
Periodic processor hang or crash (4)	Software (programming error) (4)	Total	Intermittent	Input signals will not be read. Analog and digital outputs fail low. (1)	Reboot
Periodic processor hang or crash (4)	External (excess temperature, intermittent power supply failures, electromagnetic interference, overflow) (4)	Total	Intermittent	Input signals will not be read. Analog and digital outputs fail low. (1)	Reboot
Periodic processor hang or crash (4)	Volatile software aging (resource leakage, fragmentation, numerical error accrual) (4)	Total	Intermittent	Input signals will not be read. Analog and digital outputs fail low. (1)	Reboot
Periodic processor hang or crash (5)	Electronics (e.g. electromigration) (5)	Total	Intermittent	Input signals will not be read. Analog and digital outputs fail low. (1)	Reboot
Memory lost or corrupted (4)	Electromagnetic interference, power disruption while writing or refreshing, memory corruption attack, non-volatile software aging (4)	Total	Random sudden, age-related	Unable to execute the program.	Reload program from backup memory
Loss of HMI	Electronic component failure, electrical power transient, insulation	Total	Several	Unable to process inputs and outputs.	Renew

	failure, fire, condensed moisture, mechanical shock (1)				
Ancillary damage	Wrong control signal (design or programming error)	Partial, plausible	Intermittent	No effect on PLC.	System redesign
Software incompatibility with new requirements or surrounding devices	Obsolescence	Partial, plausible	Age-related	Unable to satisfy updated functions.	Reprogram
Hardware incompatibility with new requirements or surrounding devices	Obsolescence	Partial, plausible	Age-related	Unable to satisfy updated functions.	Renew
Failure to transmit or receive data with external devices (1)	Electronic component or software failure (1)	Partial, plausible	Several	PLC continues to operate. Communications to external network devices are interrupted. Main processor diagnostics will detect and flag communications fault if application software is so designed. (1)	Module replacement
Loss of HMI display (6)	Backlight degradation (6)	Partial, non- plausible	Age-related	Outputs cannot be read, inputs can still be processed.	Renew
Loss of HMI input device	Electronic component failure, condensed moisture, mechanical shock (1)	Partial, non- plausible	Several	Outputs can still be read, inputs cannot be processed.	Replace input device if external, replace HMI if integrated

5.2 Literature and manufacturer failure data

Module level approach enables relatively easy reliability estimates if a constant failure rate can be assumed. Failure rates for the modules can be obtained from the manufacturer or generic failure rates from the literature can be used. However, generic sources rarely distinguish different failure modes or failure effects. Parts count method can be used for non-redundant systems i.e. failures rates of the modules are summed to comprise the total failure rate. Table 9 presents PLC failure rates found from the literature. Failure shares are calculated assuming that system has only one input and one output module. Annual PoF is calculated by exponential survival function based on the constant failure rate assumption.

Table 9. PLC failure rates per hour from the literature.

Source	Power supply	CPU	Digital input	Analog input	Digital output	Analog Output	Total	Annual PoF
Exida (2003)	5,00E-06	1,00E-05	1,00E-06	2,00E-06	1,00E-06	2,00E-06	2,10E-05	0,17
Choi et al. (2012) T = 30°C	3,81E-06	4,94E-06	4,92E-06	4,00E-06	4,05E-06	3,61E-06	2,53E-05	0,20
Choi et al. (2012) T = 50°C	8,11E-06	1,35E-05	1,05E-05	1,27E-05	8,11E-06	9,84E-06	6,28E-05	0,42
Held & Fritz (2009) FIDES							5,47E-06	0,05
Held & Fritz (2009) RIAC 217Plus							7,83E-06	0,07
Held & Fritz (2009) avionics field data							7,20E-06	0,06
Parashar & Taneja (2007)							5,50E-06	0,38
Rizwan (2006)							7,80E-05	0,50
Average	4,31E-06	7,23E-06	5,48E-06	6,23E-06	4,39E-06	5,15E-06	3,28E-05	0,25
Share of failures	22 %	37 %	23 %		18 %			

Failure rates have even order of magnitude differences but are for different devices operating in different environments. It is unclear whether crashes and reboots are counted as failures. Values by Choi et al. (2012) are calculated by MIL-HDBK-217F methodology, which is considered conservative. Held & Fritz (2009) compared two failure rate estimation models to actual field data. Parashar & Taneja (2007) and Rizwan (2006) obtained their estimates from industrial failure data. Table 10 presents different ways to categorize PLC failures found from the literature.

Table 10. Different ways to classify failures and their respective shares.

Source	Classification criterion	Attributes and respective shares					
Rizwan (2006)	Required minimum intervention	Repairable	Component replacement	Unit replacement	Software reinstallation		
	Share of failures	30 %	41 %	7 %	22 %		
Parashar & Taneja (2007)	Failure type	Minor failure	Major failure, repairable	Major failure, irreparable			
	Share of failures	19 %	71 %	10 %			
Siewiorek & Swarz (1998) Sun-2	Source of error	Permanent fault	Intermittent fault	Transient fault			
	Share of failures	3 %	56 %	41 %			
Siewiorek & Swarz (1998) Tandem	Source of outage	Power	Communication lines	Application software	Max files	Hardware	Corrupted files
	Share of failures	53 %	22 %	10 %	8 %	5 %	2 %

Rizwan (2006) further describes that repairable failure is e.g. input module failure. Component replacement can be caused by e.g. burnt relay or power supply failure. Unit replacement refers to a situation where the unit is burnt completely. Software needs to be reinstalled when it is corrupted and has stopped working. Parashar & Taneja (2007) state that PLC failures can be divided to three categories: power supply failure/software corruption (minor), component failure (major, repairable) or the unit can be burnt (major, irreparable). Shares of irreparable failures stated by Rizwan (2006) and Parashar & Taneja (2007) seem to be high and were not backed up by an expert opinion (Kurtti 2021). Intermittent failures can be caused by software or unstable hardware. Transient fault is caused by temporary environmental condition. In sources used by Siewiorek & Swarz (1998) permanent failures comprise 2-12 % of all system failures. Data used by Siewiorek & Swarz (1998) is quite old and considers generally computer systems instead of PLCs. However, it can still provide useful insights. Power quality on plant is probably an important contributor to failure rates.

According to HMI manufacturer Schneider Electric (2016), liquid crystal display backlight is the main life limiting factor of HMIs, and it is expected to last 50 000 hours. The backlight degrades gradually, and it is considered failed when the brightness has decreased to half. If there is a dominant age-related failure mode, constant failure rate assumption may not be appropriate. Also Siemens (2021) and EKE electronics (2017) give an MTTF of 50 000 hours to their HMI products.

Similar approach can be used as well if the basic component level is chosen. However, number of the components is substantially larger compared to the number of modules. For example, Held & Fritz (2009) conducted a study where they compared two failure rate prediction models for electronic components: FIDES and RIAC 217Plus. An avionics control unit, which consists of a power supply, an I/O board, a CPU board and a connect board was used as an example. The control unit consists of approximately 7 000 electronic components for which the failure rate has to be determined. In practice failure rates are estimated for component families rather than individual components. Table 11 presents predicted failure rates for electrical component families included in the avionics control unit.

Table 11. Predicted failure rates for electronic component families of a control unit in failures per 10^9 hours (FIT) and their contributions to the total failure rate. (Held & Fritz 2009)

Component family	number	FIDES	%	FIT/comp.	RIAC 217Plus	%	FIT/comp.
Capacitor	2'356	572	10.5	0.2	1802	23.0	0.8
Diode	633	2659	48.7	4.2	1188	15.2	1.9
IC	470	1315	24	2.8	1450	18.5	3.1
Inductor	348	50	1	0.1	2	<0.1	<0.001
Optocoupler	10	8	0.1	0.8	158	2.0	15.8
Relay	46	84	1.5	1.8	803	10.3	0.8
Connector	57	109	2	1.9	87	1.1	1.5
Resistor	2'940	433	8	0.1	1231	15.7	0.4
Transistor	84	120	2.1	1.4	978	12.5	11.6
Miscellaneous	16	115	2.1	7.2	132	1.7	8.3
Total		5465	100		7830	100	

Both of the models give high contributions to capacitors, diodes and integrated circuits (IC). The FIDES model estimates that these three account for about 80 % of the total failure rate, and according to the RIAC model these three contribute 60 %. However, the models show large differences for relays, resistors, transistors and optocouplers. Field data from more than 15 years of operation gives FID of 7 200 for the controller, meaning that both of the predictions are in reasonable range.

The FIDES prediction was recognized to be optimistic. The methodology includes an audit regarding design, production, system integration, field operation and maintenance. Flawless process was assumed as no audit was performed. (Held & Fritz 2009) Operational and environmental stress factors are included in both models. Stress parameters included in the FIDES model are: thermal, thermal cycling (case and solder joints), thermo-electrical, mechanical, chemical, electrical and humidity. The RIAC model includes operational stresses, thermal cycling (separate for case and solder joints) and electrical overstress. The

FIDES model estimates that thermal and thermal cycling are the most significant stress factors, while according to the RIAC thermal cycling of solder joints and electrical overstress are dominant. (Held & Fritz 2009) These factors impair comparability of the results as airplane and power plant have different conditions. Figure 26 illustrates how changes in ambient temperature, annual cycles, vibration and relative humidity affect the failure rate.

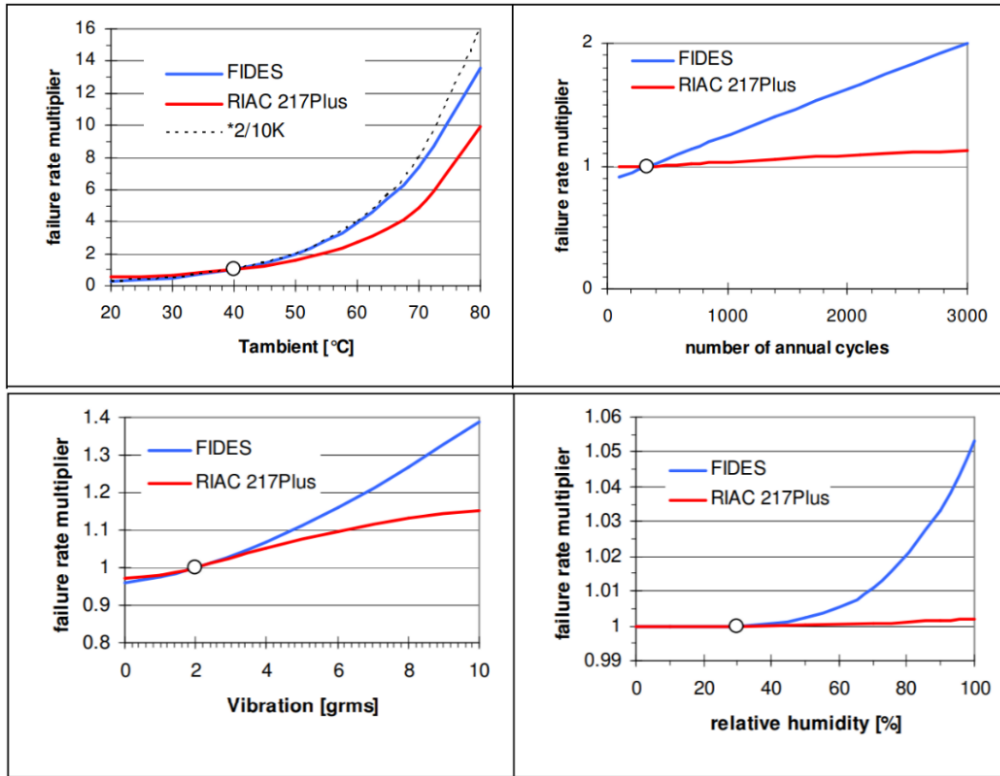


Figure 26. Effect of different stress factors on the failure rate of electronics. (Held & Fritz 2009)

The models predict fairly similar effects on moderate values, except for the number of annual cycles. According to the both models, ambient temperature has the most significant effect on the failure rate. The FIDES model estimates that the failure rate almost doubles per 10 °C increase in temperature. Impact of humidity is considered only for non-operational phases. The RIAC model assumes that humidity impacts only non-hermetic integrated circuits, while the FIDES model assumes that it affects most of the components. Estimated total failure rates are in reasonable range, but the composition differs quite significantly between the models. However, the real composition of total failure rate cannot be validated since the field data does not include root causes of failures. (Held & Fritz 2009)

5.3 Data acquisition

Data for the modeling was acquired from collaborative partner's maintenance records. Initially, the records were filtered by asset location and work type specific codes to extract interesting work orders with minimal manual work. Interesting work types are immediate maintenance, deferred maintenance and condition-based maintenance, which may refer to a failure. Location was limited to BoP. As all of the work orders were not assigned to correct location, the data acquisition was supplemented with keyword searches such as PLC, HMI, computer and automation. After filtering the work orders were interpreted manually.

Earliest work orders were from 2006. It was identified that yearly amount of work orders has been increasing over the years. The increasing trend is probably more related to eagerness to log the work orders than asset reliability. If all events were not logged in the earlier years, it can lead to a biased estimate. Earliest years were omitted when it was evident that events were not logged for the plant in question. Figure 27 gives an illustration of the relation between PLC & HMI age and annual number of failures.

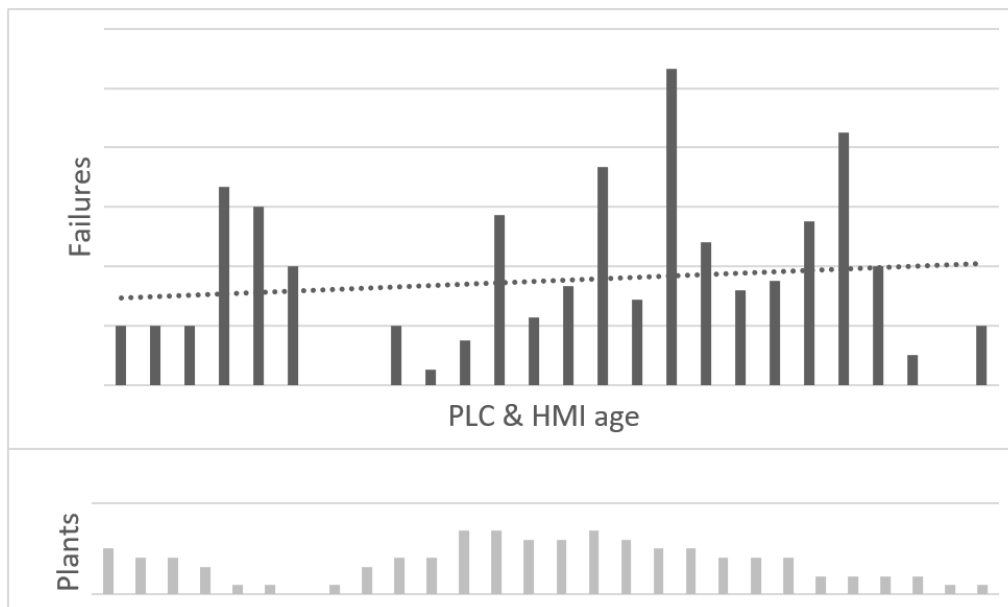


Figure 27. Average yearly number of failures as a function of PLC & HMI age.

The annual event count shows slightly increasing trend and no signs of infant mortality. As the data is from same years, devices of different ages are of different generations that may have inherently different reliabilities. Most common failure descriptions were “computer crash” and “communication error”. Majority of the failures were presumably resolved by

rebooting. It was not possible to distinguish intermittent failures from permanent failures, which complicates the modeling.

5.4 Survival analysis

Survival analysis is a statistical method used to analyze the expected time until an event occurs. Firstly, a Kaplan-Meier estimator was constructed using the acquired failure data. Kaplan-Meier estimator is built from a mission list, which contains times to a failure or censoring. Mission can start from commissioning time or time when the asset is restored to operation after a failure. Hence, the mission time is equal to the asset age for irreparable assets, but for repairable assets the age and mission time differ after the first failure. Kaplan-Meier estimator is best suited for irreparable assets or permanent failures which are repaired after their occurrence. However, in this case most of the failures were recurrent intermittent failures. As a result, most of the missions start from a reboot, and the root cause is not resolved. Figure 28 shows the obtained Kaplan-Meier estimator and distributions fitted to it.

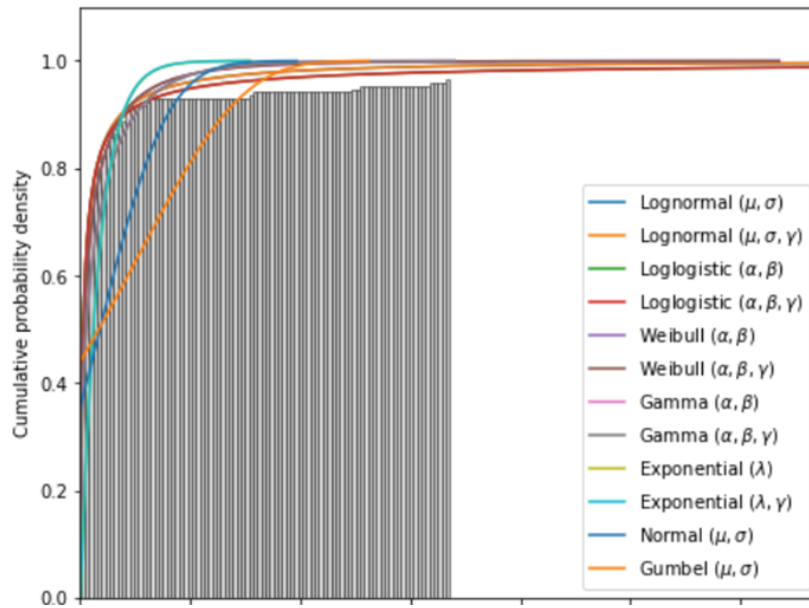


Figure 28. Kaplan-Meier cdf and fitted parametric models.

12 different distributions were numerically fitted to the Kaplan-Meier estimator using maximum likelihood estimation. The best fitting distribution was 2-parameter Lognormal distribution. Hazard function shown in Figure 29 is obtained from the best fitting parametric model, and it is a very steeply decreasing infant mortality curve. The interpretation is that probability of failure decreases as the unit survives longer. However, the right tail is speculative as none of the units survived to the flat section. Only data points in that section

were caused by absence of data from commissioning of old units to the beginning of the dataset. Cumulative hazard gives expected number of failures until time in question.

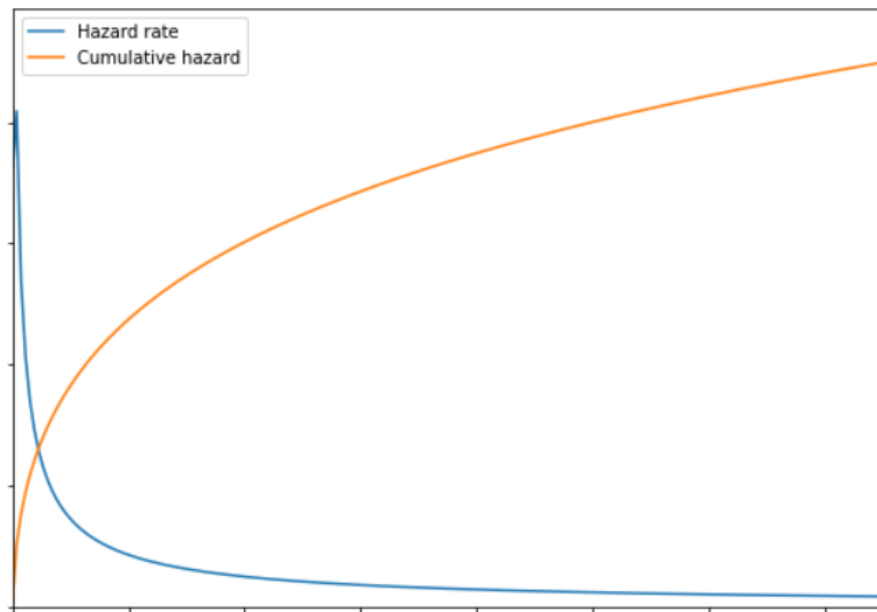


Figure 29. Hazard and cumulative hazard functions.

5.5 Survival regression

Conventional survival analysis estimates reliability only as a function time, but the analysis can be extended with survival regression to consider also effect of other variables. Log-linear Cox proportional hazards model was used to study the reliability of products from different manufacturers. Used explanatory variables are binary variables that indicate whether a unit is produced by a particular manufacturer. Figure 30 shows obtained regression parameter values.

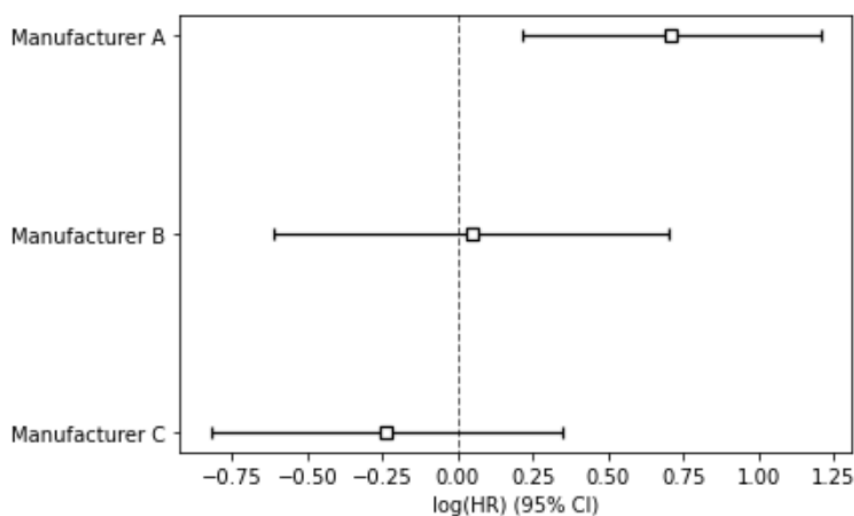


Figure 30. Comparison of the reliability of devices from different manufacturers using the Cox model.

Horizontal axis values are natural logarithms of the baseline hazard multipliers. For example, the parameter value 0.7 of the manufacturer A corresponds to a baseline hazard multiplier of $e^{0.7} \approx 2$. Number of units per manufacturer was small, and the confidence intervals are consequently wide. Therefore, this should be seen as a preliminary result and the sample size should be increased to obtain more reliable results.

5.6 Event tree models

Event tree analysis is a modelling technique used for quantitative risk assessment. The tree consists of an initiating event, chance node events and outcomes placed in logical order. Optimally, the model should be FMEA based, and every relevant failure mode should be modelled separately to consider different failure patterns of different failure modes. For example, failures caused by external events occur independently of the device's age and have thus constant rate.

Contrarily, failures caused by internal reasons can have age dependent behaviors, which should be studied from the failure data. It should be also easier to find justification for a parametric model when failure modes are treated separately. In addition to the initiating event frequency, also branch probabilities and consequences can have time dependencies. For example, low voltage ride through capability can weaken as the power supply module ages, and a same failure can have more severe consequences if spare part availability has ended. Appendix B outlines what the ideal models could be. However, realistically achievable level of detail in the models depends on failure data's level of detail. Hence, failure modes must be aggregated such that model parameters can be defined based on the data. Determining the optimal level of detail for the FMEA based models is beyond the scope of this work. In this subchapter different modelling approaches are tested and compared with a generic event tree. Event tree with a similar structure can be solved in different ways, and initiating event frequency can be estimated differently from the same dataset.

5.6.1 Conventional event tree model

Typically, an event tree is solved analytically by calculating the expected value. Recurrent failures are then modelled by multiplying expected consequence of a single failure by the expected number of failures. With this approach the yearly number of failures is deterministically predefined. Figure 31 presents schematic of the conventional model.

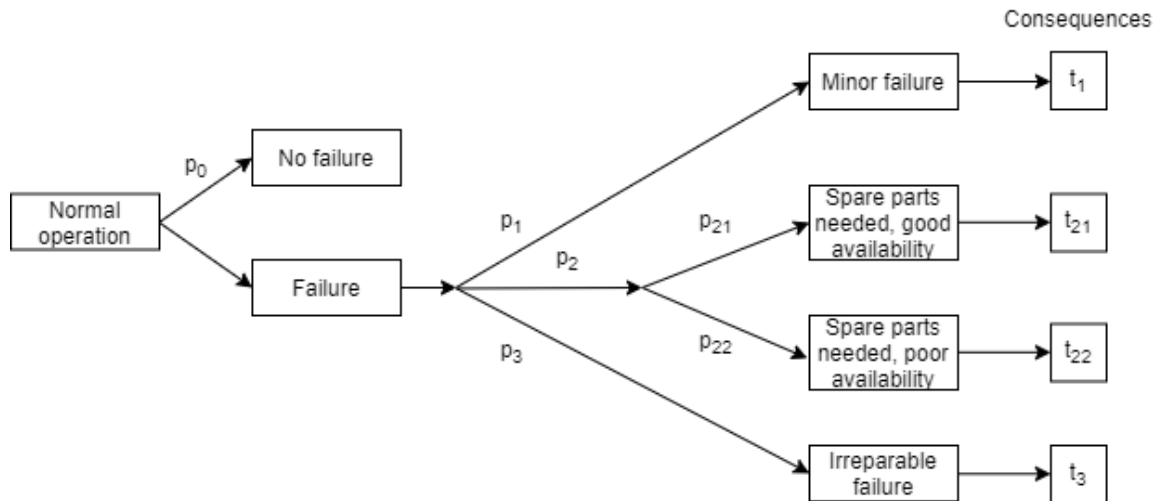


Figure 31. Schematic of the conventional model.

5.6.2 Stochastic event tree model

Frequency of an outcome can be calculated by multiplying initiating event frequency with the conditional probabilities along the path to that outcome only if they are independent. Outages resulting from failures prevent the next failure temporarily i.e. affect the initiating event frequency. If the independency condition is not met, the initiating event can be modelled as a renewal process and the tree is then Monte Carlo simulated. Figure 32 presents schematic of the stochastic model.

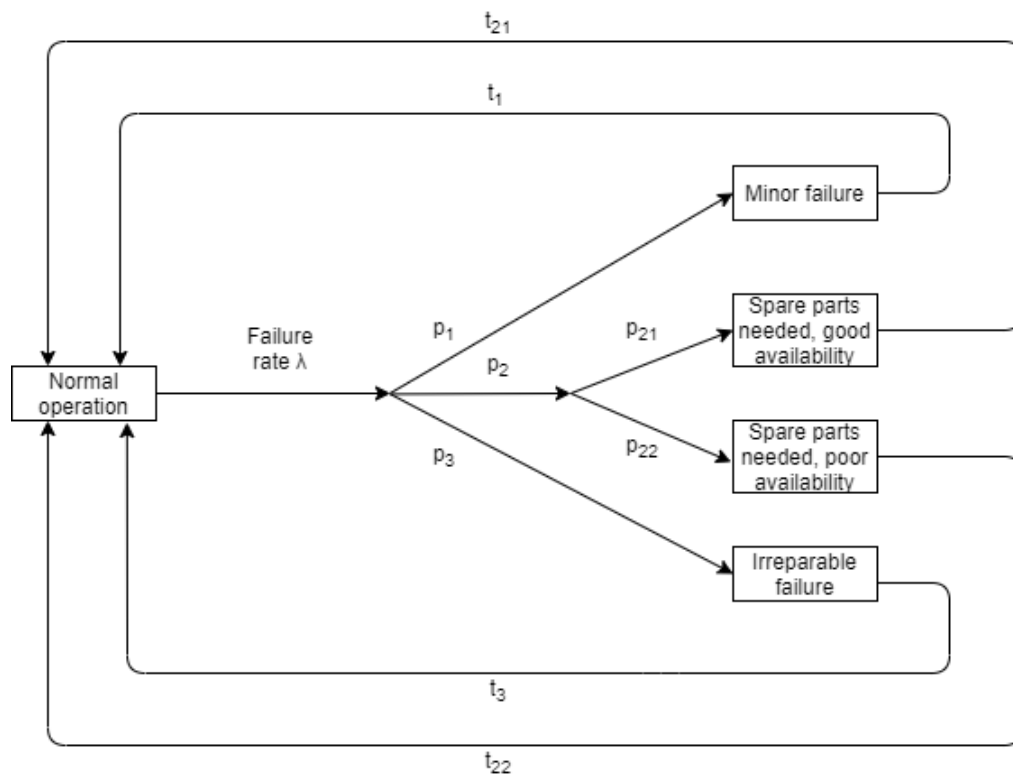


Figure 32. Schematic of the stochastic model.

In this model, initiating event, i.e. failure, is modelled as a Poisson process. When the alarm occurs, random number is drawn from a distribution corresponding to branch probabilities p , and some of the outcomes is reached. Every outcome, i.e. failure, has repair time t . After the repair time system returns to normal operation and next failure time is drawn. This cycle is repeated until cumulative time is equal to desired calculation period e.g. one year. The calculation period is simulated multiple times (e.g. 100 000) to obtain reliability and availability estimates with sufficient accuracy. In the stochastic approach there is no branch for a failure free year, but a failure free year is still possible as the first drawn time to a failure can be more than a year.

5.6.3 Initiating event frequency estimation

Initiating event frequency, or expected number of failures per year, can be estimated with two different approaches while using the same data. Either the trendline fitted to the annual event count (Figure 27) can be directly used as the initiating event frequency, or then it can be calculated from the cumulative hazard function shown in Figure 29. Figure 33 illustrates failure rate patterns obtained with different approaches.

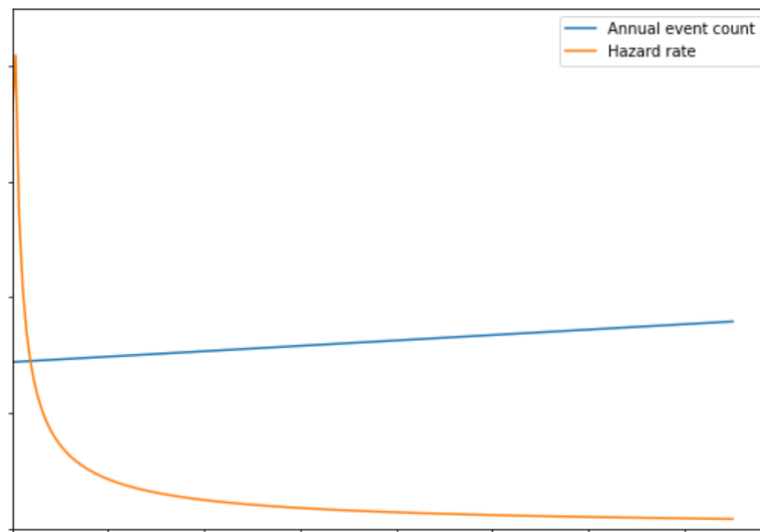


Figure 33. Annual event count vs hazard rate. Annual event count is a function of age, and hazard rate is a function of time from the last event.

The annual event count is a function of age, and hazard rate is a function of mission time. Mission time is different from the age as it goes back to zero when the asset is restored after a failure. Average failure rate for a time interval can be calculated from the hazard function using the following equation:

$$AFR(T_1, T_2) = \frac{\int_{T_1}^{T_2} h(t)dt}{T_2 - T_1} = \frac{H(T_2) - H(T_1)}{T_2 - T_1}, \quad (15)$$

where T_1 is starting time of the interval, T_2 is the ending time and $H(T)$ is the cumulative hazard function. Decision between these two approaches is important as the patterns are completely different. In this case both methods give quite similar failure rate for the first year, but after that the estimates diverge. In both of the approaches failure rate is constant during the calculation period. Year is used in this work, but it could be as well be a day to increase resolution of the discretization. The stationary models can be made stepwise time-dependent since the failure rate can change between the calculation periods.

5.6.4 Model comparison

With two different event tree modeling approaches, and two different initiating event frequency estimation approaches, there are four different modeling approaches. Initiating event frequencies are obtained from the data analysis, first branch probabilities are taken from Parashar & Taneja (2007) and outage durations are assumptions. In this demonstration also spare part availability probabilities are assumed. In real use cases spare part availabilities should vary as a function of product's lifecycle phase. Table 12 presents used parameter values. 100 000 iterations are used for the stochastic model.

Table 12. Model parameters used in the comparison.

Branch probabilities	p_1	p_{21}	p_{22}	p_3
	19 %	35,5 %	35,5 %	10 %
Outage durations	t_1	t_{21}	t_{22}	t_3
	6 h	3 d	2 weeks	20 weeks

Approaches are compared by calculating expected yearly availabilities and total numbers of failures. Figure 34 presents the estimated yearly numbers of failures, and Figure 35 presents the expected yearly availabilities. Only the first four years are compared since the extrapolated right tail of the hazard function is not considered representative.

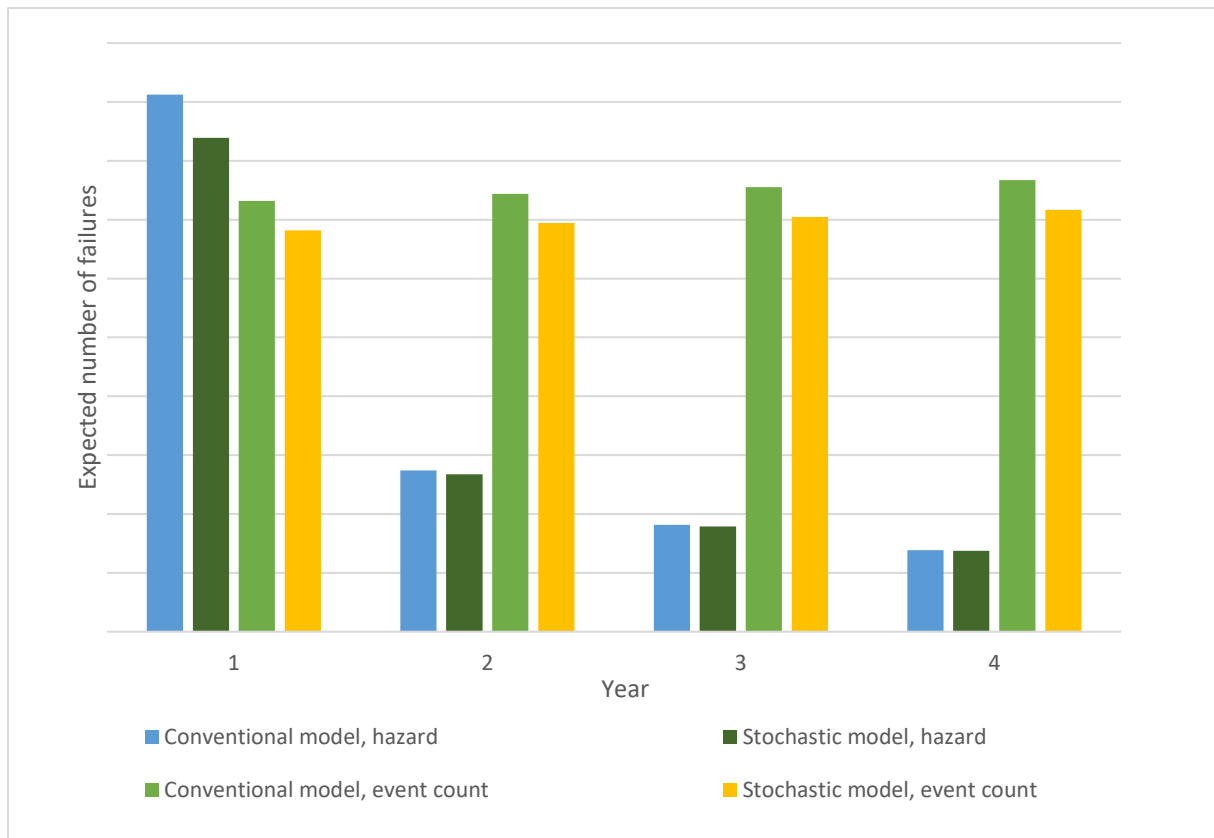


Figure 34. Comparison of yearly failures estimated by different modeling approaches.

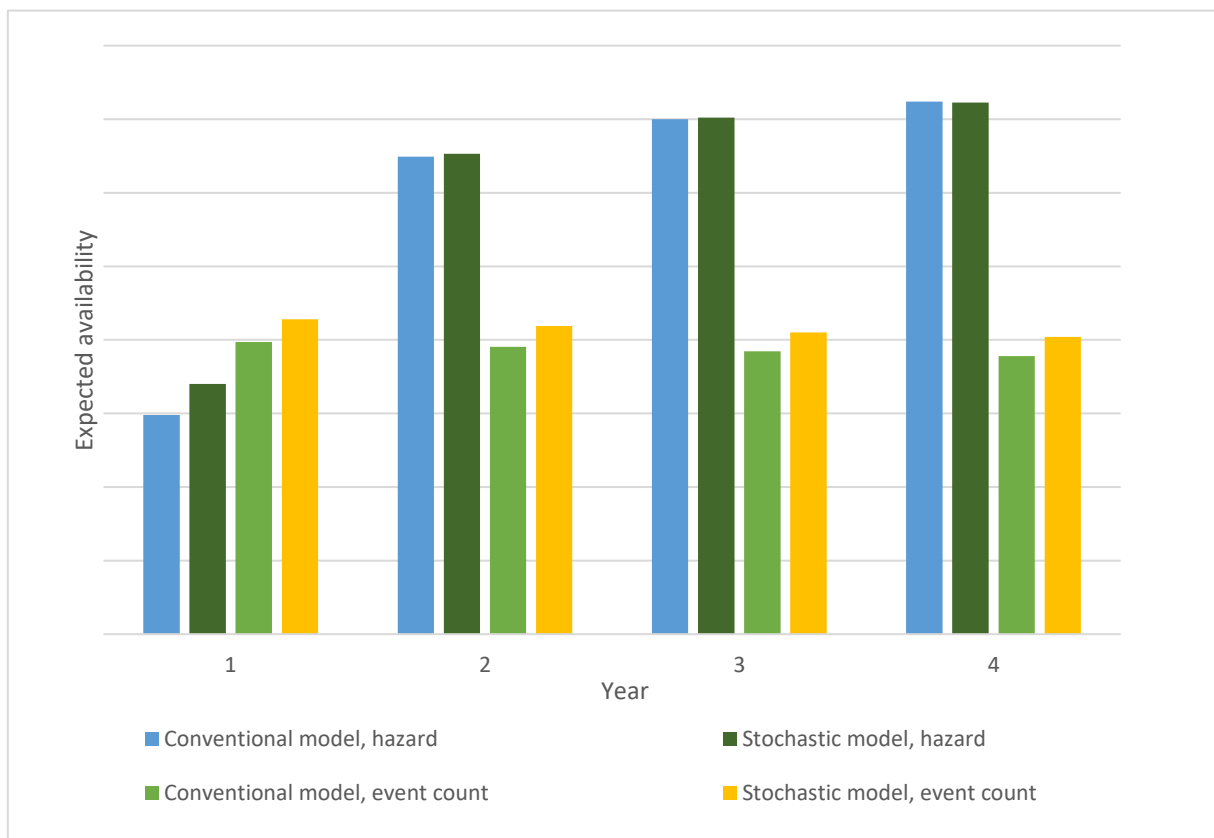


Figure 35. Comparison of availabilities estimated by different modeling approaches.

Difference in results between the initiating event frequency estimation approaches increases with time, and difference between the event tree models increases when failure rate or average outage length increases. Choice between the initiating event frequency estimation approaches depends on what is known. If time from the last failure is known, hazard function should be used. As many of the failures are consecutive, time from the last failure has significant effect on the predicted reliability. If only age is known, estimate has to be based on that. Time from the last failure is stronger condition as the reliability is almost constant as a function of age.

Benefits of simulation should offset the increased complexity and computation time to justify the stochastic model. However, results are almost the same with used failure rates. Difference increases as the system reliability weakens, and with extreme values the conventional model gives even negative availabilities. The question of what constitutes a significant difference depends on other uncertainties of the model. Nonetheless, expected number of failures for the calculation period can be decreased by increasing the discretization resolution, which reduces the difference.

6 Results, findings and discussion

One of the objects of this study was to seek validation or challenge the proposed PLC and HMI failure patterns. Manufacturers' MTBF values and handbook statistics are typically based on constant failure rate assumption. McLeod et al. (2015) stated that PLCs and computers follow the infant mortality pattern. Sometimes the constant rate is said to apply only for a so-called useful life period, which has typically undetermined or undisclosed duration. An exception to this is Texas Instruments with a statement that their certain processor has useful life of 100 000 power-on hours (11.4 years) in PLC applications, and after that the failure rate will increase (Eller 2015).

Three types of conclusions can be drawn from the data analysis. Failure pattern obtained from the survival analysis is a very steeply decreasing infant mortality pattern. Based on that, the reliability improves with the mission duration. This conclusion is backed up by the fact that many of the failures were recurrent. Alternatively, consecutive failures could have been aggregated to one failure as the root cause was probably the same for all. On the contrary, annual event count shows slightly increasing failure rate, and no signs of infant mortality. The slight increase in alarm rate may be related to older generation products or wear-out. It must be noted that age and mission time are different for repairable assets. Additionally, the survival regression revealed that failure rates vary between manufacturers. Ideally, the estimates should consider age, past failures and manufacturer. Development of such model is left for further work.

Causes of intermittent failures can be divided to four categories. Software errors that cause immediate failure have most likely decreasing rate. Failures from software errors causing volatile aging have increasing rate between reboots. External failure causes have constant rate. Intermittent failures due to hardware may have increasing rate.

In the literature review all of the PLC modules were recognized to have age-related wear-out mechanisms. Main life limiting factor of the power supply is decreased capacitance of electrolytic capacitors resulting from evaporation and leakage of the electrolyte. Optoisolator's lifetime is limited by LED degradation. Electromigration occurring in vias is the most critical wear-out mechanism of the processor. Identified limitation of these findings is that they may be valid for only certain products, and the model should represent wide population of devices from different manufacturers and years. For example, wear-out

due to electromigration is generalized from studies considering embedded processors, and PLCs are specifically designed to be ruggedized. PLC processors may have lower chip densities and larger cross-sectional areas of conductive paths, which lowers current densities making them less prone to electromigration. Also other wear-out mechanisms exist, and life limiting factors may vary between products.

If age-related failure mechanisms are not insignificant compared to random failures like voltage surges, then the constant failure rate assumption will not hold and there will be increasing failure rate. It should be noted that the ratio of internal failure causes to external causes is site specific. Non-constant failure rate is backed up by U.S. DoD, which has abandoned handbook statistics approach to reliability estimation of electronics. MIL-HDBK-217 and its progeny are recognized to contain grave deficiencies as they consider factors that actually induce wear-out as multipliers to a constant failure rate. Many degradation mechanisms accelerate significantly as temperature increases. Thus, maintaining a suitable temperature is probably the most important task that the operator can perform to achieve more reliable automation.

Quality of the maintenance records prevented validation of the FMEA and the construction of solely data-based model as initially intended. Hence, the FMEA is likely to contain failure modes that are not relevant in practice. Since majority of the work orders were vague and lacked information regarding failure causes and required actions, it was impossible to distinguish PLC and HMI related failures from other failures which cause alarms to the dispatch center. Examples of these other failures are sensor and actuator failures as well as communication failures caused by an external service provider. It would have been easier to identify relevant events if the studied system would have been the whole remote controlled automation system.

It was recognized that the used search criteria will not be able to capture all relevant events, and irrelevant events are probably included. Even though there are probably incorrect classifications, the estimate can still be unbiased assuming that misclassifications are independent of time and plant. Hence, the estimates can provide ordinal information and trends although the cardinal information would be incorrect. If quality of the work orders were improved, studies like this would be more successful in the future.

In addition to possible wear-out, also manufacturer related causes can be the driving force for renewals. If prolonged outages are considered unacceptable, the PLC and HMI renewal interval will likely be determined by obsolescence independently of the failure pattern. However, with the help of reliability models and detailed failure data, demand for spare parts could be predicted and spare parts could be purchased proactively. Then the product could be used after spare part availability has ended. Nonetheless, in addition to spare part unavailability, HMI PC's operating system can be the life limiting factor. Risk of a prolonged outage is increased also if backup copies of the program or its parameters are lost over the years. (Kurtti 2021)

7 Conclusion

Currently, the industry is experiencing so-called “Fourth Industrial Revolution”, which is strongly associated with integration of physical and digital systems. The amount of data that can be extracted from industrial processes has increased exponentially. (Carvalho et al. 2019) When processed and analysed, this data can provide valuable information. Winners of this revolution will probably be those, who best understand the value of the data and learn to take advantage of it. This will mean transition from still used "rules of thumb" and "expert judgement" towards more data-based decision making and modeling.

RCM is a structured analysis method used to identify safe and cost effective failure management strategies. Many unnecessary maintenance activities are still likely to be inherited from the beliefs of previous generations. RCM prevents this by demanding that all preventive interventions must be technically appropriate and worth doing. In other words, the planned intervention must have positive effect on reliability, and risk cost versus risk prevention cost assessment must indicate that the intervention is viable.

Physics of failure can be seen as the best reliability and RUL estimation method if the required effort is not considered. Paradoxically, the main disadvantage of the method is the expertise required, and its greatest advantage is the acquired expertise. It is also the only true prediction method that can be used before operating experience is available. Pioneers of the reliability engineering discipline, U.S. DoD and NASA, advocate physics of failure and have abandoned handbook statistics. Still, organizations with less resources for reliability assessment may have to settle for less. Artificial intelligence and stochastic cumulative damage models will evolve together with online condition monitoring technologies. Survival analysis remains as a solid method where event data is available.

Empirical part of the study consist of a FMEA, failure data analysis and modeling. 27 failure modes were recognized based on literature review. Prevalence of different failure modes could not be identified due to inadequate data quality. The issue was overcome by using values found from the literature, even though own data would have had the best representativeness. Failure pattern was studied by survival analysis, and a very steeply decreasing infant mortality curve was obtained. The observed pattern is mainly due to recurrent intermittent failures. Pattern of permanent failures could not be studied since it was not possible to distinguish permanent failures from intermittent failures. Contradictory

to the hazard function, annual event count showed slightly increasing trend. Explanation for this is that annual event count is a function of age, and hazard rate is a function of mission time, which resets when the asset is restored after a failure. Time from the last failure is more informative estimation input than age as many of the failures are recurrent. Furthermore, the effect of manufacturer was studied by survival regression, and it was observed that manufacturer has effect on reliability. The ideal model should consider both age and past failures as well as the manufacturer. Stochastic simulation model was compared to conventional event tree approach. Conventional event tree approach seemed to be sufficient for typical failure rates, but stochastic modeling may be useful for devices suffering from frequent failures.

References

- ABB. (2018). *Product lifecycle management*. Retrieved 5.7.2021, from https://library.e.abb.com/public/c2b8691e95f144c6b5d9133b3ee8c766/3ADR025047K0201_Rev2.pdf
- Abed, S., Tahar, K., & Brahim, A. (2016). *Thermodynamic and energy study of a regenerator in gas turbine cycle and optimization of performances*. International Journal of Energy Optimization and Engineering, 5, 25-44. doi:10.4018/IJEOE.2016040102
- Addepalli, S., Eiroa, D., Lieotrakool, S., François, A., Guisset, J., Sanjaime, D., et al. (2015). *Degradation study of heat exchangers*. Procedia Cirp, 38, 137-142. doi:10.1016/j.procir.2015.07.057
- Agustiady, T. K., & Cudney, E. A. (2016). *Total productive maintenance: Strategies and implementation guide* (1st ed.). Boca Raton, Florida, USA: CRC Press.
- Alphonsus, E., & Abdullah, M. (2016). *A review on the applications of programmable logic controllers (PLCs)*. Renewable and Sustainable Energy Reviews, 60, 1185-1205. doi:10.1016/j.rser.2016.01.025
- Antoun, C. (2018). *High voltage circuit breaker and power transformer failure modes and their detection*. 2018 Condition Monitoring and Diagnosis (CMD), 2018, 1-6. doi:10.1109/CMD.2018.8535655
- Arrell, D. (2006). *Next generation engineered materials for ultra supercritical steam turbines*. Siemens Westinghouse Power Corporation. doi:10.2172/896682
- Authén, S., Holmberg, J. E., Tyrväinen, T. & Zamani, L. (2015). *Guidelines for reliability analysis of digital systems in PSA context - Final report*. Stockholm, Sweden: Risk Pilot AB; Espoo, Finland: VTT Technical Research Centre of Finland Ltd. ISBN 978-87-7893-411-6.
- Azidehak, A. (2017). *Design of fault-tolerant controller for modular multi-level converters*. Ph. D dissertation, Electrical Engineering, North Carolina State University. Available: <https://repository.lib.ncsu.edu/handle/1840.20/35286>
- Babeshko, E., Kharchenko, V., & Gorbenko, A. (2008). *Applying F(I)MEA-technique for SCADA-based industrial control systems dependability assessment and ensuring*. Third International Conference on Dependability of Computer Systems DepCoS-RELCOMEX, 2008, 309-315. doi:10.1109/DepCoS-RELCOMEX.2008.23
- Babin, A., Polyakov, R., Savin, L., & Tyurin, V. (2020). *Statistical analysis of turbo generator sets failure causes*. Vibroengineering PROCEDIA, 31, 129-133. doi:10.21595/vp.2020.21331
- Basu, S. (2017). *Chapter IV - guided word hazard analysis*. In S. Basu (Ed.), *Plant hazard analysis and safety instrumentation systems*, 201-302. Academic Press. doi:10.1016/B978-0-12-803763-8.00004-2
- Bengtsson, M., & Lundström, G. (2018). *On the importance of combining “the new” with “the old” – One important prerequisite for maintenance in industry 4.0*. Procedia Manufacturing, 25, 118-125. doi:10.1016/j.promfg.2018.06.065
- Bolton, W. (2015). *Programmable logic controllers*. Oxford, United Kingdom: Newnes.
- Buhre, B., Gupta, R., Richardson, S., Sharma, A., Spero, C., & Wall, T. (2002). *PF-fired supercritical boilers - operational issues and coal quality impacts*. Callaghan, Australia: University of Newcastle.
- Böhmeke, G. (2020). *Wind Turbines*. PHYS-E6572 Advanced Wind Power Technology. Espoo, Finland: Aalto University.

- Carvalho, T. P., Soares, F. A., Vita, R., Francisco, R. D. P., Basto, J. P., & Alcalá, S. G. S. (2019). *A systematic literature review of machine learning methods applied to predictive maintenance*. Computers & Industrial Engineering, 137, 106024. doi:10.1016/j.cie.2019.106024
- Chaplin, R. A. (2009). *Steam turbine impulse and reaction blading*. Thermal Power Plants, 3. Fredericton, Canada: University of New Brunswick.
- Chen, Y., & Zheng, B. (2019). *What happens after the rare earth crisis: a systematic literature review*. Sustainability, 11(5), 1288. doi:10.3390/su11051288
- Chillarege, R., Biyani, S., & Rosenthal, J. (1995). *Measurement of failure rate in widely distributed software*. Twenty-Fifth International Symposium on Fault-Tolerant Computing. Digest of Papers, 1995, 424-433. doi:10.1109/FTCS.1995.466957
- Choi, K. C., Song, S. W., Park, G. M., & Hwang, S. J. (2012). *Reliability Analysis for Safety Grade PLC (POSAFE-Q)*, (IAEA-CN--194). International Atomic Energy Agency (IAEA).
- Clark, K. (2019). *Improving maintenance by adopting a P-F curve methodology*. Retrieved 17.3.2021, from <https://www.isa.org/intech-home/2019/march-april/features/improving-maintenance-by-adopting-a-p-f-curve-meth>
- Cox, D. R. (1972). *Regression models and life-tables*. Journal of the Royal Statistical Society: Series B (Methodological), 34(2), 187-202. doi:10.1111/j.2517-6161.1972.tb00899.x
- Dey, O., & Giri, B. C. (2014). *Optimal vendor investment for reducing defect rate in a vendor-buyer integrated system with imperfect production process*. International Journal of Production Economics; Celebrating a Century of the Economic Order Quantity Model, 155, 222-228. doi:10.1016/j.ijpe.2014.02.004
- Dilimulati, A., Stathopoulos, T., & Paraschivoiu, M. (2018). *Wind turbine designs for urban applications: A case study of shrouded diffuser casing for turbines*. Journal of Wind Engineering and Industrial Aerodynamics, 175, 179-192. doi:10.1016/j.jweia.2018.01.003
- Dorji, U., & Ghomashchi, R. (2014). *Hydro turbine failure mechanisms: An overview*. Engineering Failure Analysis, 44, 136-147. doi:10.1016/j.engfailanal.2014.04.013
- Doymaz, F., Romagnoli, J. A., & Palazoglu, A. (2001). *A strategy for detection and isolation of sensor failures and process upsets*. Chemometrics and Intelligent Laboratory Systems, 55(1-2), 109-123. doi:10.1016/S0169-7439(00)00126-X
- Dudziak, T., Hussain, T., & Simms, N. J. (2016). *High-temperature performance of ferritic steels in fireside corrosion regimes: Temperature and deposits*. Journal of Materials Engineering and Performance, 26, 84-93. doi:10.1007/s11665-016-2423-7
- EKE Electronics. (2017). *Human machine interface (HMI)*. Retrieved 13.7.2021, from <https://www.eke-electronics.com/human-machine-interface>
- Eller, R. (2015). *AM335x reliability considerations in PLC applications*. Texas Instruments Inc. Retrieved 11.7.2021, <https://www.ti.com/lit/an/sprabv9a/sprabv9a.pdf?ts=1626793365828>
- Exida. (2003). *Safety Equipment Reliability Handbook*. ISBN: 9780972723404.
- FIDES. (2009). *FIDES guide 2009 - Reliability methodology for electronic systems*. Retrieved 23.4.2021, from http://www.embedded.agh.edu.pl/www/fpga/dydaktyka/MPiMS/Data/UTE_FIDES_Guide_2009_-_Edition_A%20-%20September%202010_english_version.pdf

- Filios, A. E. (2013). *Hydraulic turbines*. Retrieved 20.5.2021, from <https://aefilios.wordpress.com/en-pages/resources/video-collect/turbomachines/hydraulic-turbines/>
- Garzon, R. D. (2002). *High voltage circuit breakers: Design and applications*. CRC Press.
- Girsang, I., Dhupia, J., Singh, M., Gevorgian, V., Muljadi, E., & Jonkman, J. (2014). *Impacts of providing inertial response on dynamic loads of wind turbine drivetrains*. IEEE Energy Conversion Congress and Exposition (ECCE), 2014, 1507-1514. doi:10.1109/ECCE.2014.6953597
- Goble, W. M., & Brombacher, A. C. (1999). *Using a failure modes, effects and diagnostic analysis (FMEDA) to measure diagnostic coverage in programmable electronic systems*. Reliability Engineering & System Safety, 66(2), 145-148. doi:10.1016/S0951-8320(99)00031-9
- Grottke, M., Matias Jr, R., & Trivedi, K. (2008). *The fundamentals of software aging*. IEEE International conference on software reliability engineering workshops (ISSRE Wksp), 2008, 1-6. doi:10.1109/ISSREW.2008.5355512
- Gündüz, N., Kufeoglu, S., & Lehtonen, M. (2018). *Customer interruption cost estimations for distribution system operators in Finland*. IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2018, 1-5. doi: 10.1109/ISGTEurope.2018.8571720
- Hansen, C. K., & Thyregod, P. (1992). *Component lifetime models based on Weibull mixtures and competing risks*. Quality and Reliability Engineering International, 8(4), 325-333. doi:10.1002/qre.4680080405
- Held, M., & Fritz, K. (2009). *Comparison and evaluation of newest failure rate prediction models: FIDES and RIAC 217Plus*. Microelectronics Reliability, 49(9-11), 967-971. doi: 10.1016/j.microrel.2009.07.031
- Holmberg, J. E. (2020). *Luotettavuus ja koherentit järjestelmät*. MS-E2117 Riskianalyysi. Espoo, Finland: Aalto University.
- International Organization for Standardization (ISO). (2014). *ISO/IEC guide 51:2014(en) safety aspects — guidelines for their inclusion in standards*. Retrieved 29.3.2021, from <https://www.iso.org/standard/53940.html>
- Jalote, P., Murphy, B., & Sharma, V. S. (2008). *Post-release reliability growth in software products*. ACM Transactions on Software Engineering and Methodology (TOSEM), 17(4), 1-20. doi:10.1145/13487689.13487690
- Jolliffe, I. T., & Cadima, J. (2016). *Principal component analysis: A review and recent developments*. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065), 20150202. doi:10.1098/rsta.2015.0202
- Kececioğlu, D. (2002). *Reliability engineering handbook: Its quantification and optimization*. Lancaster, Pennsylvania, USA: DEStech Publications.
- Klein, J., Klein, J. P., Houwelingen, J. C., Ibrahim, J. G., & Scheike, T. H. (2014). *Handbook of survival analysis (1st ed.)*. Boca Raton, Florida, USA: CRC Press.
- Korsah, K., Cetiner, S., Muhlheim, M., & Poore, W. (2010). *An investigation of digital instrumentation and control system failure modes*. ORNL/TM-2010/32. Oak Ridge, Tennessee, USA: Oak Ridge National Laboratory.
- Kumar, N., Besuner, P., Lefton, S., Agan, D., & Hilleman, D. (2012). *Power plant cycling costs*. Sunnyvale, California: National Renewable Energy Laboratory.

Kurtti, J. (2021). *Personal communication*. 26.7.2021.

Lei, Y., Li, N., Guo, L., Li, N., Yan, T. & Lin, J. (2018). *Machinery health prognostics: A systematic review from data acquisition to RUL prediction*. *Mechanical Systems and Signal Processing*, 104, 799-834. doi:10.1016/j.ymssp.2017.11.016

Letcher, T. M. (2017). *Wind energy engineering: A handbook for onshore and offshore wind turbines*. Academic Press.

Li, M. & Meeker, W. (2014). *Application of Bayesian Methods in Reliability Data Analyses*. *Journal of Quality Technology*. 46. 1-23. doi:10.1080/00224065.2014.11917951.

Lindsley, D., Grist, J., & Parker, D. (2018). *Thermal power plant control and instrumentation: The control of boilers and HRSGs: The control of boilers and HRSGs*. Stevenage: The Institution of Engineering and Technology.

Liu, W. (2014). *The failure analysis of the repeat gear tooth breakage in a 40MW steam turbine load gearbox and the butterfly in the carburized case*. *Engineering Failure Analysis*, 46, 9-17. doi:10.1016/j.eng-failanal.2014.07.024

Mahmoodian, M., & Li, C. Q. (2016). *Chapter 11 - stochastic failure analysis of defected oil and gas pipelines*. In A. S. H. Makhoul, & M. Aliofkhaezai (Eds.), *Handbook of materials failure analysis with case studies from the oil and gas industry*, 235-255. Butterworth-Heinemann. doi:10.1016/B978-0-08-100117-2.00014-5

Maisonnier, D. (2018). *RAMI: The main challenge of fusion nuclear technologies*. *Fusion Engineering and Design*, 136. doi:10.1016/j.fusengdes.2018.04.102

Malla, C., & Panigrahi, I. (2019). *Review of condition monitoring of rolling element bearing using vibration analysis and other techniques*. *Journal of Vibration Engineering & Technologies*, 7(4), 407-414. doi:10.1007/s42417-019-00119-y

Mandø, M. (2013). *4 - direct combustion of biomass*. In L. Rosendahl (Ed.), *Biomass combustion science, technology and engineering*, 61-83. Woodhead Publishing. doi:10.1533/9780857097439.2.61

Masche, M., Puig Arnavat, M., Wadenbäck, J., Clausen, S., Jensen, P., Ahrenfeldt, J., et al. (2017). *Full-scale milling tests of wood pellets for combustion in a suspension-fired power plant boiler*. Nordic Flame Days, Stockholm, Sweden, 2017.

Matic, Z., & Sruk, V. (2008). *The physics-of-failure approach in reliability engineering*. 30th International Conference on Information Technology Interfaces, 2008, 745-750. doi:10.1109/ITI.2008.4588504

McLeod, J., Bazuik, P., Calvo, R., & Rivera, S. (2015). *Failures profiles for maintenance in industrial facilities*. *Proceedings of the World Congress on Engineering*, 2015.

Menčík, J. (2016). *Concise reliability for engineers*. Rijeka: IntechOpen. doi:10.5772/62354

Mikus, M., Dziedzic, K., & Jurczyk, M. (2016). *Flue gas cleaning in municipal waste-to-energy plants – part i*. *Infrastructure and Ecology of Rural Areas*, IV, 1179–1193. doi:10.14597/infraeco.2016.4.1.086

Miller, B. G. (2005). *CHAPTER 6 - emissions control strategies for power plants*. In B. G. Miller (Ed.), *Coal energy systems*, 283-392. Burlington: Academic Press. doi:10.1016/B978-012497451-7/50006-1

Moleda, M., Momot, A., & Mrozek, D. (2020). *Predictive maintenance of boiler feed water pumps using SCADA data*. *Sensors*, 20(2), 571. doi:10.3390/s20020571

- Moubray, J. (2001). *Reliability-centered maintenance*. Industrial Press Inc.
- Muthen, B., & Muthen, L & Asparouhov, T. (2015). *Random coefficient regression*. Retrieved 16.5.2021, from https://www.statmodel.com/download/Random_coefficient_regression.pdf
- Muzakkir, S. M., Lijesh, K. P., & Hirani, H. (2015). *Failure mode and effect analysis of journal bearing*. International Journal of Applied Engineering Research, 10(16), 36843-36850.
- Nakamura, S., & Nakagawa, T. (2010). *Stochastic reliability modeling, optimization and applications*. Singapore: World Scientific.
- Nelson, W. (1980). *Accelerated life testing - step-stress models and data analyses*. IEEE transactions on reliability, 29(2), 103-108. doi:10.1109/TR.1980.5220742
- Nelson, W. B. (2009). *Accelerated testing: Statistical models, test plans, and data analysis*. John Wiley & Sons.
- Nomoto, H. (2017). *Solid particle erosion analysis and protection design for steam turbines*. Advances in Steam Turbines for Modern Power Plants, 219-239. Woodhead Publishing.
- Omron. (2021). *Technical explanation for power supplies*. Retrieved 12.7.2021, from https://www.ia.omron.com/data_pdf/guide/22/powersupply_tg_e_8_3.pdf
- Ostenso, A., & May, R. (1996). *Generic requirements specification for qualifying a commercially available PLC for safety-related applications in nuclear power plants Final report (EPRI-TR--107330)*. Palo Alto, California, USA: Electric Power Research Institute Inc.
- Padhi, N., & Lila, P. (2018). *Factory automation: A tool to create competitive advantage*. International Journal of Academic Research and Development, 3(5), 36-39. ISSN: 2455-4197.
- Panasonic. (2014). *PROGRAMMABLE CONTROLLERS - FP7 Digital I/O Units User's Manual*. Retrieved 13.7.2021, from https://www.panasonic-electric-works.com/cps/rde/xbcr/pew_eu_en/mn_fp7_digital_io_europe_en.pdf
- Pandey, M. D., Yuan, X., & van Noortwijk, J. M. (2009). *The influence of temporal uncertainty of deterioration on life-cycle management of structures*. Null, 5(2), 145-156. doi:10.1080/15732470601012154
- Parashar, B., & Taneja, G. (2007). *Reliability and profit evaluation of a PLC hot standby system based on a master-slave concept and two types of repair facilities*. Reliability, IEEE Transactions On Reliability, 56(3), 534-539. doi:10.1109/TR.2007.903151
- Pecht, P., Das, D., & Gaonkar, A. (2020). *Evaluation and comparison of FIDES and PoF-based EEE part reliability assessment*. College Park, Maryland, USA: University of Maryland.
- Peck, R. (2012). *PLC Power Supply Replacement*. Knoxville, Tennessee, USA: Control Technology Inc. Retrieved 12.7.2021, from <https://www.controltechnology.com/Files/Products/2500-Classic/2512/other/plc-power-supply-replacement>
- Pena, R., Clare, J. C., & Asher, G. M. (1996). *Doubly fed induction generator using back-to-back PWM converters and its application to variable-speed wind-energy generation*. IEE Proceedings-Electric Power Applications, 143(3), 231-241. doi:10.1049/ip-epa:19960288
- Pfaffel, S., Faulstich, S., & Rohrig, K. (2017). *Performance and reliability of wind turbines: A review*. Energies, 10(11), 1904. doi:10.3390/en10111904

- Ragheb, A., & Ragheb, M. (2010). *Wind turbine gearbox technologies*. 1st International Nuclear & Renewable Energy Conference (INREC), 2010, 1-8. doi: 10.1109/INREC.2010.5462549
- Ranjbar, K. (2010). *Effect of flow induced corrosion and erosion on failure of a tubular heat exchanger*. Materials & Design, 31(1), 613-619. doi:10.1016/j.matdes.2009.06.025
- Raqab, M. Z., Al-Awadhi, S., & Kundu, D. (2018). *Discriminating among Weibull, log-normal, and log-logistic distributions*. Communications in Statistics-Simulation and Computation, 47(5), 1397-1419. doi:10.1080/03610918.2017.1315729
- Regan, N. (2012). *The RCM solution: Reliability-centered maintenance*. New York, New York, USA: Industrial Press, Inc.
- Rizwan, S. M. (2006). *Reliability modeling strategy of an industrial system*. First International Conference on Availability, Reliability and Security (ARES'06), 2006, 6-630. doi:10.1109/ARES.2006.107
- Rodriguez, G. (2010). *Parametric survival models*. Princeton, New Jersey, USA: Princeton University.
- Sadiku, M., Shadare, A., Dada, E., & Musa, S. (2016). *Physics of failure: An introduction*. International Journal of Scientific Engineering and Applied Science, 2, 108-111. ISSN: 2395-3470
- Sakurai, S., & Isobe, N. (2010). *Life assessment for creep and fatigue of steam turbine components*. Transactions of the Indian Institute of Metals, 63(2), 281-288. doi:10.1007/s12666-010-0038-5
- Sandu, I. A., Salceanu, A., & Bejenaru, O. (2018). *New approach of the customer defects per lines of code metric in automotive SW development applications*. Journal of Physics: Conference Series, 2018, 1065(5). IOP Publishing.
- Schneider Electric. (2016). *Display specifications*. Retrieved 13.7.2021, from https://product-help.schneider-electric.com/Machine%20Expert/V1.1/en/SCUhw/SCUhw/Specifications_Intro/Specifications_Intro-4.htm
- Shavell, S. (2008). *Economics and liability for accidents*. New Palgrave Dictionary of Economics, 2nd Edition, Harvard Law and Economics Discussion Paper No. 535. SSRN: <https://ssrn.com/abstract=870565>
- Shin, J. (2008). *Lifetime reliability studies for microprocessor chip architecture*. Ph. D dissertation, Computer Engineering, University of Southern California. Available: <http://digitallibrary.usc.edu/digital/collection/p15799coll127/id/108618/>
- Si, X., Wang, W., Hu, C., & Zhou, D. (2011). *Remaining useful life estimation – A review on the statistical data driven approaches*. European Journal of Operational Research, 213(1), 1-14. doi:10.1016/j.ejor.2010.11.018
- Si, X., Wang, W., Hu, C., Chen, M., & Zhou, D. (2013). *A wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation*. Mechanical Systems and Signal Processing, 35(1), 219-237. doi:10.1016/j.ymssp.2012.08.016
- Siemens. (2014). *SIMATIC HMI Product Lifecycle Management (PLM)*. Retrieved 8.4.2021, from https://cache.industry.siemens.com/dl/files/124/91688124/att_64052/v1/2014_04_28_plm_hmi_en.pdf
- Siemens. (2021). *Product support*. Retrieved 13.7.2021, from [https://support.industry.siemens.com/cs/document/23493978/how-high-are-the-mtbf-\(mean-time-between-failure\)-values-for-the-simatic-panel-pcs-677-and-877-?dti=0&lc=en-WW](https://support.industry.siemens.com/cs/document/23493978/how-high-are-the-mtbf-(mean-time-between-failure)-values-for-the-simatic-panel-pcs-677-and-877-?dti=0&lc=en-WW)
- Siewiorek, D. P., & Swarz, R. S. (1998). *Reliable computer systems: Design and evaluation*. AK Peters/CRC Press.

- Sikorska, J. Z., Hodkiewicz, M., & Ma, L. (2011). *Prognostic modelling options for remaining useful life estimation by industry*. Mechanical Systems and Signal Processing, 25(5), 1803-1836. doi:10.1016/j.ymssp.2010.11.018
- Singh, R., & Shukla, A. (2014). *A review on methods of flue gas cleaning from combustion of biomass*. Renewable and Sustainable Energy Reviews, 29, 854-864. doi:10.1016/j.rser.2013.09.005
- Singpurwalla, N. D. (1995). The failure rate of software: Does it exist?. IEEE Transactions on Reliability, 1995, 44(3), 463-469. doi:10.1109/24.406582
- Sinha, A. K. (2010). *Aspects of failure of condenser tubes and their remedial measures at power plants*. AKS/Journal.
- Sinocruz, W. (2007). *Failure Modes and Effects Analysis (FMEA) for the Tricon Version 10.2 Programmable Logic Controller*. Invensys. Retrieved 3.3.2021, from <https://www.nrc.gov/docs/ML0932/ML093280223.pdf>
- Slama, J. B. H., Hellali, H., Lahyani, A., Louati, K., Venet, P., & Rojat, G. (2007). *Optocouplers ageing process: Study and modeling*. International Conference on Electrical Engineering Design and Technologies, Hammamet, 2007.
- Society of Automotive Engineers (SAE). (2009). Evaluation Criteria for Reliability-Centered Maintenance (RCM) Processes JA1011_200908. Retrieved 9.4.2021, from https://www.sae.org/standards/content/ja1011_200908/
- Sorrels, J. L., Randall, D. D., Schaffner, K. S., & Fry, C. R. (2019). *EPA Air Pollution Control Cost Manual*. Washington, D.C, USA: U.S. Environmental Protection Agency.
- Sullivan, L. (2016). *Survival analysis*. Boston, Massachusetts, USA: Boston University.
- Sun, Y. T., Luo, L. F., Zhang, Q., & Qin, X. R. (2019). *Reliability analysis of stochastic structure with multi-failure modes based on mixed copula*. Engineering Failure Analysis, 105, 930-944. doi:10.1016/j.eng-failanal.2019.06.021
- Sun, Y., Zhang, Q., Qin, X. R., & Lifu, L. (2018). *Multiple failure mode reliability modeling and analysis in failure crack propagation based on time-varying copula*. Journal of Mechanical Science and Technology, 32, 4637-4648. doi:10.1007/s12206-018-0911-4
- Thaduri, A., Verma, A. K., & Kumar, U. (2013). *Comparison of reliability prediction methods using life cycle cost analysis*. Proceedings Annual Reliability and Maintainability Symposium (RAMS), 2013, 1-7. doi:10.1109/RAMS.2013.6517747.
- Tinh, T., Choi, J., Jeon, D., Chu, J., Thomas, R., & Billinton, R. (2021). *A study on optimal reliability criterion determination for transmission system expansion planning*. KIEE International Transactions on Power Engineering, 5(1), 62-69.
- Torshizi S. E. M., & Jahangiri, A. (2018). *Analysis of Fatigue–Creep crack growth in the superheater header of a power plant boilers and estimation of its remaining lifetime*. Journal of Failure Analysis and Prevention, 18(1), 189-198. doi:10.1007/s11668-018-0400-1
- Troyer, D. (2018). *Reliability engineering principles for the plant engineer*. Retrieved 29.4.2021, from <https://supplychainminded.com/reliability-engineering-principles-for-the-plant-engineer/>
- UK Health and Safety Executive (HSE). (2021). *Risk management: Expert guidance*. Retrieved 22.3.2021, from <https://www.hse.gov.uk/managing/theory/index.htm>

- UK Office of Gas and Electricity Markets (OFGEM). (2017). *DNO common network asset indices methodology*. Retrieved 14.1.2021, from https://www.ofgem.gov.uk/sites/default/files/docs/2017/05/dno_common_network_asset_indices_methodology_v1.1.pdf
- U.S. Bureau of Reclamation (USBR). (2014). *Hydrogenerator Start / Stop Costs*. Retrieved 17.2.2021, from <https://www.usbr.gov/research/projects/detail.cfm?id=1880>
- U.S. Bureau of Reclamation (USBR). (2017). *Federal Replacements Units, Service Lives, Factors*. Retrieved 1.3.2021, from https://www.usbr.gov/power/data/2017_Federal_Hydropower_Replacements_Book_BW_1.1.pdf
- U.S. Department of Defense (DoD). (1995). *MIL-HDBK-217F Military Handbook: Reliability Prediction of Electronic Equipment*. Washington D.C., USA.
- U.S National Institute of Standards and Technology (NIST). (2013) *e-Handbook of Statistical Methods*. doi:10.18434/M32189
- U.S National Research Council. (2015). *Reliability growth: Enhancing defense system reliability*. Washington, D.C., USA: The National Academies Press. doi:10.17226/18987
- U.S. Naval Surface Warfare Center (NSWC). (2011). *Handbook of reliability prediction procedures for mechanical equipment*. West Bethesda, Maryland, USA.
- Veers, P., Dykes, K., Lantz, E., Barth, S., Bottasso, C. L., Carlson, O., et al. (2019). *Grand challenges in the science of wind energy*. Science, 366(6464). doi: 10.1126/science.aau2027
- Vega, A., Bose, P., & Buyuktosunoglu, A. (2017). *Rugged embedded systems: Computing in harsh environments*. Amsterdam, Netherlands: Morgan Kaufmann.
- Vundela, S., Kaushik, S., Tyagi, S., & Panwar, N. L. (2010). *An approach to analyse energy and exergy analysis of thermal power plants: A review*. Smart Grid and Renewable Energy, 1(3). doi:10.4236/sgre.2008.13019
- Wang, L., Chu, J., & Wu, J. (2007). *Selection of optimum maintenance strategies based on a fuzzy analytic hierarchy process*. International Journal of Production Economics, 107(1), 151-163. doi:10.1016/j.ijpe.2006.08.005
- Wang, L., Zhang, Z., Long, H., Xu, J., & Liu, R. (2016). *Wind turbine gearbox failure identification with deep neural networks*. IEEE Transactions on Industrial Informatics, 13(3), 1360-1368. doi:10.1109/TII.2016.2607179
- Webber, A. (2020). *Calculating useful lifetimes of embedded processors*. Texas Instruments Inc. Retrieved 11.7.2021, from https://www.ti.com/lit/an/sprabx4b/sprabx4b.pdf?ts=1626436409262&ref_url=https%253A%252F%252Fwww.google.com%252F
- Weibull, W. (1951). *A statistical distribution function of wide applicability*. Journal of Applied Mechanics, 18(3), 293-297.
- Wheeldon, J., & Shingledecker, J. P. (2013). *Materials for boilers operating under supercritical steam conditions*. Ultra-Supercritical Coal Power Plants, 81-103. doi:10.1533/9780857097514.1.81
- Xu, B., Li, H., Pang, W., Chen, D., Tian, Y., Lei, X., et al. (2019). *Bayesian network approach to fault diagnosis of a hydroelectric generation system*. Energy Science & Engineering, 7(5), 1669-1677. doi:10.1002/ese3.383

- Xu, Z., Mo, W., Gui, L., Ma, Z., & Xiao, X. (2020). *Practical test method for the sensitivity of programmable logic controller to voltage sags and short interruptions*. IET Circuits, Devices & Systems, 14(6), 830-837. doi:10.1049/iet-cds.2019.0490
- Yang, H., & Xu, H. (2011). *Reliability analysis of gas turbine based on the failure mode and effect analysis*. Asia-Pacific Power and Energy Engineering Conference, 2011, 1-4. doi:10.1109/APPEEC.2011.5749000.
- Ye, Z., & Chen, N. (2014). *The inverse gaussian process as a degradation model*. Technometrics, 56(3), 302-311. doi:10.1080/00401706.2013.830074
- Zagala, M., & Abdelaal, H. (2017). *Flue gas cleaning systems-A review paper*. Contemporary Problems of Power Engineering and Environmental Protection, 75-82.
- Zhang, P. (2008). *3 - system interfaces for industrial control*. In P. Zhang (Ed.), *Industrial control technology*, 259-427. Norwich, New York, USA: William Andrew Publishing. doi:10.1016/B978-081551571-5.50004-9
- Zhang, P. (2010). *Advanced industrial control technology* (1st ed.). Amsterdam, Netherlands: William Andrew/Elsevier.
- Zhang, X., Wu, X., Liu, R., Liu, J., & Yao, M. (2017). *Deformation-mechanism-based modeling of creep behavior of modified 9Cr-1Mo steel*. Materials Science and Engineering, 689, 345-352. doi:10.1016/j.msea.2017.02.044
- Zhao, Q., Jia, X., Guo, B. & Cheng, Z. (2018). *Real-time bayes estimation of residual life based on multisource information fusion*. Prognostics and System Health Management Conference (PHM-Chongqing), 2018, 215-222. doi:10.1109/PHM-Chongqing.2018.00043
- Ziegler, D., Puccinelli, M., Bergallo, B., & Picasso, A. (2013). *Investigation of turbine blade failure in a thermal power plant*. Case Studies in Engineering Failure Analysis, 1(3), 192-199. doi:10.1016/j.csefa.2013.07.002
- Zonta, T., da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., & Li, G. P. (2020). *Predictive maintenance in the industry 4.0: A systematic literature review*. Computers & Industrial Engineering, 150, 106889. doi:10.1016/j.cie.2020.106889

A. Appendix – Digital instrumentation and control failure modes

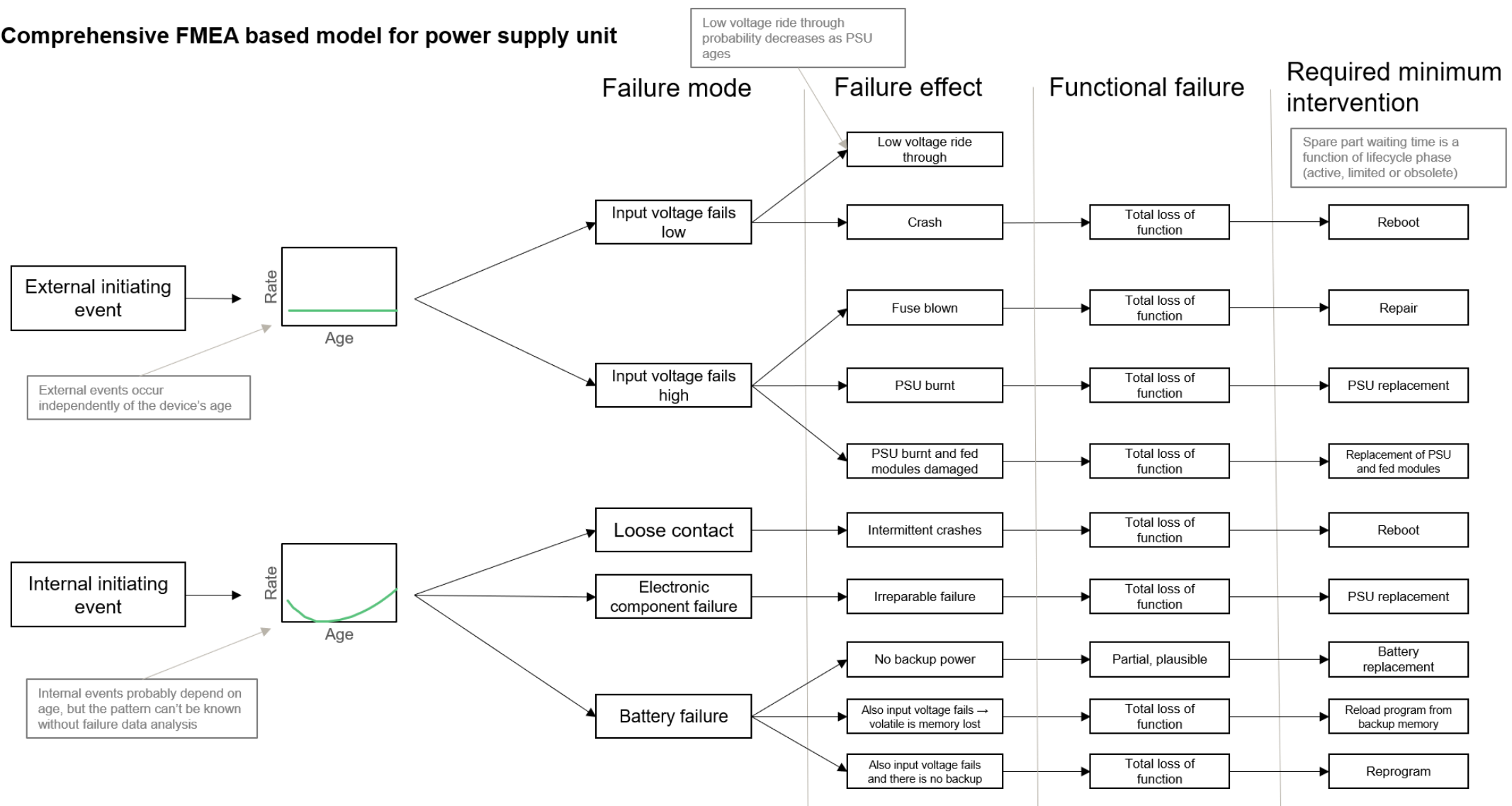
Digital instrumentation and control failure modes identified by Korsah et al. (2010):

Failure mode	Failure cause	Failure character
CPU lockup	Incompatibility of hardware	Detectable/Preventable before failure
Incorrect firmware coding	Programming error OR Requirements error/misinterpreted requirements	
Unresponsive in auto mode.	Incorrect interpretation of requirements	
Failure to communicate data to remote computer	Programming Error	
Encoder Output Error		
Instrument air pressure drop		
Task crash [Loss of asynchronous system traps (AST)]	Programming Error OR Incomplete Requirements Specifications	
Faulty program calculation	Requirements error	
Loss of communication (PLC)	Requirements error / Incomplete requirements description OR Misinterpreted Requirement	
Erroneous/false output	Human error	
Open breaker		
Loss of communication	Inadequate software V&V	
Erratic/unstable output		
Incorrect PLC output		
False output	Requirements error OR Incorrect interpretation of requirements	
Software lockup	Programming error OR Requirements error	
Communication lockout due to accumulation of timeout errors	Programming error AND/OR Operating outside specifications	
PPI unresponsive (lock up)	Operating Outside of Specifications	
Loss of communication		
Spurious performance (CPU board)		
NAND gate output failed in a quasi-trip state (would not provide a true “HI”)		
Open fuse (caused by voltage spike)		
Failed to establish communication	Installation error; also operating outside of specifications	
Degradation of battery	Degradation/Age-related	Age-Related
Voltage regulator card failed due to aging		
Degradation of UPS battery		
Out of Tolerance (drifting) due to unstable clock		
Short Circuit		
Incorrect functioning of CPU or clock		

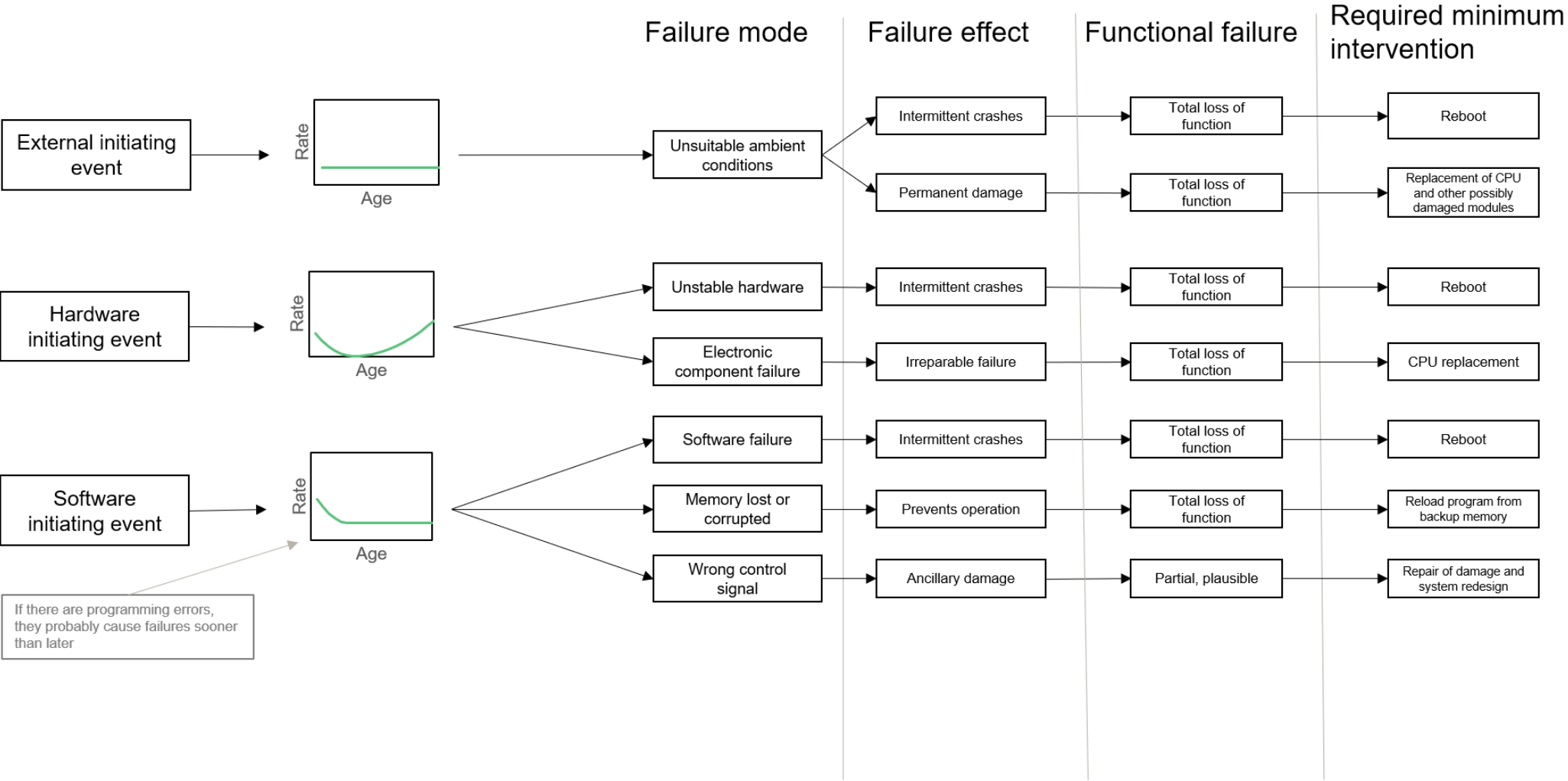
Loss of Vdc power		
Failure of Control Rod Element Assembly to move specified distance on command.		
Electrolytic capacitor failure (Actual mode of failure not specified)		
Damaged capacitors (mode of failure not indicated)		
Damaged components on output cards (actual failure mode not indicated)		
Spurious performance (isolator card)		
Intermittent Loss of Power	Equipment Aging	
Loss of Communication/ Common bus failure	Corrosion	
Degraded pulse-to-analog converter signal	NI	
Output degradation (due to static buildup)	Operating Outside of Specifications	
Erratic/fluctuating	Unknown	
Unable to reset	Unknown	
Communication Dropout/Loss of communication	Unknown	
Erratic Output	Unknown	
Component failure (actual failure mode not indicated)	Unknown	
Variable Frequency Drive controls failed (mode of failure not indicated)	Excessive traffic (interference or data storm) on the connected plant network	Random
“Open circuit/loss of continuity”	NI	
FPLA failed (mode of failure not indicated)	Unknown	
Tracking driver card output failed high	Unknown	
Loss of logical network connection	Operating beyond limited software resources	
No output indication	NI	
Communication Dropout	Maximum accrued timeouts	
Failed output of address decoder chip	Unknown	
Failed Output (high or low)	Unknown	
Network switch disconnected	Loss of power	
Unscheduled clock reset	Memory corruption of recorder software	
Computer lockup	Unknown	
PLC failed to reboot	Unknown	
Loss of memory	Battery failure	
Failed analog input card	Unknown	
Processor hang up	Unknown	
Shorted capacitor		
Shorted operational amplifier. Overpressure Delta-T setpoint failed high.	Electronic component failure	
Failed output (HI or LO)	Cold/bad solder joint	
Loss of trip signal	Failure of rotary switch or relay	
Periodic processor hang-up	Inadequate environmental control	Intermittent

B. Appendix – Comprehensive models

Comprehensive FMEA based model for power supply unit



Comprehensive FMEA based model for processor module



Comprehensive FMEA based model for I/O modules

