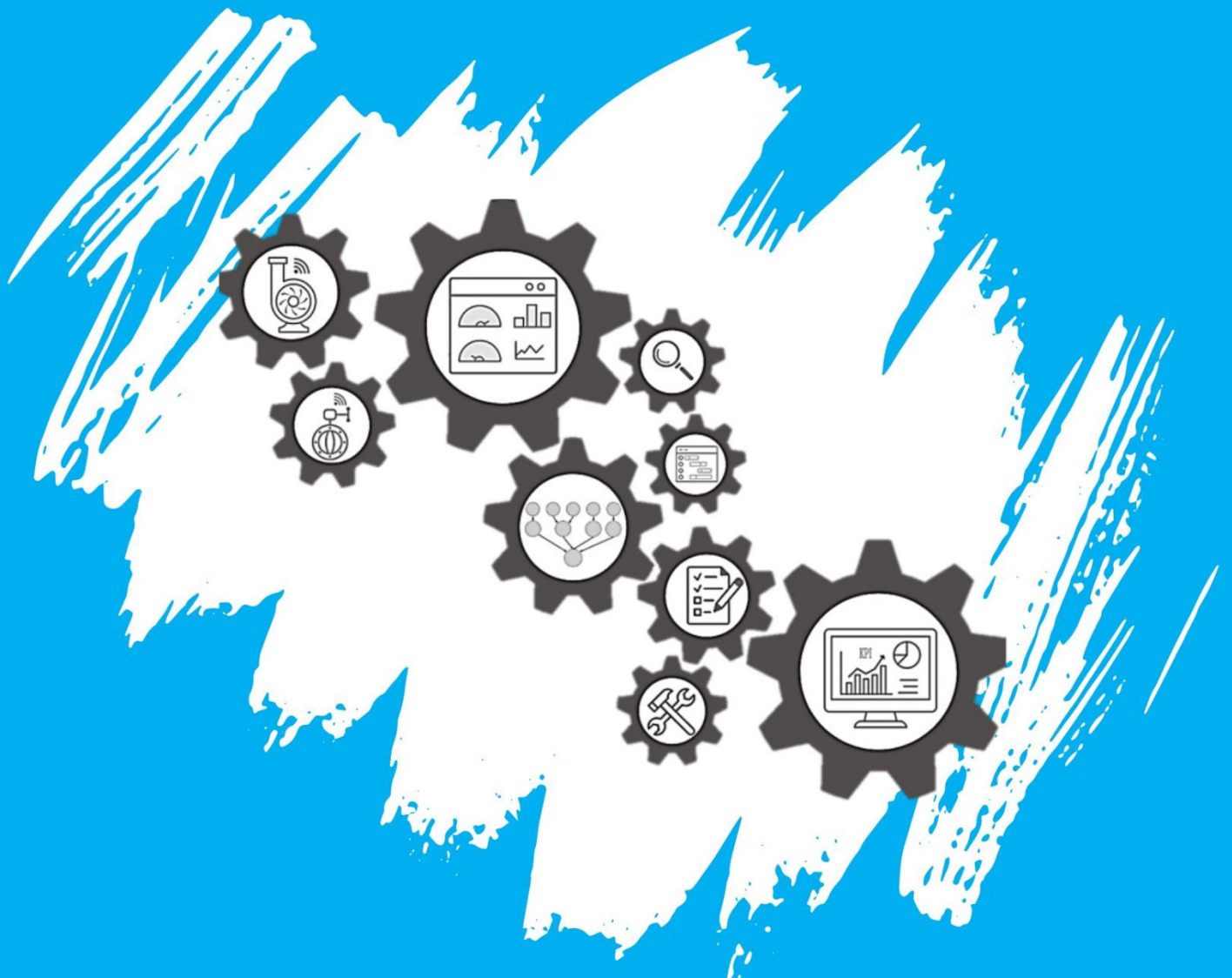


Reliability improvement of a soft drink production line using a Bayesian network

P.M.H. Freling



Reliability improvement of a soft drink production line using a Bayesian network

Master Thesis

By

P.M.H. Freling

in partial fulfilment of the requirements for the degree of

Master of Science
in Mechanical Engineering

Master track Multi-Machine Engineering
at the Department Maritime and Transport Technology
of Faculty Mechanical, Maritime and Materials Engineering
of Delft University of Technology

To be defended publicly on Wednesday September 27, 2023 at 10:00 AM.

Student number: 4558146
MSc track: Multi-Machine Engineering
Report number: 2023.MME.8809

Thesis committee:	Prof. dr. R. R. Negenborn	Chair, TU Delft
	Dr. ir. Y. Pang	Supervisor, TU Delft
	Ir. G. Monchen	Company supervisor, Bellt
	Ir. B. Zwerink	Company supervisor, Bellt
	Ir. M. Borsotti	Committee member, TU Delft

Date: September 20, 2023

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

It may only be reproduced literally and as a whole. For commercial purposes only with written authorization of Delft University of Technology. Requests for consult are only taken into consideration under the condition that the applicant denies all legal rights on liabilities concerning the contents of the advice.

Preface

Welcome dear reader,

The research in this master thesis focuses on developing a model to indicate how the reliability of a soft drink production line can be improved. The aim of the research is to provide insight into the possibilities of how a predictive maintenance strategy can contribute to improving reliability. This research is done for the company Batenburg Bellt. Furthermore, this graduation research falls under the supervision of Delft University of Technology.

From my interest in making everyday systems more efficient and reliable, the subject for my graduation research is something that fits in perfectly with this. I am grateful for the experiences I had during my graduation. They were special and instructive months marked by great challenges to arrive at this thesis. For this, I would like to thank Ben and Guido from the company Bellt. From the very first moment, they supported me with their knowledge and enthusiasm. Even at times when things were difficult, they were there to think along with me. I also want to thank Marcel, who was willing to answer all my questions about how the production line works in the factory and how the data had to be interpreted.

In addition, I would like to thank my supervisor at TU Delft, Dr. ir Y. Pang, for his patience and feedback during the process. I would also like to express my thanks for the feedback from Prof. dr. R.R. Negenborn during the milestones in this process.

Finally, I would like to thank my friends and family. Over the past months, they supported me with good advice, nice walks and positive energy. This helped me a lot during the many beautiful, but sometimes also very difficult moments throughout the process.

With the completion of my graduation project, my time at Delft University of Technology also officially comes to an end. Over the past few years, this was a place where I was allowed to grow as a person and make many wonderful friendships. I am looking forward to the new adventures and challenges that will now cross my path.

Enjoy reading my thesis!

Pauline Freling
September 2023

Summary

One of the biggest contemporary challenges within the beverage industry are the problems and consequences surrounding unplanned downtime in production lines. These are times when production lines shut down due to the failure of a component in the line. A production line within the beverage industry consists of tanks, pipes, valves and pumps. Once a valve or pump fails then the liquid cannot be pumped to the next tank. The moment the line shuts down during production, any liquid (syrup or soft drink) in the line has to be flushed out according to Food and Consumer Product Safety Authority rules. Unplanned stoppages not only cause the line to shut down at that time, but also cause delays in production planning. One of the causes of unplanned downtime can be traced back to the maintenance strategy used in the overall process industry and thus also in most beverage plants. The most common maintenance strategy within this industry is corrective maintenance. Here, maintenance is carried out only at the moment a component is broken. However, this does cause a lot of unplanned downtime and also causes high costs of flushing and overtime to catch up. Looking at other industries such as aviation and marine, these industries are many steps ahead in terms of proactive maintenance strategies. Proactive means that maintenance is performed even before a component has a chance to fail. In the aviation industry, this is a reassuring thought, after all, if a component suddenly fails there and, for example, the aircraft comes to a standstill, the consequences are incalculable. Within proactive maintenance strategies, there are a few more subcategories. There is preventive maintenance and condition-related maintenance. Preventive maintenance is often schedule-based, but does not look at the condition of the component. If the schedule says that the component needs to be replaced then it happens. Condition-related maintenance works on the basis of condition data on, for example, the valves and pumps to determine what the condition is and then make a prediction on how long the component can still function before it will fail. This is also known as predictive maintenance and has been on the rise in recent years. Within the beverage industry, this development of predictive maintenance is still hardly used, one of the reasons being that it is perceived as very difficult to develop a predictive maintenance strategy. In addition, it requires investment but is not yet seen as a sustainable investment that saves money on maintenance in the long run. However, predictive maintenance is already known to help reduce unplanned downtime and improve the reliability of components in the line, and so the reliability of the line itself.

Line reliability is defined in this thesis as: "The ability of a system or component to perform its required functions for a specified time under specified conditions". This research will look at what is already being done within the beverage industry, particularly focusing on the soft drink industry, to move from corrective to predictive maintenance. It will also look at what benefits this would bring to soft drink manufacturers. This research focuses on the question:

How can a predictive maintenance strategy contribute to improving the reliability of a soft drink production line?

For this research, the focus is on a production line at a soft drink manufacturer in the Netherlands. Within this production line are several valves and pumps. These are the components within the scope of this research. All other components are left out of consideration.

First, it will be examined what data is already being collected regarding the condition of the components within the production line. Next, the methods for using the condition data to create a model for predicting the failure probability of a component are examined. This research has shown that there is only limited data available that can be used to determine the condition of the valves. The condition data used for the valves is the time it takes the valve to open or close, this is called the looptijd. It is known that if the looptijd increases, i.e. the valve takes longer to open or close, this indicates that valve needs maintenance.

No condition data is available on the pumps. The choice was made to determine, based on literature, which parameters are needed to determine the condition of the pumps. The parameters used in this research for the condition of the pumps are the probability of cavitation, which is determined using the Net Positive Suction Head margin, and at how much percent of the Best Efficiency Point the pump is operating at. Cavitation is the formation and imploding of vapour bubbles on the impeller of the pump, this puts a lot of forces on the impeller and this causes fatigue phenomena and reduction in service life.

A pump functions best when it operates near its BEP. The further the pump is away from its BEP, the more likely it is to fail and the shorter its lifetime is.

In this research, the best method to fit the available data and knowledge was considered in order to arrive at a model to predict whether a component needs maintenance. This revealed that a Bayesian Network is the most accessible method for this research. In the BN, the parameters and all components in the line are represented as nodes with connections to indicate the relationships between them. This allows determining, based on the condition data collected from the line, and a synthetic dataset for the pumps, the probability of failure of a component and therefore the probability of unplanned downtime of the line. The model, with the BN method, is coded in Python using the PGMPY package.

To demonstrate the impact of using or not using the model on production line reliability, two situations are considered. Both situations work with a production schedule for a year, this is based on the available historical production

schedules. In the first situation, the model is not used, assuming only corrective maintenance and a small part of preventive maintenance in the two production free weeks per year.

In the second situation, the model is used and predictive maintenance and preventive maintenance are used based on the condition data. Because there was not enough data available from the production line regarding the condition monitoring data (running time, probability of cavitation and at what percentage of BEP), it was chosen to work with a synthetic dataset in this research. That is, a data set is created for the condition monitoring data based on historical data and assumptions from the literature. For this, historical process data is also used to determine how long a particular process takes and what products are produced across the line. Before starting production, the operator can enter this synthetic process data set for the valves and pumps into the model. After entering this data into the model, feedback is provided on the probability of failure and which components, if any, could fail during the process. Based on this data, the operator can decide whether or not to start the process.

To indicate how the model affects the reliability of the production line, KPIs are used. The first KPI looks at the ratio between the number of hours of planned maintenance performed during planned downtime hours compared to the total number of hours of downtime, also called the Maintenance Downtime Index (MDI). The total number of hours of downtime is the number of hours of planned downtime plus the number of hours of unplanned downtime. Line reliability increases as the KPI value increases.

The second, third and fourth KPI look at maintenance costs. It determines what percentage of the total maintenance cost goes on corrective, preventive and predictive maintenance. If a larger percentage is spent on predictive maintenance compared to corrective maintenance, this is an indication that the reliability of the line is increasing. Based on these KPIs, the two situations can be compared.

The results showed that line reliability improves when the developed model is used. In the case of the first KPI, there are fewer total hours of downtime. This is because there are fewer unplanned downtime hours. In addition, more hours of downtime were used to carry out planned maintenance. For the situation where the model is used, there is a 5% reduction in downtime hours. Looking at the cost KPIs, it can be seen that the total maintenance costs in the situation where the model is used are 53% lower than in the situation without the model. In addition, the percentage of corrective maintenance costs drops dramatically. The results from the KPIs indicate that line reliability improves the moment predictive maintenance is used instead of corrective maintenance.

Samenvatting

Een van de grootste hedendaagse uitdagingen binnen de beverage industrie zijn de problemen en gevolgen rondom ongeplande stilstand in de productielijnen. Dit zijn momenten waarop de productielijnen stilvallen als gevolg van het falen van een component in de lijn. Een productielijn binnen de beverage industrie bestaat uit tanks, leidingen, kleppen en pompen. Zodra er een klep of een pomp faalt dan kan de vloeistof niet naar de volgende tank worden gepompt. Op het moment dat de lijn stilvalt tijdens het produceren moet alle vloeistof (siroop of frisdrank) die in de lijn zit worden weggespoeld volgens de regels van de Voedsel en Waren autoriteit. Ongeplande stilstanden zorgen niet alleen voor het stil vallen van de lijn op dat moment, maar ook voor vertraging in de productie planning. Een van de oorzaken van de ongeplande stilstanden is te herleiden naar de onderhoudsstrategie die wordt gebruikt in de algehele proces industrie en daarmee ook in de meeste beverage fabrieken. De meest voorkomende onderhoudsstrategie binnen deze industrie is correctief onderhoud. Hierbij wordt er pas onderhoud uitgevoerd op het moment dat er een component kapot is. Dit zorgt echter wel voor veel ongeplande stilstand en zorgt ook voor hoge kosten van het wegspoelen en de overuren om de opgelopen achterstand in te lopen. Kijkend naar andere industrieën zoals de luchtvaart en de marine lopen deze industrieën vele stappen voor op het gebied van proactieve onderhoudsstrategieën. Proactief wil zeggen dat er al onderhoud wordt uitgevoerd voordat een component de kans krijgt om te falen. In de luchtvaart industrie is dat een geruststellende gedachte, immers als er daar ineens een component faalt en bijvoorbeeld het vliegtuig valt stil dan zijn de gevolgen niet te overzien.

Binnen de proactieve onderhoud strategieën zijn er nog enkele vertakkingen. Er is preventief onderhoud en conditie gerelateerd onderhoud. Preventief onderhoud is vaak op basis van een schema, maar kijkt niet naar de conditie van het component. Als er op het schema staat dat het component vervangen moet worden dan gebeurt dat. Conditie gerelateerd onderhoud werkt op basis van conditie data over bijvoorbeeld de kleppen en de pompen om te bepalen wat de toestand is en daarmee een voorspelling te doen over hoe lang het component nog kan functioneren alvorens het zal falen. Dit wordt ook wel voorspellend onderhoud genoemd en is de laatste jaren bezig aan een opmars. Binnen de beverage industrie wordt er nog nauwelijks gebruik gemaakt van deze ontwikkeling, een van de redenen hiervan is dat het als zeer moeilijk wordt ervaren om een voorspellend onderhoudsstrategie te ontwikkelen. Daarnaast moet er worden geïnvesteerd, maar wordt het nog niet gezien als een duurzame investering die op de lange termijn geld bespaart op onderhoud.

Van het voorspellend onderhoud is echter al bekend dat het bijdraagt aan het reduceren van ongeplande stilstand en de betrouwbaarheid van de componenten in de lijn, en daarmee ook de betrouwbaarheid van de lijn zelf, verbetert. De betrouwbaarheid van de lijn is in dit onderzoek gedefinieerd als: "Het vermogen van een systeem of onderdeel om gedurende een bepaalde tijd onder bepaalde omstandigheden de vereiste functies te vervullen". In dit onderzoek zal er worden gekeken naar wat er al gedaan wordt binnen de beverage industrie, met name gericht op de frisdrank industrie, om van correctief naar voorspellend onderhoud te gaan. Ook wordt er gekeken naar welke voordelen dat zou opleveren voor de frisdrank fabrikanten. Dit onderzoek richt zich op de vraag:

Hoe kan een voorspellend onderhoudsstrategie bijdragen aan het verbeteren van de betrouwbaarheid van de frisdrank productielijn?

Voor dit onderzoek ligt de focus op een productielijn bij een frisdrank fabrikant in Nederland. Binnen deze productielijn zitten meerdere kleppen en pompen. Dit zijn de componenten die binnen de scope van dit onderzoek vallen. Alle andere componenten worden buiten beschouwing gelaten.

Allereerst wordt er gekeken welke data er alreeds wordt verzameld met betrekking tot de conditie van de componenten binnen de productielijn. Vervolgens wordt er gekeken welke methodes er zijn om met de conditie data tot een model te komen waarmee er voorspellingen gedaan kunnen worden over de faalkans van een component. In dit onderzoek is naar voren gekomen dat er slechts gelimiteerd data beschikbaar is die kan worden gebruikt voor het bepalen van de conditie van de kleppen. De conditie data die voor de kleppen wordt gebruikt is de tijd die de klep erover doet om open of dicht te gaan, dit heet de looptijd. Er is bekend dat als de looptijd hoger wordt, dat wil zeggen dat de klep er langer over doet om open of dicht te gaan, dit aangeeft dat klep onderhoud nodig heeft.

Voor de pompen is geen conditie data beschikbaar. Hier is gekozen om op basis van literatuur te bepalen welke parameters nodig zijn om de conditie van de pompen te kunnen vast stellen. De gebruikte paramaters in dit onderzoek voor de conditie van de pompen zijn de kans op cavitatie, wat wordt bepaald met de Net Positive Suction Head marge, en op hoeveel procent van het Best Efficiency Point de pomp opereert. Cavitatie is het vormen en imploderen van dampbellen op de waaier van de pomp. Hierdoor komen er veel krachten op de waaier terecht en dat zorgt voor vermoeiingsverschijnselen en reductie van levensduur.

Een pomp functioneert het beste als deze in de buurt van zijn BEP werkt. Hoe verder de pomp van het BEP af zit hoe meer kans op falen en hoe korter de levensduur van de pomp is.

In dit onderzoek is er gekeken met behulp van het onderzoek van Sikorska naar welke methode het beste past bij de beschikbare data en kennis om op basis daarvan tot een model te komen waarmee kan worden voorspeld of een component onderhoud nodig heeft. Hier is uit voort gekomen dat een Bayesian Network voor dit onderzoek de meest toegankelijke methode is. In het BN zijn de parameters en alle componenten in de lijn weergegeven als nodes met verbindingen om de onderlinge relaties aan te duiden. Hiermee kan er worden bepaald op basis van de conditie data

die uit de lijn wordt verzameld, en een synthetische dataset voor de pompen, wat de kans is op falen van een component en daarmee de kans op ongeplande stilstand van de lijn. Het model, met daarin het BN, is gemodelleerd in Python met het PGMPY package.

Om aan te kunnen tonen wat de invloed is van het wel of niet gebruiken van het model op de betrouwbaarheid van de productielijn wordt er gekeken naar twee situaties. Beide situaties werken met een productie planning voor een jaar, dit is gebaseerd op de beschikbare historische productieplanningen. In de eerste situatie wordt het model niet gebruikt, hiermee wordt er aangenomen dat er alleen wordt gewerkt met correctief onderhoud en een klein stukje preventief onderhoud in de twee productie vrije weken per jaar.

In de tweede situatie wordt er wel gebruik gemaakt van het model en wordt er gewerkt met voorspellend onderhoud en preventief onderhoud op basis van de conditie data. Omdat er niet voldoende data beschikbaar was van de productielijn omtrent de conditie monitoring data (looptijd, kans op cavitatie en op hoeveel procent van het BEP), is ervoor gekozen om in dit onderzoek te werken met een synthetische dataset. Dat wil zeggen dat er op basis van historische data en aannames uit de literatuur een data set wordt gecreëerd voor de conditie monitoring data. Hiervoor wordt ook de historische proces data gebruikt om te bepalen hoe lang een bepaald proces duurt en welke producten er over de lijn worden geproduceerd. Alvorens een productie wordt gestart kan de operator deze synthetische proces dataset voor de kleppen en de pompen invoeren in het model. Na het invoeren van deze data in het model wordt er teruggekoppeld wat de kans op falen is en welke componenten eventueel kunnen falen gedurende het proces. Op basis van deze data kan de operator beslissen over het wel of niet starten van het proces.

Om aan te kunnen geven wat de invloed is van het model op de betrouwbaarheid van de productielijn wordt er gewerkt met Key performance indicators (KPIs). De eerste KPI kijkt naar de ratio tussen het aantal uur gepland onderhoud dat wordt uitgevoerd tijdens geplande stilstand uren ten opzichte van het totaal aantal uur stilstand. Het totaal aantal uur stilstand is het aantal uur geplande uur stilstand plus het aantal uur ongeplande stilstand, ook wel genaamd de Maintenance Downtime Index (MDI). De betrouwbaarheid van de lijn neemt toe als de KPI waarde hoger wordt.

De tweede, derde en vierde KPI kijken naar de onderhoudskosten. Er wordt bepaald hoeveel procent van de totale onderhoudskosten op gaat aan correctief, preventief en voorspellend onderhoud. Indien er een groter percentage opgaat aan voorspellend onderhoud in vergelijking met correctief onderhoud dan is dit een indicatie dat de betrouwbaarheid van de lijn toeneemt.

Op basis van deze KPIs kunnen beide situaties worden vergeleken.

Uit de resultaten is naar voren gekomen dat de betrouwbaarheid van de lijn verbeterd als het ontwikkelde model wordt gebruikt. In het geval van de eerste KPI is er minder totaal aantal uur stilstand. Dit komt omdat er minder ongeplande stilstand uren zijn. Daarnaast zijn er meer uren van de stilstand gebruikt om gepland onderhoud uit te voeren. Voor de situatie waarin het model wordt gebruikt is er een reductie van 5% op het aantal stilstand uren. Als er wordt gekeken naar de kosten KPIs dan valt daaruit op te maken dat de totale onderhoudskosten in de situatie waarbij het model wordt gebruikt 53% lager zijn dan in de situatie zonder model. Daarnaast daalt het percentage van de correctief onderhoudskosten drastisch. De resultaten uit de KPIs wijzen erop dat de betrouwbaarheid van de lijn verbeterd op het moment dat er wordt gewerkt met voorspellend onderhoud in plaats van correctief onderhoud.

List of figures

Figure 1: Overview beverage industry, with highlight on the carbonated soft drinks [2]	1
Figure 2: Early soft drink production manual equipment [8]	2
Figure 3: Different lines in soft drink factory defined by Koss [9]	2
Figure 4: Production line, the valves are indicated with the blue circles and de pumps with orange circles [2]	3
Figure 5: Schematic overview of the production line for this research	6
Figure 6: Double seat valve [23]	7
Figure 7: Double seat valve [24]	7
Figure 8: Butterfly valve [26]	8
Figure 9: Butterfly valve [27]	8
Figure 10: Centrifugal pump schematic [33]	9
Figure 11: Centrifugal pump [34]	9
Figure 12: Overview of types of downtimes [36]	9
Figure 13: Overview of different maintenance strategies based on EN13306 [43]	12
Figure 14: Overview of different prognostic PdM models [51]	16
Figure 15: Pie chart of the hours of the production line	19
Figure 16: Pie chart of time production line in use hours	20
Figure 17: Pie chart of downtime hours	20
Figure 18: Overview of valve failures 01-02-2021 till 03-02-2023	21
Figure 19: Increasing looptijd of valve V5.2	22
Figure 20: Graph of logistic function with different m-values	23
Figure 21: Cavitation occurrence [88] [89]	24
Figure 22: NPSH margin [97] [98]	25
Figure 23: BEP, POR and AOR [96]	25
Figure 24: Pump curve reliability [101]	27
Figure 25: Pump curve reliability Barringer [100] [103]	27
Figure 26: Pump curves of pumps P _{4.1} and P _{5.1}	28
Figure 27: LEM overview [72]	29
Figure 28: LEM overview [72], with criteria for selection	33
Figure 29: Example BN	33
Figure 30: Multi-layer structure [135]	39
Figure 31: BN with multi-layer structure	40
Figure 32: Line segment 1, part of BN	40
Figure 33: States of B _{Px,x}	41
Figure 34: Example of logistic function	43
Figure 35: Last section of BN	45
Figure 36: Best case scenario PL(0)	49
Figure 37: Worst case scenario PL(0)	49
Figure 38: Combination of Figure 35 and Figure 36	50
Figure 39: Reliability and maintenance costs [151]	54
Figure 40: Costs for different maintenance strategies	54
Figure 41: Flow chart	60
Figure 42: Overview of failure and maintenance without model	64
Figure 43: Pie chart of maintenance hours, without model	65
Figure 44: BN output PL(1) for warning and failure values	66
Figure 45: BN output LS(1) for warning and failure values	66
Figure 46: Overview of time used by every component on moment of maintenance	67
Figure 47: Pie chart of maintenance hours with model	68
Figure 48: Comparison of maintenance hours with and without model	68
Figure 49: Comparison of maintenance costs with and without model	69

Figure 50: Comparison of type of maintenance costs with and without
Figure 51: Steps of CIP

69
91

List of tables

Table 4.1: Values of nodes for calculations example BN	36
Table 5.1: Values best and worst case scenarios for all different combinations of LT nodes of LS3	48
Table 6.1: Warning and potential failure values of PL and LS nodes	57
Table 6.2: Summary of all the costs	58
Table 6.3: Summary of the hours for inspection and maintenance	58
Table 7.1: Maintenance costs without model	65
Table 7.2: Maintenance costs with model	68
Table 7.3: Comparison of maintenance hours with and without model	70
Table 7.4: Results cost KPIs	70

Abbreviations

AI	Artificial Intelligence
ANN	Artificial Neural Networks
AOR	Allowable Operating Region
ARIMA	Auto Regressive Integrated Moving Average
ARMA	Auto Regressive Moving Average
ARMAX	Auto Regressive Moving Average model with eXogenous inputs
BEP	Best Efficiency Point
BN	Bayesian Network
CBM	Condition-based Maintenance
CIP	Clean In Place
CPT	Conditional Probability Table
DAG	Directed Acyclic Graph
DL	Deep Learning
DT	Downtime
KPI	Key Performance Indicator
LEM	Life Expectancy Models
MDI	Maintenance Downtime Index
MLE	Maximum Likelihood Estimation
MTBF	Mean Time Between Failure
NPSH _A	Net Positive Suction Head Available
NPSH _R	Net Positive Suction Head Required
PDF	Probability Density Function
PdM	Predictive Maintenance
PHM	Proportional Hazard Model
PLC	Programmable Logic Controller
POR	Preferred Operating Region
RUL	Remaining Useful Life

Symbols / Variables

LT	Looptijd
Ω	Set of states
LT _{V1.1}	Looptijd of line segment 1 valve 1
LT _{V1.2}	Looptijd of line segment 1 valve 2
BP _{1.1}	Parameter BEP of line segment 1 pump 1
GP _{1.1}	Parameter cavitation of line segment 1 pump 1
V _{1.1}	Valve; line segment 1 valve 1
V _{1.2}	Valve; line segment 1 valve 2
P _{1.1}	Pump; line segment 1 pump 1
LS _I	Line segment 1 (or I)
LT _{V2.1}	Looptijd of line segment 2 valve 1
LT _{V2.2}	Looptijd of line segment 2 valve 2
BP _{2.1}	Parameter BEP of line segment 2 pump 1
BP _{2.2}	Parameter BEP of line segment 2 pump 2
GP _{2.1}	Parameter cavitation of line segment 2 pump 1
GP _{2.2}	Parameter cavitation of line segment 2 pump 2
V _{2.1}	Valve; line segment 2 valve 1
V _{2.2}	Valve; line segment 2 valve 2
P _{2.1}	Pump; line segment 2 pump 1
P _{2.2}	Pump; line segment 2 pump 2
LS _{II}	Line segment 2 (or II)
LT _{V3.1}	Looptijd of line segment 3 valve 1
LT _{V3.2}	Looptijd of line segment 3 valve 2
LT _{V3.3}	Looptijd of line segment 3 valve 3
LT _{V3.4}	Looptijd of line segment 3 valve 4
LT _{V3.5}	Looptijd of line segment 3 valve 5
LT _{V3.6}	Looptijd of line segment 3 valve 6
LT _{V3.7}	Looptijd of line segment 3 valve 7
LT _{V3.8}	Looptijd of line segment 3 valve 8
BP _{3.1}	Parameter BEP of line segment 3 pump 1
GP _{3.1}	Parameter cavitation of line segment 3 pump 1
P _{3.1}	Pump; line segment 3 pump 1
V _{3.1}	Valve; line segment 3 valve 1
V _{3.2}	Valve; line segment 3 valve 2
V _{3.3}	Valve; line segment 3 valve 3
V _{3.4}	Valve; line segment 3 valve 4
V _{3.5}	Valve; line segment 3 valve 5
V _{3.6}	Valve; line segment 3 valve 6
V _{3.7}	Valve; line segment 3 valve 7
V _{3.8}	Valve; line segment 3 valve 8
LS _{III}	Line segment 3 (or III)
LT _{V4.1}	Looptijd of line segment 4 valve 1
LT _{V4.2}	Looptijd of line segment 4 valve 2
LT _{V4.3}	Looptijd of line segment 4 valve 3
LT _{V4.4}	Looptijd of line segment 4 valve 4
LT _{V4.5}	Looptijd of line segment 4 valve 5
BP _{4.1}	Parameter BEP of line segment 4 pump 1
GP _{4.1}	Parameter cavitation of line segment 4 pump 1
P _{4.1}	Pump; line segment 4 pump 1
V _{4.1}	Valve; line segment 4 valve 1
V _{4.2}	Valve; line segment 4 valve 2
V _{4.3}	Valve; line segment 4 valve 3
V _{4.4}	Valve; line segment 4 valve 4
V _{4.5}	Valve; line segment 4 valve 5
LS _{IV}	Line segment 4 (or IV)
LT _{V5.1}	Looptijd of line segment 5 valve 1

LT _{V5.2}	Looptijd of line segment 5 valve 2
LT _{V5.3}	Looptijd of line segment 5 valve 3
LT _{V5.4}	Looptijd of line segment 5 valve 4
BP _{5.1}	Parameter BEP of line segment 5 pump 1
GP _{5.1}	Parameter cavitation of line segment 5 pump 1
P _{5.1}	Pump; line segment 5 pump 1
V _{5.1}	Valve; line segment 5 valve 1
V _{5.2}	Valve; line segment 5 valve 2
V _{5.3}	Valve; line segment 5 valve 3
V _{5.4}	Valve; line segment 5 valve 4
LS _V	Line segment 5 (or V)
JP ₁	Joint probability line segment 1
JP ₂	Joint probability line segment 2
JP ₃	Joint probability line segment 3
JP ₄	Joint probability line segment 4
JP ₅	Joint probability line segment 5
PL	Production line

Table of contents

Preface	I
Summary	II
Samenvatting	IV
List of figures	VI
List of tables	VIII
Abbreviations	IX
Symbols / Variables	X
1. Introduction	1
1.1 History of soft drinks.....	1
1.2 Production line.....	3
1.3 Problem definition.....	3
1.4 Research objectives	4
1.5 Research questions.....	4
1.6 Methodology.....	5
1.7 Thesis outline	5
2. Production process and components	6
2.1 Production process.....	6
2.1.1 Production line case study.....	6
2.2 Clean-In-Place.....	7
2.3 Valves and Pumps	7
2.3.1 Valves; types and working principles	7
2.3.2 Valve failures.....	8
2.3.3 Pumps; working principle.....	8
2.3.4 Pump failures	9
2.4 Different types of downtimes	9
2.4.1 Condition monitoring of valves.....	10
2.4.2 Condition monitoring of pumps	10
2.4.3 Case study at soft drink factory	10
2.5 Current maintenance strategy and requirements	10
2.6 Conclusion	11
3. Maintenance strategies	12
3.1 Maintenance standard EN 13306.....	12
3.1.1 Corrective maintenance.....	12
3.1.2 Preventive maintenance	13
3.1.3 Predetermined maintenance.....	13
3.1.4 Condition-based maintenance	13
3.1.5 Predictive maintenance	13
3.1.6 Non predictive condition-based maintenance	14
3.1.7 Prescriptive maintenance	14
3.2 Maintenance strategy for this problem	14
3.3 Predictive maintenance	14
3.3.1 Diagnostics	15
3.3.2 Prognostics	16

3.4 Model selection	17
3.5 Conclusion	18
4. Model development	19
4.1 Data analysis.....	19
4.1.1 Overall line analysis	19
4.1.2 Valve analysis	21
4.1.3 Pump analysis.....	23
4.1.3.1 NPSH margin	24
4.1.3.2 Pump efficiency	25
4.1.3.3 BEP.....	28
4.1.4. Data conclusion.....	28
4.2 Method selection	29
4.2.1 LEM stochastic methods	29
4.2.1.1 Aggregate reliability functions	29
4.2.1.2 Conditional probability functions	30
4.2.2 LEM Statistical methods.....	31
4.2.2.1 Trend evaluation.....	31
4.2.2.2 Autoregressive methods	31
4.2.2.3 Proportional hazards modelling	31
4.3 LEM method selection.....	32
4.4 Bayesian Network	33
4.4.1 Probability theory	34
4.4.2 Example calculations.....	36
4.5 Conclusion	38
5. Method for the model	39
5.1 Multi-layer structure.....	39
5.2 Bayesian Network	39
5.2.1 Component nodes.....	42
5.2.2 Subsystem level nodes.....	43
5.2.3 Overall system level node	45
5.3 Conditional probability tables	46
5.4 Forward inference	46
5.5 BN verification	46
5.5.1 Verification of nodes and relations.....	47
5.5.2 Verification: Sensitivity analysis.....	47
5.5.3. Verification conclusion.....	51
5.6 Conclusion	52
6. Model performance	53
6.1 Reliability.....	53
6.2 Key Performance Indicators	54
6.2.1 Maintenance downtime index	54
6.2.2 Cost key performance indicators	54
6.2.2.1 Corrective maintenance costs.....	55
6.2.2.2 Preventive maintenance costs.....	55
6.2.2.3 Predictive maintenance costs	56
6.2.2.4 Cost KPIs	57
6.3 Overview of the model.....	57
6.3.1 BN.....	57

6.3.2 MTBF	58
6.3.3 Cost analysis	58
6.3.4 Flow chart of the model	59
6.3.4.1 Preparations before use of the model	62
6.3.4.2 Verification of the dataset	62
6.3.4.3 Assumptions	62
6.4 Conclusion	63
7. Results	64
7.1 Without model	64
7.2 With model	65
7.3 Comparison	68
7.4 Results KPIs	69
7.4.1 Results MDI KPI	69
7.4.2 Results cost KPIs	70
7.5 Conclusion	71
8. Conclusion	72
9. Discussion	74
10. Recommendations	77
Literature sources	79
Appendix A: Scientific paper	84
Appendix B: CIP	91
Appendix C: Vapour pressure and density	92
Appendix D: Pump curves	95
Appendix E: Method selection	97
Appendix F: Overall Bayesian Network	104
Appendix G: Conditional probability tables	105
Appendix H: Independence checks	125
Appendix I: Mean Time Between Failure	126
Appendix J: Synthetic production planning	127

1. Introduction

Fresh, sparkling water with or without a flavour, who doesn't drink it? About 1,458,000,000 litres of carbonated soft drinks are consumed in the Netherlands annually. Converted, that is 85L per person per year in the Netherlands, which is equivalent to 57 bottles of 1.5L with carbonated soft drinks [1].

The soft drink industry is part of the much larger beverage industry. In this industry, a distinction is made between alcoholic and non-alcoholic beverage [2], as shown in Figure 1. Within alcoholic beverages, a distinction can be made on the basis of the produced beverage: fruit/juice or grain-based. For the non-alcoholic beverage, a distinction can be made between non-carbonated, in which case no carbon dioxide is added, and carbonated drinks, to which carbon dioxide is added [2].

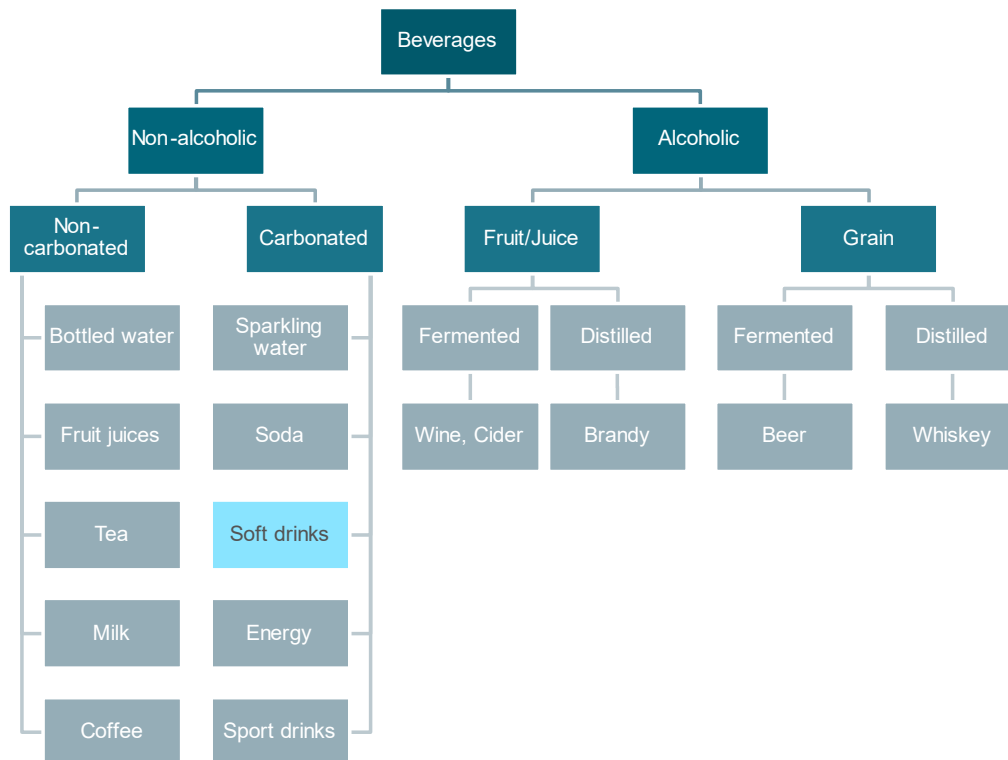


Figure 1: Overview beverage industry, with highlight on the carbonated soft drinks [2]

Within this research, the focus is on carbonated soft drinks. In this research, the carbonated soft drinks are defined as: "water-based flavoured drinks usually with added carbon dioxide and with nutritive, non-nutritive, and/or intense sweeteners with other permitted food additives" [3].

1.1 History of soft drinks

The origins of carbonated soft drinks can be traced back to 1767, when Joseph Priestly experimented with impregnating water with fixed air [4]. However, it became well known among the wider public after 1783 when Jacob Scheppe started producing carbonated waters in glass bottles under the brand name "Schweppes" [5]. Over time, more brands were added, providing a wide variety of flavours and colours. Packaging also became increasingly diverse. Where it used to be sold only in glass bottles, soft drinks also became available in cans since 1948 [6]. Later, in 1978, soft drinks were first sold in PET bottles [7]. In the early years, the process of producing soft drinks was completely manual, see Figure 2, even the glass bottles were blown piece by piece. By the time of the second industrial revolution around 1890, automated machines emerged [7].

LEMONADE AND SODA-WATER MACHINE.

NATHANIEL GEORGE WILCOCKS,
ENGINEER AND
Soda Water Machine Manufacturer.

There are over 1000 Manufacturers in the United Kingdom using these Machines for making

Soda, Potass,
Seltzer, Carrara,
Lithia, Magnesia,
and
Tonic Waters, etc. ;



Lemonade, Ginger
Beer, Ginger Ale,
Orangeade,
Gingerade, Nectar
and
Champagne Cider.

Directions and Recipes for the Manufacture of all Aërated Drinks given to purchasers of Machines, etc.

MACHINES FROM £30 TO £200.

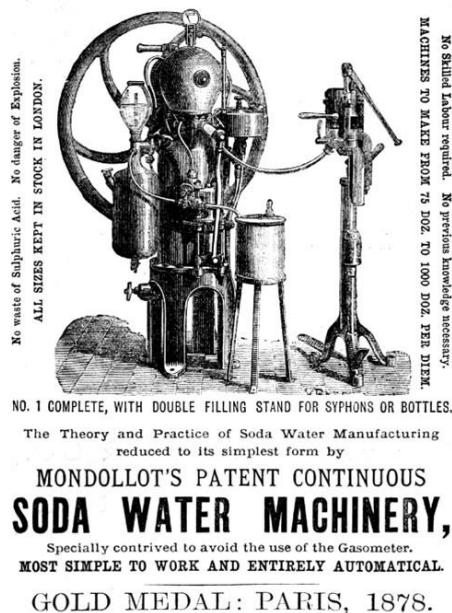


Figure 2: Early soft drink production manual equipment [8]

Nowadays, all soft drinks are produced in soft drink factories and it is an automated process.

There are a number of large soft drink producers in the Netherlands that produce for both national and international markets. This industry is responsible for every imaginable type, flavour and colour of soft drink on the market. This research includes a case study performed in collaboration with a soft drink producer in the Netherlands.

Soft drinks consist mostly of water to which sugar, flavourings and colourings are added. Therefore, inside soft drink factories several processes must take place to turn the raw ingredients into soft drinks. These processes are defined by Koss [9] as different production lines with the following classification, see also Figure 3:

- **The container line:** this is where the bottles and cans are unwrapped or shaped. Afterwards, these are cleaned/washed and prepared to be filled.
- **The product line:** here the ingredients are compounded, mixed and subjected to other processes to comply with the laws and regulations of the food and commodity authorities. Other processes can include heating to kill bacteria or, for example, adding carbon dioxide for carbonated soft drinks. Furthermore, the product is prepared to be put into bottles and cans.
- **The filling/closing line:** this is where the container line and product line meet. The bottles and cans are filled and then sealed with a lid or cap.
- **The container treatment line:** here, the bottles and cans are subjected to a heating process or pasteurization, if necessary. Afterwards, the bottles and cans are labelled and checked for the last time.
- **The product packaging line:** here the individual products are assembled in, for example, crates or as trays.
- **The storage preparation line:** here the bulk packaging is loaded onto pallets and sealed. These are then stored until the product is transported to the customer.



Figure 3: Different lines in soft drink factory defined by Koss [9]

After each production batch, the lines have to be cleaned before a new batch can be started. This research focuses on the product line of a soft drink producer in the Netherlands. The soft drink producer has several lines through which the different types of soft drinks are produced, in this research the focus is on one of those lines. All other lines as defined by Koss [9] will be disregarded.

1.2 Production line

The production line within the soft drink factory is responsible for bringing together and mixing all the ingredients to obtain the soft drink [9]. This part of the line is also responsible for carbonisation, adding the carbon dioxide to the soft drink, heating and cooling. Once the soft drink is ready on this line it only has to go to the fill and packaging lines. Figure 4 is an example of the product line.

In principle, the line is only used between Monday morning and Friday afternoon, in case there are run-offs, work continues over the weekend to make up those hours. It works with a batch process, meaning that one batch of a particular soft drink is produced each time, then it is cleaned and then a new batch can be started.

To do this, the line consists of all kinds of different components. For instance, there are tanks where all the ingredients are stored, there are pumps to pump the ingredients through pipes to the next tank and valves for dosing and sending them to the right destination.

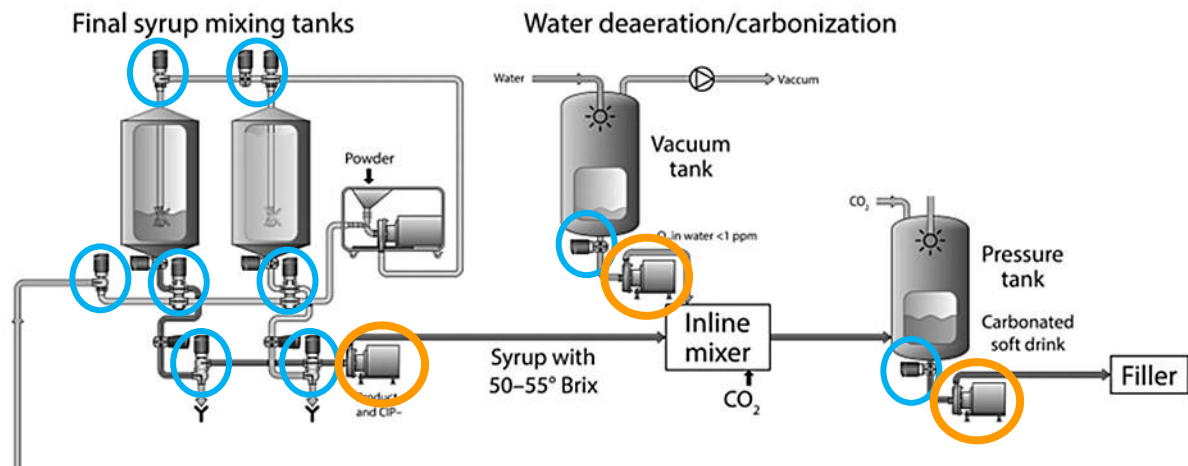


Figure 4: Production line, the valves are indicated with the blue circles and de pumps with orange circles [2]

In this research, a case study will be done on the valves and pumps in a part of one production line at a Dutch soft drink manufacturer. For this purpose, 25 valves and 6 pumps are included in this research. Chapter 2 looks at the operation of the valves and pumps in more detail.

1.3 Problem definition

The problem that is considered in this research can be divided into a practical problem of the soft drink company and a scientific problem.

A well-known problem within the soft drink manufacturing industry are unplanned downtimes [10]. This is the unexpected stoppage of a production line due to the failure of components within that line. Maintenance must then be carried out immediately such that the line can be restarted afterwards. As a result of the unplanned downtimes, a lot of production time is lost and batches have to be rejected because quality cannot be guaranteed. All in all, unplanned downtimes cause unnecessary production losses and lower line reliability [10].

The process industry, and therefore also the soft drink industry, is known to be a slow adopter when it comes to adopting new technologies and strategies [11]. This is also reflected when looking at current maintenance strategies. In this industry, there is still a lot of either corrective maintenance or preventive maintenance. Corrective maintenance is defined as carrying out maintenance after equipment has failed. Whereas preventive maintenance is defined as replacing parts on a fixed schedule or carrying out maintenance without any further reason.

However, these maintenance strategies still result in many unplanned downtimes. This is also the case at the soft drink product line for this research. The company suffers from unplanned line downtimes due to the failure of mainly valves and pumps. It uses mainly corrective maintenance, where maintenance is carried out if a component is broken. Furthermore, the company has been collecting data for several years that can say something about the condition of the equipment in the line. However, so far no step has been taken to integrate the data to form a predictive maintenance strategy for the valves and pumps within a production line. The aim of the predictive maintenance strategy is to reduce the number of unplanned downtimes and thus improve the reliability of the production line.

Within the literature, there is a considerable amount of information on the ways that maintenance strategies can be adapted to reduce unplanned downtimes. In many cases, data is collected to form an idea of the condition of the

equipment within a system. Models are then used to determine the expected remaining useful life. This can indicate when a component needs maintenance so that it does not reach the end of its useful life and fail. From research by Jimenez and Vingerhoeds [12], this emerges as characteristics of a predictive maintenance strategy. Predictive maintenance is already used in many other industries, such as aviation [13] and marine [14], to prevent unplanned downtimes.

For the beverage industry, there are some studies by Tsarouhas [15] [16] in which equipment in the production line is monitored and data is collected to determine whether the line needs maintenance or not. This is often done at system level rather than component level. It can then be determined when which line needs maintenance, but not so much which component within the line.

In contrast to previous research looking at how condition monitoring of an entire system is used to set up a predictive maintenance strategy, this research will focus on the component level.

Within the literature, few examples can be found for the beverage industry on how the condition monitoring data of components can be used. Whereas in the aviation and oil industries there are numerous examples of models where condition monitoring data is used as a tool for, for instance, a maintenance plan, this is still a grey area for the soft drink industry. This research will look at how the beverage industry can use condition monitoring data at component level for a predictive maintenance strategy. It will also look at what benefits the predictive maintenance strategy brings to the soft drink industry.

1.4 Research objectives

In line with the problem statement for this research, the following research objectives can be defined.

It is necessary to investigate what information condition monitoring data provide at the component level. It is then necessary to determine how condition monitoring data can be related to component failure probability. It is also necessary to examine how the condition of a component affects the failure probability of the entire production line. The failure probability should also be related to the probability of unplanned downtime. All the above should be brought together in an accessible model.

Part of the model should be able to determine, using a method, the failure probability of the whole system from given input data; this is forward reasoning. To determine which method can best do this, a survey of existing methods is first carried out and which method best suits this research. In addition, the method must also be capable of being updated as new data becomes available.

The aim of the research is to arrive at an accessible method that can be used to create a model. The model will therefore include the method described earlier and some additions to be able to arrive at a working model. A final model should be able to determine a failure probability for the entire line based on input data, condition monitoring data of the system and defined relationships between the data. This can then be used to calculate the probability of unplanned downtime. The model must be able to perform these actions on both existing data and a synthetic dataset. Using the model's output, it should become possible to determine whether a component needs maintenance before the next process is started. This will allow then to perform predictive maintenance. Indeed, synthetic data can be run through the model to determine whether a component can perform its function long enough until the next scheduled stopping point.

The model should further provide insight into how line reliability changes when timely maintenance is performed compared to the situation where no model would be used.

1.5 Research questions

In this research, a main research question has been formulated to arrive at a solution to the previously stated problem. To formulate an answer to the main research question, some sub-research questions have been drawn up and will be answered throughout the research. The answer to the main question is further supported by the outputs of the model.

Main question:

How can a predictive maintenance strategy contribute to improving the reliability of a soft drink production line?

Sub-research questions:

1. What are the main components in the soft drink production line?
2. What is the best maintenance strategy for this problem compared to the current maintenance strategy?
3. Based on the available data and knowledge, what kind of method is most appropriate to develop a model with for this problem?
4. What are the steps to develop and verify the method of the model?
5. In what way is the implementation of the model contributing to the reliability of the production line?

1.6 Methodology

To arrive at a predictive maintenance strategy that can improve production line reliability, a literature review must first be conducted. It will consist of a chapter on the process and function of valves and pumps. Then, different maintenance strategies will be discussed. Here, the current maintenance strategy and its shortcomings should be discussed. Next, a predictive maintenance strategy will be examined and how this type of maintenance strategy can complement the shortcomings of the current maintenance strategy. Current applications of predictive maintenance are also discussed with a focus on the beverage industry. It will then be considered how a predictive maintenance strategy can be modelled. The model will consist of several parts. This includes a part that can determine the probability of failure of the production line based on a method. For the method, the first step is to see what methods are available to determine this. These methods are then subjected to some previously established criteria specific to this research. This allows the best method for this research to be determined.

Once the method is chosen, it will be determined which factors contribute to the likelihood of unplanned downtime in which way. Theoretical approaches will be used to link certain variables. The relationships between the variables and the probability of downtime will be incorporated into the method, creating a method that reflects the components in the production line.

Then the method will be subjected to some verifications.

The method is part of the model. In this model, the method is used to determine the probability of failure of the production line. In addition, the model determines some other factors that, finally, can be used to indicate whether or not maintenance is required before starting new production.

To determine the effect of the model (and thus the predictive maintenance strategy) on line reliability, a synthetic dataset is created. This makes it possible to compare two situations, namely the situation where the model is not used (corrective maintenance) and the situation where the model is used (predictive maintenance). It is done on the basis of KPIs. By comparing the values of the KPIs of the two situations, the impact of a predictive maintenance strategy on the reliability of the production line can be determined.

In short, data analysis, theoretical relationships and a model containing a method for determining probability of failure are used to determine the impact of the predictive maintenance strategy on the reliability of the soft drink production line.

1.7 Thesis outline

The structure of the thesis is based on the sub-research questions. [Chapter 2](#) looks at the processes that take place in a soft drink factory. It also considers the valves and pumps used and the ways in which these components can fail. The different types of downtime and the current maintenance strategy are briefly discussed.

[Chapter 3](#) focuses on the second sub-research question and looks at the different maintenance strategies. More explanation is given of the different types of methods that can serve as a basis for modelling a predictive maintenance strategy. Finally, the criteria to be met by a method for this research are briefly touched upon.

[Chapter 4](#) starts with a data analysis to determine what data are available and how they can be related to the failure probability. This is done for the production line in general and specifically for the valves and pumps. Methods are then discussed. The different methods are subjected to a number of criteria, ultimately leaving one method that best fits this research. A brief general explanation of the chosen method is then given.

[Chapter 5](#) deals with the fourth sub-research question. Here, the model created for this research is explained step by step. Next, verification is carried out.

[Chapter 6](#) discusses the performance of the model and thus partially addresses the fifth sub-research question. Here, it looks at the KPIs that can reflect the reliability of the production line. Next, the flowchart of the full model for this research is discussed. It then discusses the two situations, with and without the model, and how to calculate the KPIs for these.

[Chapter 7](#) presents the results of the model as described in the flowchart for both situations. It also looks at the fifth sub-research question here to answer it based on the results from the model. This also looks at the outcomes of the KPIs for both situations. Finally, the results are briefly compared.

[Chapter 8](#) brings forward the conclusion of this research. [Chapter 9](#) presents the discussion and [Chapter 10](#) closes this thesis with the recommendations for a follow-up research.

2. Production process and components

This chapter focuses on the sub-research question: *What are the main components in the soft drink production line?* To answer this question, first, the production process will be looked at in more detail. There will be also discussed what happens in the line when producing. It also looks at the CIP (Clean-In-Place) process. After it is clear what processes take place, the valves and pumps in the line are looked at. The operation of these components and known failures are identified. This is followed up by looking at what downtimes of the production line are and how valves and pumps affect them. The chapter concludes with a brief review of the current maintenance strategy and desires for a different strategy.

2.1 Production process

As stated in [Chapter 1.2](#), only the production line is considered. Defined by Koss [9], this is the line where all the ingredients are processed and mixed to arrive at the final product.

Before the products can be mixed, these must first be collected. First, the main ingredient of soft drinks; water. Soft drinks consist of at least 85% water [17]. Most soft drink factories pump the water themselves or otherwise extract it from the earth's surface or other water sources such as rivers or lakes [17]. The water must undergo some treatments, including filtering, venting and pH processing, so that it complies with Food and Commodity Authority Regulations. The water is then stored in tanks ready to use.

Another main ingredient in soft drinks is sugar. There are soft drink factories that process sugar beet into sugar themselves. Other factories use ready-made sugar. Sugar is dissolved in water to make a syrup. In zero and light products, sweeteners are used instead of sugar. These too are dissolved in water to make a syrup [18]. After dissolving the sugar or sweetener in water, flavouring and colouring agents must also be added to make a syrup. This can be done through metering valves connected to tanks containing fruit juice extracts and colourings. Alternatively, there is a premix of the flavourings and colourings that has to be dissolved with sugar in water [7]. All this has to be mixed well in so-called mixing tanks. After mixing in the tank, the syrup is pasteurised. After pasteurisation, the syrup must be de-aerated. This is because air may have entered the syrup during mixing. De-aerating can be done by letting the syrup rest for a while in a tank or by using a vacuum de-aerator [19].

After de-aerating the syrup, the next step is to mix the right amount of syrup with water and carbon dioxide. In many cases, the water is first impregnated with the carbon dioxide gas under high pressure and low temperature. The lower the temperature of the water, the better the carbon dioxide can dissolve in it. The process of adding the water with the carbon dioxide gas is called carbonisation. The carbonated water is then mixed with the syrup to become a carbonated soft drink. In many cases, it is then pasteurised again to avoid unwanted processes [17]. After this, the soft drink is ready to go to the filling line.

2.1.1 Production line case study

The line analysed for this research starts at the point the syrup is mixed. After this, therefore, the steps of heating, cooling, resting and carbonisation still have to take place.

Below, see [Figure 5](#), is a schematic representation of the line that is central to this research.

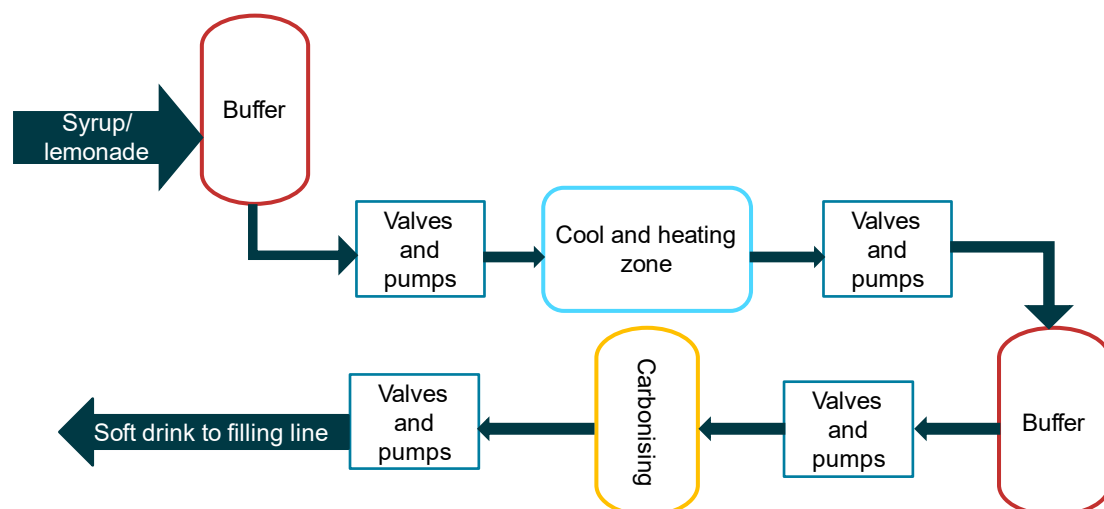


Figure 5: Schematic overview of the production line for this research

The follow-up steps such as filling the bottles or cans are beyond the scope of this research.

Besides the production process, there is another process that takes place within the soft drink factory. This is the cleaning process of all equipment within the process, i.e. the valves, tanks, pipes and pumps. It is called the Clean-In-Place (CIP) system. This will be briefly addressed as it affects the lifetime of the components. It also provides background information on what to do when the line needs cleaning after maintenance has been carried out.

2.2 Clean-In-Place

There are many different ways in which CIP can be carried out.

The purpose of the CIP system is to clean the production line and thereby also remove bacteria [7] [20]. There are many different ways in which CIP can be carried out, but three ways of cleaning can always be distinguished [21].

1. **Mechanical:** using impact and turbulence to remove residues, often this is done using only water pumped through the line at high speed to create turbulence which is then used to flush away the residues left behind.
2. **Chemical:** using chemical actions to clean up the remaining residues, smaller residues are flushed out of the line.
3. **Sterilisation:** to kill the micro-organisms.

Cleaning the line can be done in many different ways, but generally there are a few steps that are always performed. First, everything will be flushed with water and then cleaned with different cleaning agents. After using cleaning agents, rinsing is always done with water. [Appendix B](#) provides an overview of all the different steps, cleaning agents, temperature and time frame.

Which type of CIP is used depends on the type of product that is produced.

Often, the water and cleaning agents for CIP are stored in tanks in another part of the factory. Through pipes, pumps and valves, it is pumped into the right part of the line that has to be cleaned.

There are some factors to consider during the CIP process. For example, it is important that the flow rate is 1.5 m/s [20]. This is because then the cleaning agents can do their job best. If the flow rate is higher than this it can give "water hammer" causing the equipment within the system to break down [20]. Water hammer can occur in pipelines that pump fluid and contain valves to direct the fluid in the right direction. If a pump stops pumping or a valve suddenly closes when it should have been open, the fluid comes to a sudden stop. The momentum of the liquid then causes a pressure wave that is reflected back into the pipe. As a result of the resulting interplay of forces, the valve or pump may fail and have to be replaced. This creates unplanned downtime [20]. Furthermore, the total time of cleaning is difficult to determine because it depends on many factors, which must be taken into account when making the production planning.

CIP makes extensive use of water pumped through the line at high speed. In addition, various chemical cleaning agents, often at high temperature, are pumped through the line and components. All this affects the service life of valves and pumps within the line [21]. Therefore, for this research, CIP, or rather the influence of CIP on service life, must be included.

2.3 Valves and Pumps

From the description of the production process and CIP, it can be seen that different types of equipment are needed to run the entire process. Within this research, the main focus will be on valves and pumps within the production process.

2.3.1 Valves; types and working principles

Many valves are used in most soft drink factories, as is the case in the factory in this research. The valves ensure that the products go to the right tanks or pipes, so that it can be further processed there into the final product. There are also metering valves that ensure that the right amount flows into or out of a tank. Other valves are usually located in the pipes or in a valve matrix [22]. A common type of valve is the valve with double seat. In this, two chambers separated by a stem can separate liquids. This type of valve can also be used to direct fluids to another line [23]. See [Figure 6](#) [23] and [Figure 7](#) [24] for an example of the double seat valve.

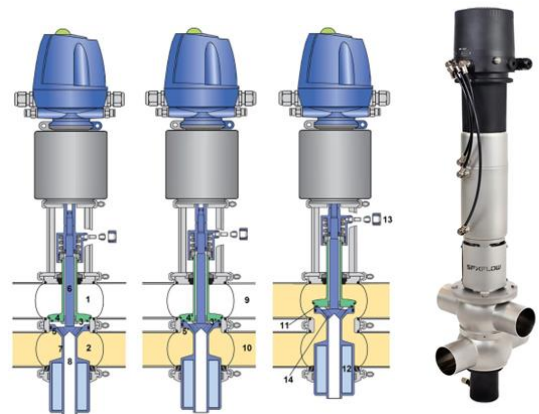


Figure 6: Double seat valve [23]

Figure 7: Double seat valve [24]

Another commonly used type of valve is the butterfly valve. This is a type of valve that can only be open or closed. It works with a disc in the centre that can be turned 90 degrees so that the valve is open. To close, the disc is turned 90 degrees again. This causes the disc to fall back into the rubber which ensures that nothing can pass [25], see [Figure 8](#) [26] and [Figure 9](#) [27].

All valves in the line are air-operated [22].



Figure 8: Butterfly valve [26]



Figure 9: Butterfly valve [27]

The operator controls the Programmable Logic Controllers (PLCs) connected to the valves. The PLCs transmit the command and then send a signal when the valves are in the desired position. Also, the PLC logs if anything changes in the position of the valve [28]. Furthermore, the valves can be classified by diameter in addition to the type. In this research, twelve double seat valves and thirteen butterfly valves will be included. This gives a total of 25 valves.

2.3.2 Valve failures

The most common problem with valves is leakage. This is usually caused by wear of either the stem or the seat [29]. Also, the O-rings between the different chambers can have wear which prevents the valve from closing properly. Possible causes for this wear include frequent use, cleaning agents, water hammer and influence of process parameters such as temperature or pressure [30]. It is noteworthy here that both the production process and CIP influence valve failures.

Another type of failure is when a valve has been stationary in a certain position for too long, causing the valve to malfunction [31].

Failure due to water hammer suddenly involves a lot of water being pumped through the pipes, often because something else malfunctions at that point, putting a lot of force on the valves. This can knock the O-rings out of the valves or the stem can no longer hold and fails [31].

If a valve fails then it causes downtime of the entire production line. Only when maintenance has been carried out or the valve is replaced can the process be restarted, after cleaning.

2.3.3 Pumps; working principle

Soft drink factories often use centrifugal pumps.

The pumps are used to transport the liquids through the pipes [32]. In many cases, centrifugal pumps are used that operate with impellers to provide energy to the liquid (syrup, water or cleaning fluid) to be pumped, see [Figure 10](#) [33] and [Figure 11](#) [34]. The inlet (suction side) is where the liquid is sent into the pump. It then reaches the impeller which processes it to the outlet (discharge) of the pump. Centrifugal pumps can also fail. There will be six centrifugal pumps involved in this research.

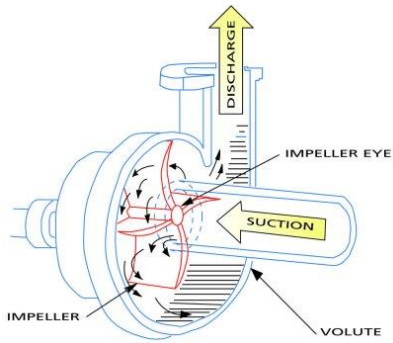


Figure 10: Centrifugal pump schematic [33]



Figure 11: Centrifugal pump [34]

2.3.4 Pump failures

Centrifugal pump failures can be classified into three categories [29]:

1. **Hydraulic failures** → These include all failures caused by pressure changes. This leads to cavitation and high thrust on the impellers.
2. **Mechanical failures** → The failures that occur to parts of the pump. These include fatigue, failing bearings, leaking seals and degradation of lubricants due to high temperatures.
3. **Corrosion and erosion** → Corrosion occurs due to a chemical attack on the material which changes as a result. Erosion is defined as the physical wear and tear of equipment [35].

It is known that if the pumps build up too much pressure and the fluid has nowhere to go that then the pumps clamp and break down. If a pump fails, it causes downtime for the entire production line. Only when maintenance has been carried out or the pump replaced can the process be restarted, after cleaning.

2.4 Different types of downtimes

One of the biggest challenges currently facing the soft drink industry is to reduce downtimes. Downtime (DT) is defined as: “period of time in which production machinery is not allowed to perform its output because it is not working” [36].

Two categories can be distinguished within downtimes [36], see also Figure 12 (based on [36]):

- **Planned downtimes:** these are organized stops of the process to perform maintenance, cleaning or preparing the process for another product, for example.
- **Unplanned downtimes:** these are the moments when the process suddenly stops due to equipment failure.

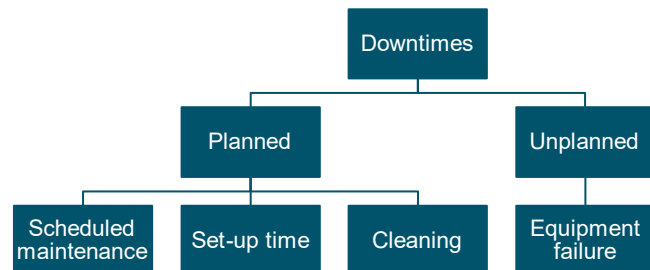


Figure 12: Overview of types of downtimes [36]

Within the soft drink industry, the biggest problem is unplanned downtimes, due to equipment failure in the production line [37].

There are known examples where unplanned downtimes are 37% of operational time. This means that only 63% of the operating time can be used to produce products [38]. On average, beverage manufacturers lose about 25 hours of total operating time per month to unplanned downtimes [39].

An annoying consequence of unplanned downtimes within the beverage industry is that it causes wastage of products. This is due to the very strict laws and regulations of the Food and Consumer Product Safety Authority. Due to unplanned downtimes, quality can no longer be guaranteed and the batch has to be rejected [40].

Unplanned downtimes due to equipment failure within the beverage industry often do not happen all at once. In this case, it is often a gradual degradation of the equipment, which is reflected in the output of the products, think of varying product quality [41]. By using condition monitoring tools, the condition of equipment can be monitored. Chapter 3 will take a closer look at different maintenance strategies that use condition monitoring techniques.

2.4.1 Condition monitoring of valves

If valves can no longer close properly, this causes major problems, for example, fluid can enter pipes where it is not wanted. If there is a leak, the result may be that fluid leaks out of the system. One of the possible consequences could then be that the composition of the product is affected and thus food safety. Also, fewer bottles or cans may then be filled than previously calculated. Once failure occurs within the valves during producing then the entire line must be shut down, usually this is an unplanned downtime, to allow maintenance to be carried out. After maintenance, the line has to be cleaned before new production can be started.

For valves, there are several condition monitoring techniques that can be used. Some of those are: valve torque, travel time (time between valve opening and closing) (in Dutch called: looptijd) [42], position monitoring, acoustic monitoring and valve flow coefficient [29].

2.4.2 Condition monitoring of pumps

The consequences are significant if the pumps come to a stop. Most of the time it causes an unplanned downtime. Fluid can then no longer be transported through the rest of the line. There are a few scenarios that can then occur; the liquid comes to a standstill in the line and has a chance to adhere to the equipment there with all the consequences. Another possibility is that all liquid is let out of the system through the valves, however, there is always a chance that something will be left behind, but cleaning is not an option because those liquids cannot be pumped either. Either way, in both cases it creates an undesirable situation that should be avoided if at all possible.

There are also condition monitoring techniques for centrifugal pumps. Some of them are: vibration monitoring, lubricant sampling, head-flow measurement, shut-off head method, temperature difference and thermography [29] [32].

2.4.3 Case study at soft drink factory

In this research, data available on the valves and pumps in a single production line will be used. In the case of the valves, this is historical data on the looptijd, the time between the command open or close and whether the valve is in the correct position. In addition, it is known where in the line a particular valve is located, which processes have occurred and when valves have failed as well as what position the valve was in during the failure. The historical data is supplemented with real-time data on looptijd, valve position, type of process (production or CIP) and moments of failure. There is also a lot of knowledge on the factory floor about how to interpret the data.

From the previously mentioned condition monitoring techniques for valves, looptijd, position monitoring and process monitoring are used.

For the pumps, head-flow measurement, pressure differences and pump performance measures will have to be considered. Because it is clear when a production process was going on and when cleaning was done, it is possible to see how this affects unplanned downtimes.

2.5 Current maintenance strategy and requirements

Currently, the plant uses a combination of corrective and preventive maintenance. In many cases in the production line, the "Fix it, when it is broken" principle, or corrective maintenance [43], prevails and maintenance is carried out the moment a component from the line fails. This component is then repaired or replaced. Fixed maintenance moments are also used, during which maintenance is carried out on the lines scheduled at that time. This falls under the preventive maintenance principle [43].

At the moment, only the real-time data coming in is looked at; if this shows deviating values, the operator can intervene. In many cases, this first involves resetting the components not working properly in the line and controlling them again. If this does not work, the entire line will have to be shut down to perform maintenance and there is thus an unplanned downtime. After there has been an unplanned downtime, the entire line has to be cleaned before a new batch can be started. This comes at the cost of production time over that line. So far, no other way of maintenance has been tried to reduce the number of unplanned downtimes. As a result, the line is down several times a month.

However, data has been collected and stored for several years with which nothing is currently being done. This data could play a major role in improving the maintenance strategy to reduce unplanned downtimes. The company wants to use the data to determine the condition of individual components in the line. In addition, the company also wants to look at how, based on the data and expert knowledge, a forecast can be made about the expected time in which a component will fail. Based on those forecasts, it can then determine when a component needs maintenance so that the amount of unplanned downtimes will reduce. To improve the maintenance strategy from reactive to proactive, it is therefore necessary to look at ways of creating the desired output based on data and expert knowledge.

2.6 Conclusion

This chapter focuses on the sub-research question: *What are the most important parts of the soft drink production line?* Clearly, the valves and pumps are the most important parts of the production line. In the production line used in this study, there are 25 valves and six pumps. It was then investigated how these components function and in what ways these components can fail. It was also recorded how the valves and pumps cause unplanned downtime and what the consequences are. In addition, the techniques that can be used to determine the condition of the components are discussed.

The case study is carried out at a soft drink factory in the Netherlands. This factory currently operates with mainly a corrective maintenance strategy. There are a few moments during a calendar year when the line is scheduled to stop and maintenance is carried out on the components currently on a list, often based on data from the manufacturer on how often a component needs maintenance. With these modes of maintenance, many unplanned downtimes are experienced. Steps have been taken in recent years by collecting data on the condition of valves and pumps in the line. However, this is not currently being used to reduce the number of unplanned downtimes. It can be gathered from the literature that one of the ways to reduce unplanned downtimes is to work with a proactive maintenance strategy instead of a reactive maintenance strategy (reactive includes the corrective maintenance strategy).

The case study will therefore need to look further into the ways of moving from the reactive to a proactive maintenance strategy and how this contributes to reducing the number of unplanned downtimes. [Chapter 3](#) will elaborate on the different maintenance strategies and which proactive maintenance strategy is the best fit for this research.

3. Maintenance strategies

The central sub-research question in this chapter is: *What is the best maintenance strategy for this problem compared to the current maintenance strategy?*

To answer this research question, different maintenance strategies will first be discussed. Then it will be examined which of the maintenance strategies best suits the requirements that have emerged in [Chapter 2.5](#).

The best-fit maintenance strategy, for the problem in this research, will then be discussed in more detail to clarify what the functionalities and essential elements are.

3.1 Maintenance standard EN 13306

When looking at the term maintenance, there are many different forms and definitions. This research will use the definition of maintenance as stated in EN 13306 [43]; "combination of all technical, administrative and managerial actions during the life cycle of an item intended to retain it in, or restore it to, a state in which it can perform the required function".

Within EN 13306 [43], the first distinction is whether maintenance is performed before or after the failure of a component. This is also called reactive or proactive by Shukla et al [44]. Reactive maintenance takes place after the failure of a component. Proactive maintenance, on the other hand, takes place before the component has a chance to fail. The overall scheme can be seen in [Figure 13](#) [43].

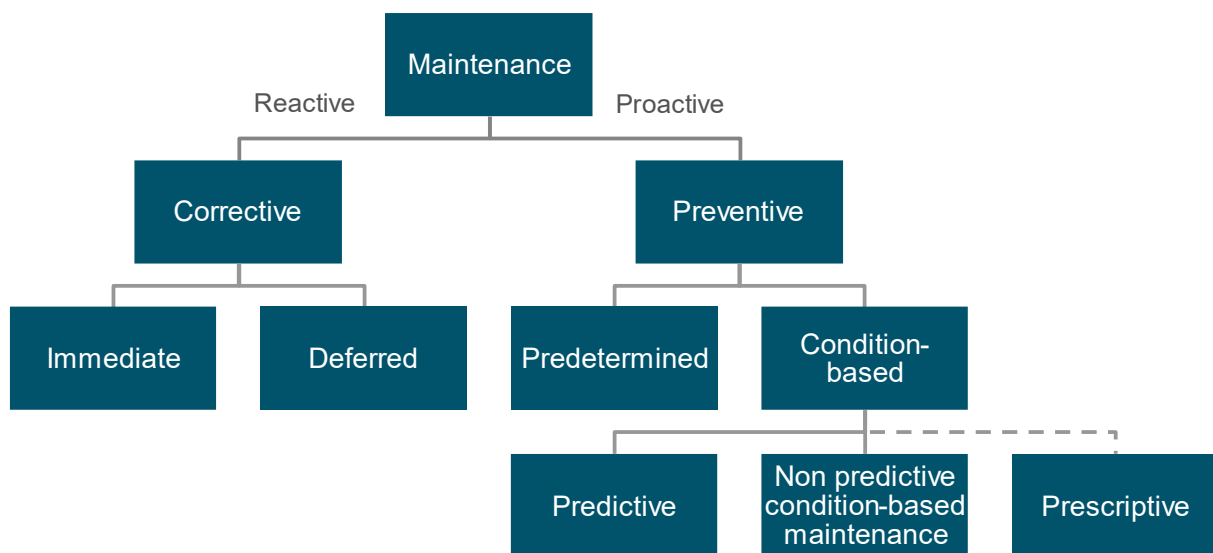


Figure 13: Overview of different maintenance strategies based on EN13306 [43]

3.1.1 Corrective maintenance

Corrective maintenance falls under the category of reactive maintenance. It is also known as: "run-to-failure" [44] or "break-down maintenance" [45]. Corrective maintenance is characterised by the system continuing to function throughout its lifetime and only being replaced after failure. The definition of corrective maintenance according to EN 13306 is as follows: "maintenance carried out after fault recognition and intended to restore an item into a state in which it can perform a required function" [43].

According to Merkt [46] systems with corrective maintenance are set up to deal with pre-known failures and damages. However, as time goes by, new failures and associated patterns emerge with the use of the system.

Corrective maintenance can be further divided, according to EN 13306, into immediate corrective and deferred corrective maintenance. As the name suggests, immediate corrective maintenance involves immediate maintenance of the failing component to avoid unpleasant consequences [47].

With deferred corrective maintenance, on the other hand, maintenance is carried out later because other things have higher priority or parts are out of stock [48].

Inside the proactive maintenance category, there is the preventive maintenance category according to EN 13306. With this strategy, maintenance is carried out before a component has a chance to fail.

3.1.2 Preventive maintenance

When maintenance is carried out before a failure has occurred, it is called preventive maintenance. According to EN 13306, preventive maintenance has the following definition: "maintenance carried out intended to assess and/or to mitigate degradation and reduce the probability of failure of an item" [43].

According to Merkt [46] the biggest challenge of preventive maintenance is that it does nothing with the past. This means that no data from the past is kept and analysed regarding, for example, abnormal behaviour or maintenance actions that were carried out to prevent failure. Preventive maintenance defines a set of actions to be performed before the system or component even has a chance to fail.

In line with EN 13306, this category is further divided into predetermined maintenance and condition-based maintenance.

3.1.3 Predetermined maintenance

Predetermined maintenance is defined according to EN 13306 as follows: "preventive maintenance carried out in accordance with established intervals of time or number of units of use but without previous condition investigation". These intervals are determined with the knowledge that exists about the possible failure options and after how much time or cycles this should/could happen. Equipment degradation is not monitored here. According to Kothamasu et al [49], the underlying assumption is that each component is assumed to always go off the same operational curve. This can then be used to determine when maintenance should take place. Merkt [46] states that predetermined maintenance includes time-based maintenance. In which maintenance is carried out based on a pre-conceived time schedule. With this, the parts are used until it may fail according to the manufacturer, once that moment is reached the part is replaced without actually being at the end of its service life. Again, there is no monitoring of the actual degradation but is based on pre-supplied specifications.

Kothamasu et al [49] argue that time-based maintenance is also known as constant time interval. In addition to this strategy, Kothamasu et al [49] suggest that there are two other strategies that are common in preventive maintenance. These are age-based maintenance and imperfect maintenance. Age-based maintenance, as the name suggests, is based on the age of a component. If the component fails at time t , then the next maintenance moment can be scheduled after t time has elapsed again. According to Kothamasu et al [49], compared to the constant time interval strategy, this strategy reduces the number of maintenance moments.

Imperfect maintenance, where the other strategies assume that a component works as new again and thus starts a life cycle anew, here it is assumed that a component no longer works as new. It is brought back to a working state, but not starting a new life cycle. The imperfect maintenance strategy takes the uncertainty of a component's current state into account when planning maintenance moments.

3.1.4 Condition-based maintenance

Condition-based maintenance is defined according to EN 13306 as:

Preventive maintenance that includes assessment of physical condition, analysis and possible consequent maintenance actions. Condition assessment can be done by operator observation, and/or inspection, and/or testing, and/or condition monitoring of system parameters, etc. according to a schedule, on demand or continuously [43].

Merkt [46] explains condition-based maintenance with that it is a strategy of monitoring equipment and intervening when there is evidence of degradation or other deviations from normal behaviour of the system or component. In addition, Key performance indicators (KPIs) or other health indicators can also be calculated and analysed to detect patterns or disturbances.

According to EN 13306 [43], condition-based maintenance is further split into predictive maintenance and non-predictive condition-based maintenance.

3.1.5 Predictive maintenance

Predictive maintenance (PdM) is used in EN 13306 with the following definition: "condition-based maintenance carried out following a forecast derived from repeated analysis or known characteristics and evaluation of the significant parameters of the degradation of the item".

Merkt [46] adds that predictive maintenance uses both historical and real-time data. The data is processed into prognostic models using machine learning and other methods, which can be used to make accurate predictions about the future status of the equipment. Predictive models assume that at a certain point in the equipment's life-time the failure rate will increase. This can then be used to predict when maintenance is required to prevent failure.

3.1.6 Non predictive condition-based maintenance

Non-predictive condition-based maintenance does not use prognosis on the degradation of the component. Only the degradation is monitored but no analyses are performed. This makes it similar to normal condition-based maintenance [43].

3.1.7 Prescriptive maintenance

Merkt [46] states that in addition to these maintenance terms, there is another term that is not specified in the EN 13306 standards, namely prescriptive maintenance. Prescriptive maintenance is defined by Merkt [46] as:

A recommendation of one or more courses of action based on the results of corrective and predictive maintenance models. When a predictive model raises an alarm before the fault occurs, the prescriptive model will work in the direction of reducing the probability that this alarm will rise in the future, by modifying the working parameters and variables of the asset or the process affected by the fault. When the fault is confirmed, the prescriptive models will work to minimize its impact of the work context and to re-routing assets to the non-faulty production lines [46].

Marques and Giacotto [50] state that:

In other words, the prescriptive maintenance not only is based on the failures' prediction accordingly to the analysis of data patterns and trends, but also taking the specific company's maintenance process into consideration to provide detailed recommendations, and supports the solution-finding process [50].

Prescriptive maintenance is several steps further than predictive maintenance. It uses self-learning systems that can detect abnormalities, predict and indicate where the problem is and how to fix it. The system also learns from the failure so that it can provide even more detailed information next time.

3.2 Maintenance strategy for this problem

Chapter 2.5 listed some requirements to be considered in the process of choosing a maintenance strategy for the problem of unplanned downtimes. The current maintenance strategy is closest to corrective maintenance, namely repairing or replacing at the time an unplanned downtime occurs, in a few cases there is predetermined maintenance. This manifests itself in the fact that there are maintenance moments in which the replacement of parts is put at the top of a list, often based on the knowledge or data of the supplier who has determined in advance the time interval at which maintenance must be performed. Here, there is no further reasoning as to why a part should be replaced. In summary from Chapter 2.5, the proactive maintenance strategy should ensure that, based on condition monitoring data and expert knowledge, a signal can be given in time that a certain component in the line is not functioning properly and which needs maintenance or replacement.

Combining both the requirements and the earlier given descriptions of the different maintenance strategies, it can be seen that condition-based maintenance strategies should be looked at. That leaves predictive, non-predictive and prescriptive maintenance as possible options. For the problem in this research, prescriptive maintenance is too complex, as it also requires data on how maintenance is performed on the various components and which parts of, for instance, a valve need to be replaced and also the method how to replace that specific part. This requires much more detailed data than is currently available.

There is data available from the past few years that can be used to draw up trends and analyses on how certain components fail, which is why non-predictive condition-based maintenance is also discarded, as this strategy does not use analyses.

The most suitable option for a proactive maintenance strategy in this research is predictive maintenance.

3.3 Predictive maintenance

As defined earlier in Chapter 3.1.5, PdM can be used to predict the approaching end-of-life, in other words; failure, of a component or system based on condition monitoring and historical data. Many examples can be found in other industries where PdM is already widely used. Within the oil and gas industry, it is used to monitor pumps, engines and compressors and carry out timely maintenance [51]. One of the driving forces behind this is the very costly downtimes it can reduce [52].

PdM is also increasingly being used within aviation engineering. Here, the PdM strategy is applied to critical components [13]. It ensures that the parts that are likely to fail can be ordered and replaced in time. This not only prevents an aircraft from being grounded unnecessarily, but also improves safety [53].

Furthermore, many other examples can be found within the literature where PdM has been applied, including for wind turbines [54] [55] [56], marine industry [14] [57], car production lines [58] and rail networks [59].

From the above examples, it can be seen that many forms and applications of PdM are possible. It is therefore important to examine what are essential features of a predictive maintenance strategy.

Now the first thing it will be looked at is what exactly makes a maintenance strategy predictive.

PdM has two main applications: diagnostic and prognostic. According to Jimenez and Vingerhoeds [12], the diagnostics approach is used to determine the current health condition of a component or system. Jimenez et al [60] add that diagnostics is also used to identify the cause of faults.

Prognostics, according to Jimenez and Vingerhoeds [12], is dedicated to predicting the future health status or failures of the system and/ or the Remaining Useful Lifetime (RUL).

Jardine et al [61] argue that diagnostics can be viewed as a posterior event analysis and prognostics as a prior event analysis. Predictive maintenance can use either diagnostic or prognostic approaches or a combination of both. From the descriptions of both approaches, it can also be seen that both approaches complement each other. Suppose the prognostic approach fails and a fault suddenly occurs anyway, diagnostics is important to find out what caused the fault.

Diagnostics and prognostics can be performed in two ways according to Vingerhoeds et al [62]. The first is on-line. For on-line applications, data is collected, processed and analysed in real time while the system is operating. With this, alarms or other notifications for maintenance can then be issued while the system is operating [62]. On-line must take into account that there is only a limited time frame to gather the information, process it and plot actions.

The other way is off-line. With this, data is collected that is later used and analysed [62].

Since it is apparent from the descriptions that diagnostics and prognostics can both be used and are complementary to each other, both will be discussed in more detail. Diagnostics will be discussed first.

3.3.1 Diagnostics

Diagnostics is also known as condition monitoring, anomaly detection or root cause failure method. For condition monitoring, sensors can be used to monitor the functions of the equipment.

There are several commonly used condition monitoring techniques known that can be used for individual equipment within the system. Some commonly used techniques are described below.

- **Vibration monitoring:** This technique is capable of detecting many different forms of fatigue, wear, misalignment, turbulence and so on. It is commonly used to determine the condition of pumps, motors and turbines [63].
- **Process parameters:** This often falls under the normal monitoring of systems, but can be used as a basis for determining the condition of the system. This is because it is also applicable to non-mechanical parts of a system such as pipes and boilers. These include: Process efficiency, heat loss, machine temperature, fluid pressure and looptijd [29].
- **Acoustic monitoring:** based on frequencies, it can be checked and determined whether the equipment is still working properly [29].
- **Visual inspection:** based on inspectors' observations, the condition of the equipment is determined [64].

In addition to the above techniques, many others can be found in the literature that can determine the condition of the equipment within a system.

In [Chapter 2.4](#), condition monitoring techniques for valves and centrifugal pumps were already discussed. There it was noted that, according to source [29]: for valves, looptijd, position monitoring, acoustic monitoring and valve flow coefficients can be applied. For centrifugal pumps, vibration monitoring, lubricant sampling and temperature difference are commonly used techniques.

Besides monitoring the equipment, what meanings can be derived from the data should also be considered. Based on the data collected, the first step is to analyse whether the data falls within the margins defined as safe. One way to do this is with anomaly detection. In fact, anomaly detection is also known as outlier detection [65]. If the analysis shows abnormal values, a notification can be issued. Based on this, further investigation can then be done.

Based on condition monitoring and analysis of the data, a diagnosis can be made about the state of the system and whether any abnormalities have been detected.

Earlier, Jimenez and Vingerhoeds stated that diagnostics can also be used to identify the cause of a failure. This is done on the basis of observed symptoms [12].

All this together explains the first part of a predictive maintenance strategy. As indicated earlier, there is also the prognostic part.

3.3.2 Prognostics

As stated earlier, the prognostic part of a predictive maintenance strategy is to determine what the RUL is or to predict what the future condition or failure may be of the equipment.

Within a prognostic approach, a distinction can be made between roughly three types of models. These models themselves have other models underneath them as shown in Figure 14 [51]. It is also possible to use a combination of the models [60] [66].

- Knowledge-based models
- Data-driven models
- Physics-based models

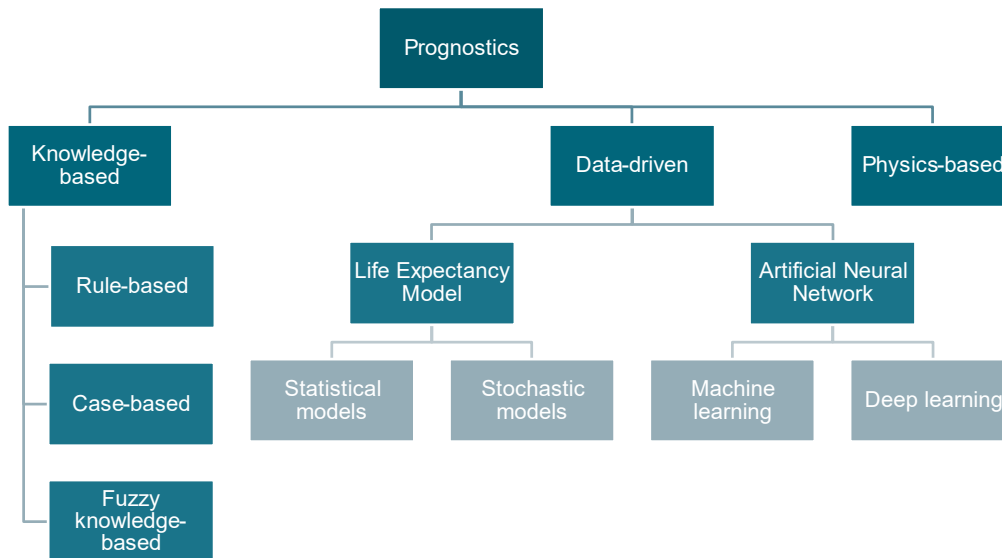


Figure 14: Overview of different prognostic PdM models [51]

Knowledge-based models

These models are based on experience and knowledge [44]. Experience and knowledge can be gained from facts, rules/laws or expert knowledge gathered over the years on the operation and maintenance of the system [62] [67]. It can be used to describe and predict component failures or degradation [68] [69]. Within the literature, three subcategories can be distinguished for knowledge-based models [60]:

Rule-based models

"If Then" relationships are used. The input is compared with knowledge in the database and then it is determined what a logical output would be [70].

Case-based models

These models are based on the knowledge previously gained from other cases or obtained from previous experiences or situations [71]. The new case is compared with others to see what the best solution is. Useful when no clear relationship between facts can be found [62] [71].

Fuzzy knowledge-based models

It is similar to the rule-based models but instead of just being right or wrong, it is now possible for it to be partly right or partly wrong [72] [73]. Other terms like hot, cold, big, small can also appear here.

There are some caveats about knowledge-based models. It is limited when looking at prognostics. It is difficult to predict anything just based on past knowledge/experience. However, it can serve as a supplement or basis if other models are used for the RUL [60].

Data-driven models

These models are based on data. This data can include operational data, environmental data, equipment data and many other types of data. With the information obtained from the data, something can be determined about the health condition of the component or system, such as degradation or RUL [60]. Data-driven models can be divided into two different subcategories according to Sikorska et al [72]:

Life expectancy models (LEM): these models aim to determine the life expectancy of individual components with respect to the expected risk of deterioration under known operating conditions [72]. LEM models can be divided into two sub categories:

Statistical models

These models analyse the behaviour of random variables by comparing it with recorded data from the past. For predictive maintenance, it is employed to be able to determine degradation and RUL, first a distribution function has to be established with which a trend analysis can be done [60]. Examples of statistical approaches include: regression analysis [74], autoregressive models [72] and Bayesian models [75].

Stochastic models

This type of models are probability models that focus on the evolution of a random variable over time [76] For this, historical data is used in most cases [77]. Two stochastic processes have been identified that can both be used: Gaussian processes and Markov processes.

Artificial Neural Network (ANN): these models can directly or indirectly calculate an expected output for RUL. For this purpose, an observation dataset is used to create a mathematical representation of the component. The physical understanding of the failure processes is not considered [72]. Common known sub categories are:

Machine learning models

These models use specialised learning algorithms, like AI, to build models from data [78]. Machine learning models, unlike physics-based, statistical or stochastic-based models, are able to work with complex relationships. The most important thing about machine learning models is the learning process. This depends on the application and the data available for the system [79]. Machine learning can also be divided into several subcategories like supervised learning, unsupervised learning and reinforcement learning [80][81].

Deep Learning

Deep Learning (DL) is a combination of machine learning and AI. It can be compared to data processing in the human brain [82]. The technology behind DL uses multiple layers to keep the relevant data and make connections in order to build computational models. A very large database is needed to create these kinds of models [82].

Physics-based models

This type of model uses the laws of physics to determine the degradation of components. With the necessary mathematical and physics capabilities and knowledge, it is possible to simulate the behaviour of a system in a model. Those models can accurately simulate how cracks or other fatigue phenomena occur, for example [83] [84]. It is very dependent on accuracies in describing physics phenomena. Also, it is difficult to include external influences in these models, although this can have a major impact on the behaviour of the system [75].

3.4 Model selection

To determine which model or method is the best fit, the complexity of the system must be considered [12]. It is also necessary to determine how much expert knowledge there is about both the system and modelling techniques [61]. In addition to this knowledge, the data can also be looked at, both the data per se and the knowledge present about this data. Here, it is important to be able to form a complete picture of what is present in terms of data and knowledge so that a model can be chosen that matches it. As an example, it is irrational to choose a physical model if there is no knowledge about how the degradation of components can be described mathematically and physically. Another example could be that if few data are available, a deep learning method should not be chosen, because this method needs a lot of data to build a mathematical model with.

3.5 Conclusion

This chapter has answered the sub-research question: *What is the best maintenance strategy for this problem compared to the current maintenance strategy?*

The current maintenance strategy is a kind of combination of corrective and preventive maintenance. Corrective maintenance is reflected in that a component is replaced or repaired only when it is broken and therefore has caused unplanned downtime. Preventive maintenance is that there are scheduled maintenance moments when replacement of parts is started at the top of a list. However, there is no reasoning or rationale behind this. The proactive maintenance strategy therefore requires looking at how data combined with expert knowledge can be used to predict when a component is going to fail. The best-suited maintenance strategy is therefore predictive maintenance, with which data-based analyses can be done to see when a particular component degrades and consequently increases the likelihood of failure. For PdM, there are two applications, diagnostic and prognostic, which are complimentary to each other to be able to form a picture of the current condition and predict when a failure will occur. For both applications, there are different ways and methods that can be used. As indicated in [Chapter 3.4](#), data analysis can be used to see which method best fits the problem statement. The [Chapter 4](#) will elaborate on the data analysis and which method best fits the outcomes from the data analysis.

4. Model development

Within this chapter, the central sub-research question is: *Based on the available data and knowledge, what kind of method is most appropriate to develop a model with for this problem?*

To answer this sub-research question, this chapter first conducts a data analysis. It will provide a picture of the available data and knowledge. In order to determine which method fits best, the second part of the chapter looks at the methods in more detail. In the third part of the chapter, a method is chosen based on established criteria. This method is then briefly explained so that the basic principles are clear.

4.1 Data analysis

To determine the best method and model to be used in this research, a data analysis must first be done. The purpose of this is to identify what data is available and usable, what knowledge about the data is present and how this can be used in the predictive maintenance strategy to improve the reliability of the production line.

4.1.1 Overall line analysis

As previously stated in [Chapter 2.1](#), here the focus will be on part of a production line. It concerns the piece between the lemonade storage tank and the filling process. In this piece of line there is a cooling and heating section, a storage tank, carbonisation tank and filter. Only the valves and pumps in this part of the line will be included. This amounts to 25 valves and six pumps.

The parts in the line all have a PLC that can be used to control the component and give reports back to the operator. This data is all stored in a database.

For data analysis, data between 01-02-2021 and 03-02-2023 was used. The line is most frequently used between Monday morning and Friday evening. Sometimes there is run-out and production is completed on weekends. This gives the dataset used in the further research. Furthermore, besides data on and from the components, data is also available on when the line was used for production and when for cleaning.

The historical data can be divided into production planning data and process data. Process data refers to the data that follows from the PLCs and other sensors from the line. The historical production schedule data will be used to understand how long the processes take and when the line is down.

The historical data used is from the period 01-02-2021 to 03-02-2023. In terms of hours, that is 17568 hours.

From this analysis, it can be seen that on average in a year (8784 hours), the line is in use 57% of the time, see [Figure 15](#). Of this, 53% is for production and 4% for cleaning. The remaining 43% is spent on scheduled and unscheduled downtimes and weekends. Scheduled downtimes in this research are the times when the line is stopped outside weekends. Every year, there are two weeks when the line is not used. These weeks can be used for preventive maintenance.

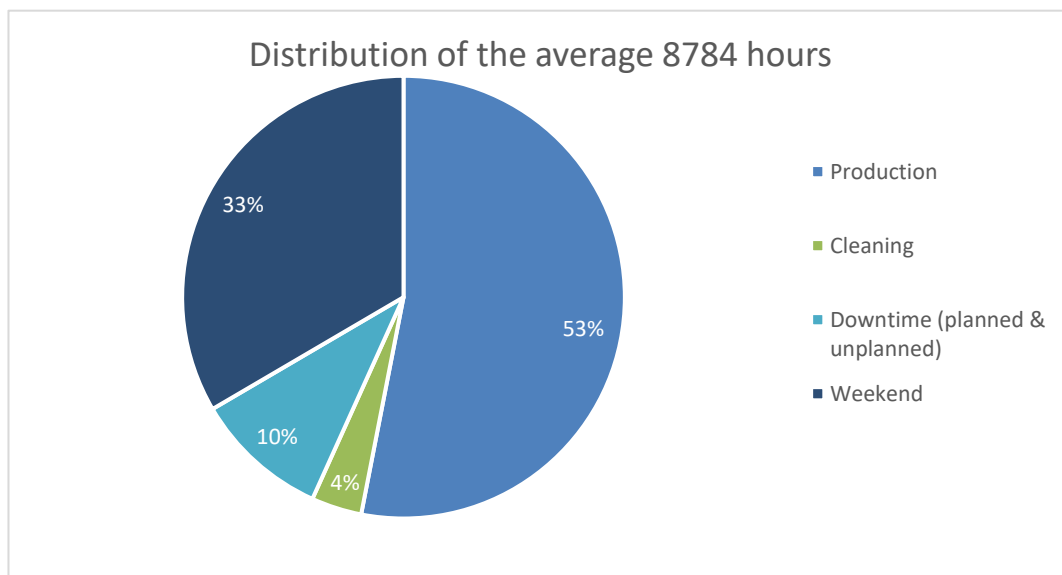


Figure 15: Pie chart of the hours of the production line

When zooming in further, the time the line is in use is 57% of the time, which corresponds to 4985 hours. Since the production process consists of the steps; start-up, effective production and shutdown, another distinction can be made

between this. Shutdown is the last step of the production process. It means that the production process is stopped either because production is finished or because failure of one of the components occurs. Together, this gives the distribution of time as shown in [Figure 16](#).

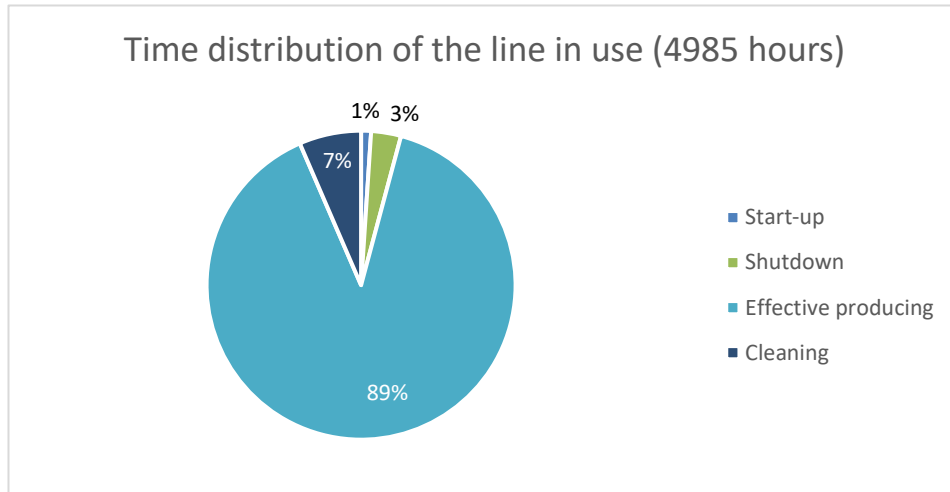


Figure 16: Pie chart of time production line in use hours

Finally, attention is paid to the number of hours the line is down. This could be due to planned downtimes. In these, the line is made ready for the next process or planned downtime weeks to carry out preventive maintenance. Weekends are not considered here. Then it comes down to a total of 862 hours of downtimes on an annual average (excluding weekends, including planned downtime weeks). Of this, 54% is planned downtime and 46% is unplanned downtime, as shown in [Figure 17](#).

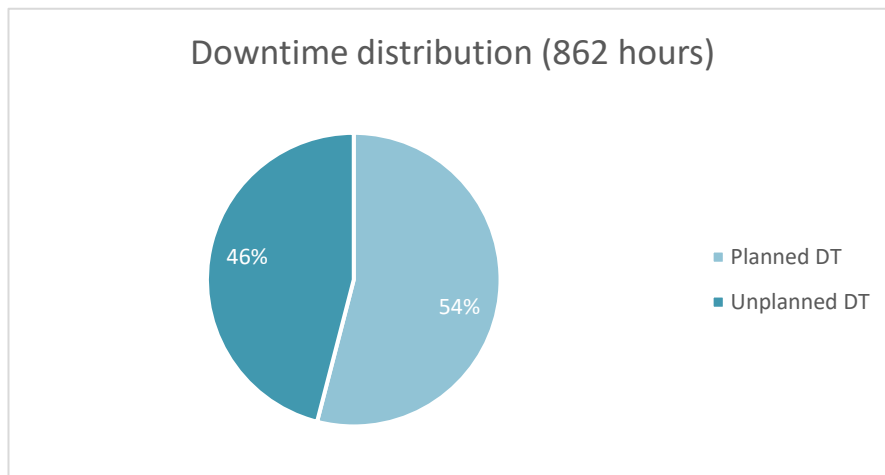


Figure 17: Pie chart of downtime hours

4.1.2 Valve analysis

When looking at valve data, the analysis should clarify the behaviour of the valves and how this is reflected in the data in combination with expert knowledge. First, the analysis looked at what kind of data logs are stored.

In the case of the valves, data is stored on looptijd, which is the time it takes the valve to open or close, and command open or close. Looptijd is Dutch for run time, both terms can be used and have the same definition in this research. In the case of the butterfly valve, this is the time it takes the valve to turn the disc in the centre 90 degrees so that the valve is either open or closed. In the case of mixproof valves, looptijd is the time it takes the valve to open or close a particular seat. The expert knowledge was used to determine when the valve will fail, in other words, what are the anomalies and trends to look out for when analysing the data.

This research will look at looptijd to determine the failure probability of a valve.

Results of analysis

Most valves in the system have a maximum looptijd of 15 seconds. If the valve takes 15 seconds to open or close, a signal is fed back to the PLC indicating that the valve is not yet in the correct position. This was pointed out by the experts. It is possible for a valve to restore itself when it is reopened and closed. In that case, good signals would be transmitted to the PLC. Therefore, if this is the case, there was a false fault signal earlier. To filter out the false fault signals, a check is made after how much time the PLC receives normal signals again from the valve. If this is within 30 minutes of giving the 15-second looptijd message, it means the valve did not fail. If the valve did fail, it takes more than 30 minutes to stop the process, repair the valve and restart the process.

Some valves have a different maximum looptijd, such as 20, 25, 30 or 50 seconds. However, the principle here is the same as described above for the looptijd of 15 seconds. The first analysis of the data looked at how often each valve failed. An overview of this is shown in Figure 18. Figure 18 shows that often only one valve failed at a time. There are three times when three valves failed at the same time.

This shows that there were 70 unplanned downtimes, spread over 57 unique failure moments, in the 2-year time frame.

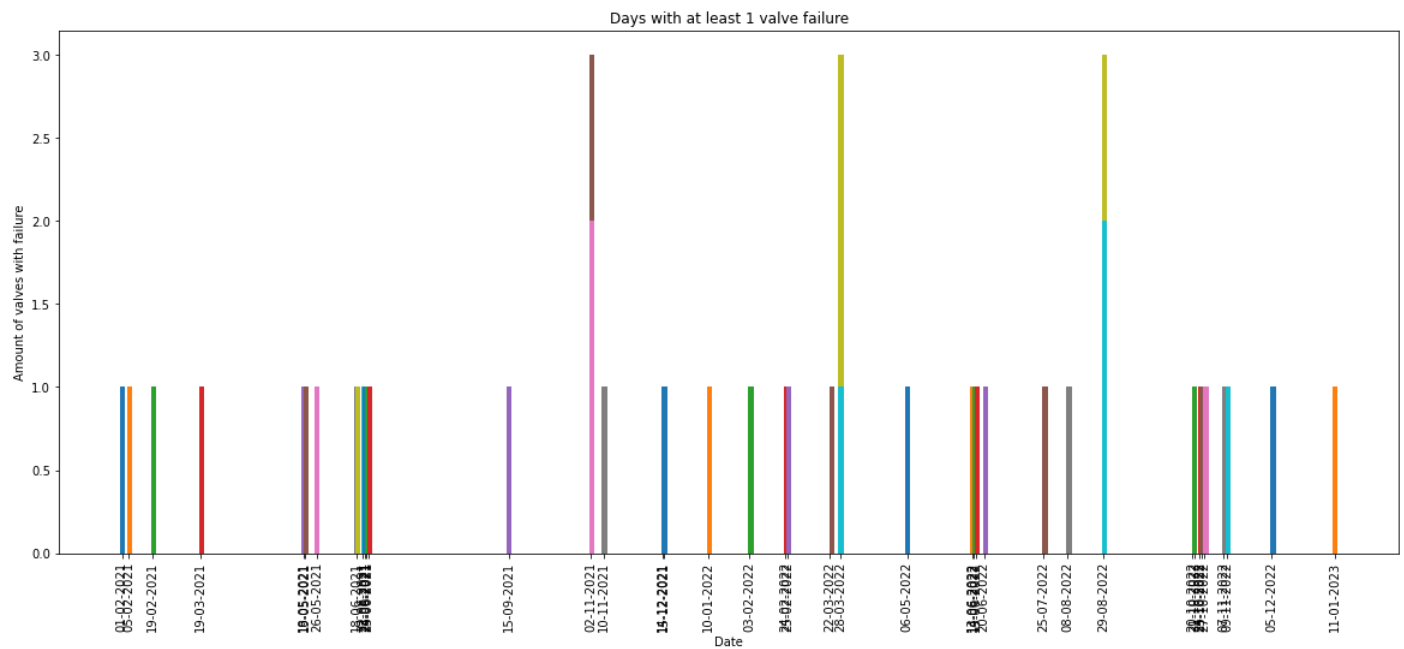


Figure 18: Overview of valve failures 01-02-2021 till 03-02-2023

From the data analysis, it can be further observed that just before failure, the valve already shows more high looptijd values. Figure 19 shows how the looptijd of valve V5.2 increase more and more. Until finally the value of 15 seconds looptijd is reached and the valve fails.

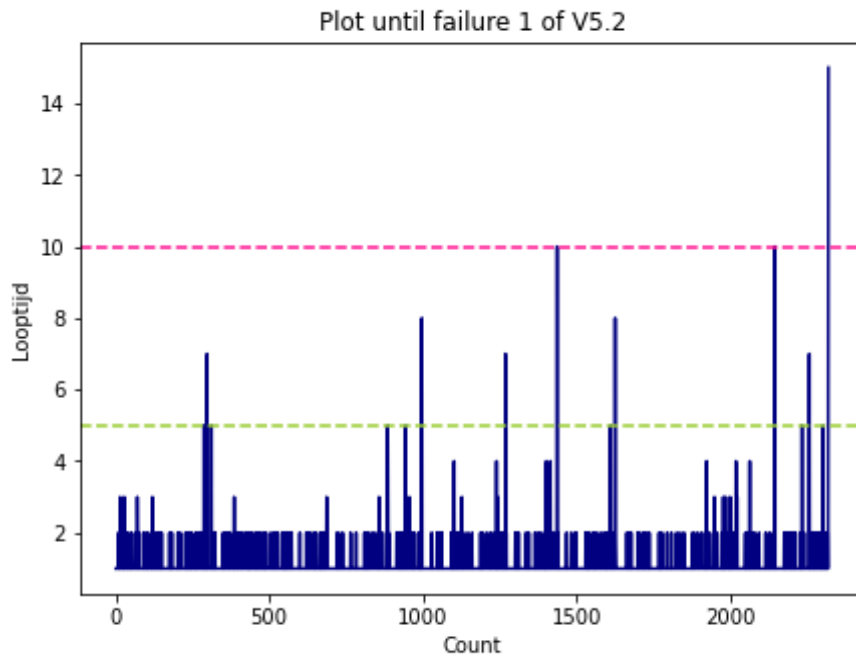


Figure 19: Increasing looptijd of valve V5.2

The data analysis further shows that four of the 25 valves have no dates on which the valves fail. It was decided to leave these valves out of the research. The research will subsequently focus on the 21 valves for which there is data on failure.

Furthermore, it was concluded from expert knowledge that the higher the looptijd, the higher the probability of failure. According to experts, the logistic function [85] best represents the relationship between the looptijd and the failure probability of a valve. The logistic function for this research expresses the probability of failure given a certain value for the looptijd. This can be expressed using the following Eq. 4.1 [85]:

$$\text{probability of failure} = \frac{L}{(1 + e^{(-k*(looptijd-m)})})}$$

Eq. 4.1

- L* is the maximum value, in this case 1, because then probability of failure is 100%
- k* is the slope of the graph
- m* is the average looptijd
- looptijd* is the time it takes for the valve to open or close

Chapter 5.2.1 shows how the parameters of Eq. 4.1 are determined for each valve.

This function gives the following graph, see [Figure 20](#) with different values for m :

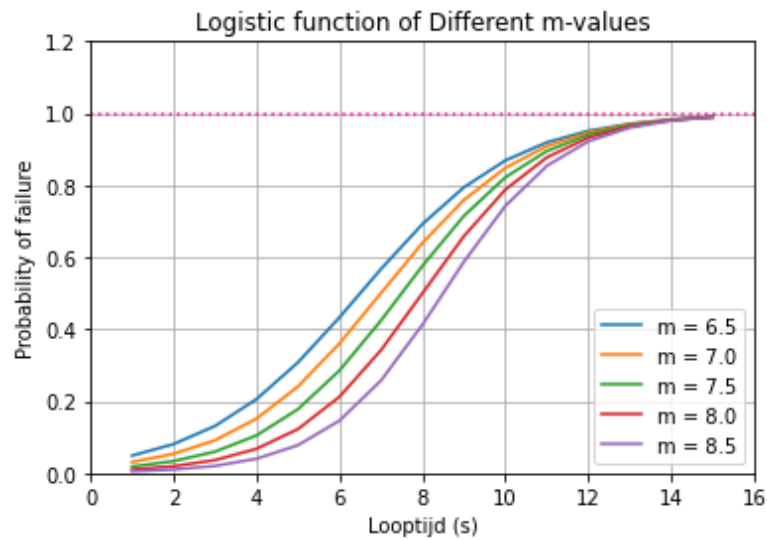


Figure 20: Graph of logistic function with different m -values

The value of m , average looptijd, determined when the graph passes through the point of 0.5 failure probability. For each valve, an unique logistic function can be constructed that best represents the behaviour of the valve. For this, parameter k can also still be used to determine the correct shape of the graph.

If a valve has a looptijd different than 15 seconds then a logistic function can also be constructed.

In short, a probability of failure can be determined from the incoming input data using the logistic function of the valve.

4.1.3 Pump analysis

There are six pumps in the line section being analysed. These are all centrifugal pumps. Since there is only data available on the type of pump, theoretical data and knowledge is used for the pumps. It is chosen for this to still be able to paint a picture of what parameters are needed and how that can be used in the later model.

As stated earlier in [Chapter 2.3.4](#), the ways in which a pump fails can be divided into three categories. If really specific causes are considered, the examples below are the common known causes of pump failure within the case study soft drink factory.

- Cavitation
- Leaking shaft seal
- O-ring leaking

All these failure mechanisms affect the pump performance.

Cavitation causes many problems when it occurs. Cavitation occurs when the pressure in the pump is lower than the vapour pressure of a liquid, vapour pressure being the point at which a liquid turns into gas. Often this occurs at the inlet (suction side) of the pump because this is where the pressure is reduced. This creates gas and/or air bubbles. Then the action of the pump increases the pressure again (to above the vapour pressure of the liquid), but now there are bubbles in the liquid that implode, often this happens in the pump at the impeller [86][87]. This creates irregularities on the surface of the impeller which can leave liquid behind, this causes cavitation. This process is shown in [Figure 21](#).

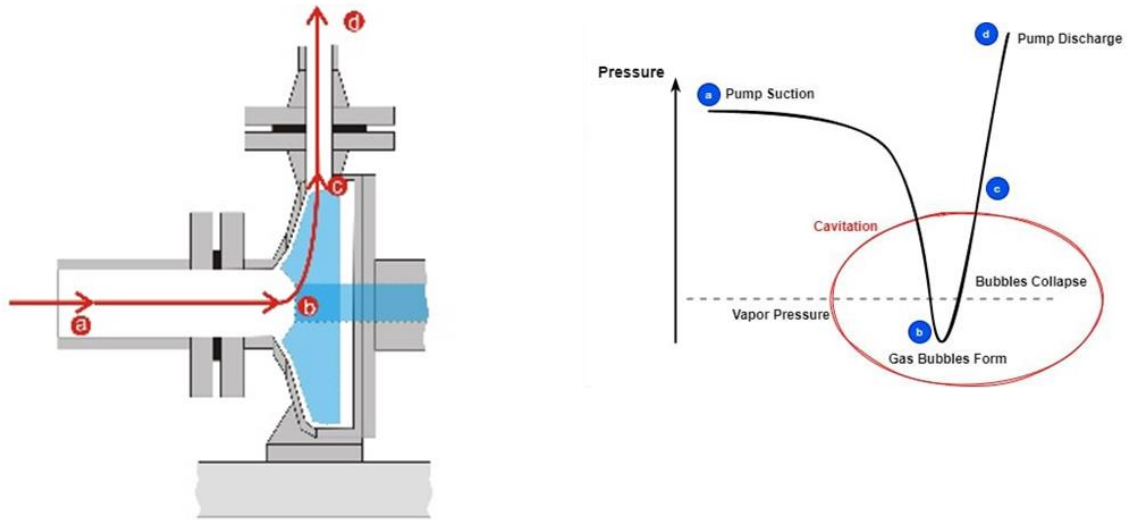


Figure 21: Cavitation occurrence [88] [89]

Briefly, cavitation can be said to be caused by the temperature of the fluid, the vapour point and the inlet pressure [90]. Temperature affects the vapour pressure of the fluid.

Cavitation causes vibration and noise. The vibrations in particular have a major impact on the other mechanical parts in the pump. These cause bearings to get out of alignment and change the interplay of forces on the components. This affects the service life of the mechanical components and the overall life of the pump [91].

The shaft seal can fail mainly because it gets dry, so it is important to apply enough lubrication so that the pump can keep working. Often, this type of failure is only noticeable when oil or other type of seal leaks from the pump [92].

For the O-ring, it ensures that no leakage occurs between different parts of the pump. The ring is exposed to the temperature and pressure in the pump. The rubber of the ring has a certain life span that depends on the temperature and the type of fluid forced through the pump. Failure of the O-ring produces leakage and can be observed when moisture is encountered at the pump [92].

4.1.3.1 NPSH margin

Of the failure mechanisms mentioned earlier, cavitation is the one that can be used quantitatively in this research. To determine whether cavitation is likely to occur, this research looks at the difference between $NPSH_A$ (Net Positive Suction Head Available) and $NPSH_R$ (Net Positive Suction Head Required).

$NPSH_R$ [m] is given by the manufacturer and $NPSH_A$ [m] can be determined using Eq. 4.2 [93]:

$$NPSH_A = \frac{p_i}{\rho g} + \frac{V_i^2}{2g} - \frac{p_v}{\rho g}$$

Eq. 4.2

p_i	is the pressure at inlet of the pump [Pa] or [N/m ²] or [kg/(ms ²)]
V_i	is velocity at pump inlet [m/s]
ρ	is fluid density [kg/m ³]
g	is acceleration of gravity [m/s ²] = 9.81 m/s ²
p_v	is vapor pressure of the liquid [Pa] or [N/m ²] or [kg/(ms ²)]
$NPSH_A$	is net positive suction head available [m]

For the segment of line being analysed, the liquid pumped through the line consists of sugar and water. Or by water and detergents if cleaning is done. The inlet pressure (p_i) and inlet velocity (V_i) should be measured by sensors.

The vapour pressure (p_v) and density (ρ) can be calculated. These depend on the composition of the liquid. See Appendix C for the calculations.

To prevent cavitation from occurring, the minimum must apply, see Eq. 4.3 [94]:

$$NPSH_A > NPSH_R$$

Eq. 4.3

Another way is to determine the NPSH margin. In ANSI/HI 9.6.1-2012 Guideline for NPSH Margin [95], values are given for the minimum margin the NPSH margin must meet [96], see Eq. 4.4 for the NPSH margin.

$$NPSH \text{ margin} = \frac{NPSH_A}{NPSH_R}$$

Eq. 4.4

In this research, according to ANSI/HI 9.6.1-2012 Guideline for NPSH Margin [95], the margin is equal to 1.1 [96]. If it is lower than this margin, $NPSH_R > NPSH_A$ resulting in cavitation will occur, as shown in Figure 22.

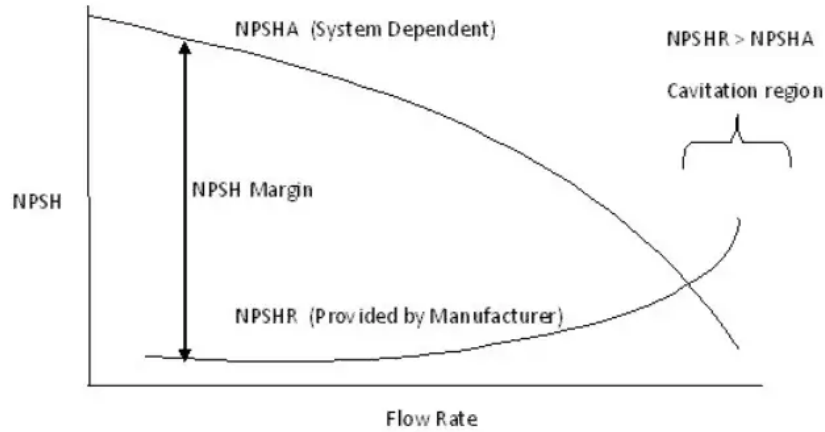


Figure 22: NPSH margin [97] [98]

4.1.3.2 Pump efficiency

Besides looking at whether and when there is a risk of cavitation, the performance of the pump can also be looked at to determine its condition. In this research, efficiency is used to determine whether the pump is in good condition. Pump efficiency averages between 40% and 92% [99].

The pump will have the longest life if it is operating at the BEP (Best Efficiency Point). This is the point at which the head-flow curve reaches maximum efficiency [35]. It is not realistic to assume that a pump will always operate at or around the BEP. Therefore, a POR (Preferred Operating Region) is often given.

This is often between 80-110% of the BEP [100]. In addition, an AOR (Allowable Operating Region) can also be determined, which is a wider region than POR. In the AOR, the pump life is shorter and there will be more noise and vibrations than in the POR, see Figure 23. Anything outside these regions should be avoided, as this causes damage to the pump and a consequence is pump failure [99].

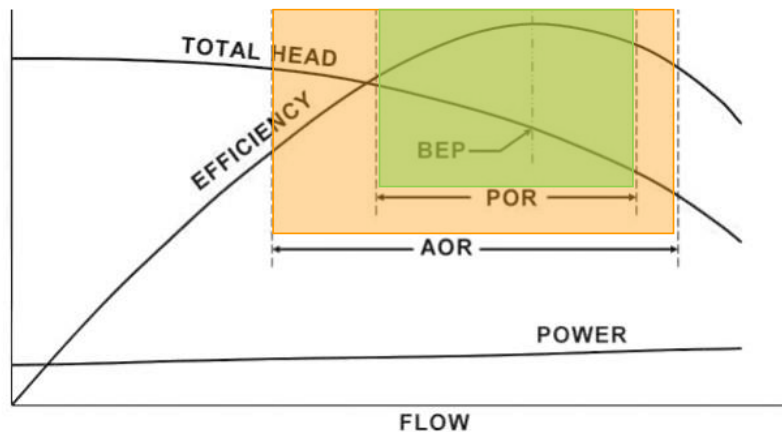


Figure 23: Graphical view of BEP, POR and AOR [96]

Pump efficiency can be determined using Eq. 4.5 [93]:

$$\eta_{CP} = \frac{P_{out}}{P_{in}} = \frac{\rho g H Q}{T \omega}$$

Eq. 4.5

η_{CP}	is pump efficiency
ρ	is fluid density [kg/m ³]
g	is acceleration of gravity [m/s ²]= 9.81 m/s ²
H	is head [m]
Q	is flow [m ³ /s]
T	is shaft torque [Nm] or [kg·m ² /s ²]
ω	is shaft angular velocity [1/s]

P_{in} is the input power on the suction side of the pump. P_{out} is the output power at the discharge side of the pump. Based on the given data from the manufacturer, the input power can be determined. However, a flow meter is needed to determine what the flow rate is at the inlet of the pump [35].

Another way is to place a watt or power meter just before the inlet of the pump to use it to determine the P_{in} .

P_{out} depends on more factors. Here, the head and flow rate at the outlet of the pump have to be taken into account. For this, flow meters and pressure sensors should be used. The pressure sensors should be present both at the inlet and at the outlet. The pressure sensors can be used to determine the difference in height and thus the head [35].

If it is assumed that there are no further losses then the head can be calculated using the Eq. 4.6 [93]:

$$H \approx \frac{p_2 - p_1}{\rho g}$$

Eq. 4.6

ρ	is fluid density [kg/m ³]
g	is acceleration of gravity [m/s ²]= 9.81 m/s ²
H	is head [m]
p_1	is pressure at inlet of the pump [Pa] or [N/m ²] or [kg/(m·s ²)]
p_2	is pressure at outlet of the pump [Pa] or [N/m ²] or [kg/(m·s ²)]

Head of a pump indicates the height to which water can be pumped against gravity. This is expressed in metres [93]. This can then be used to determine what the values of P_{out} are. Enter this in Eq. 4.5 and the efficiency value comes out.

Repeating this for all values observed by the sensors can determine what the pump's efficiency curve is.

Checking whether there is a chance of cavitation and determining the pump efficiency curve are two of the many ways to look at the condition of a pump. The values from the analyses given earlier of whether or not cavitation is likely to occur and determining efficiency can also be used to determine the failure probability of the pump.

One way to determine the failure probability using the options given earlier is to use the reliability curve of the pump [35].

Furthermore, some other failures are shown when those will occur, see Figure 24.

The reliability curve is maximum at the BEP point of the pump. As more deviation from this point occurs, the reliability of the pump decreases, as shown in Figure 24 [101].

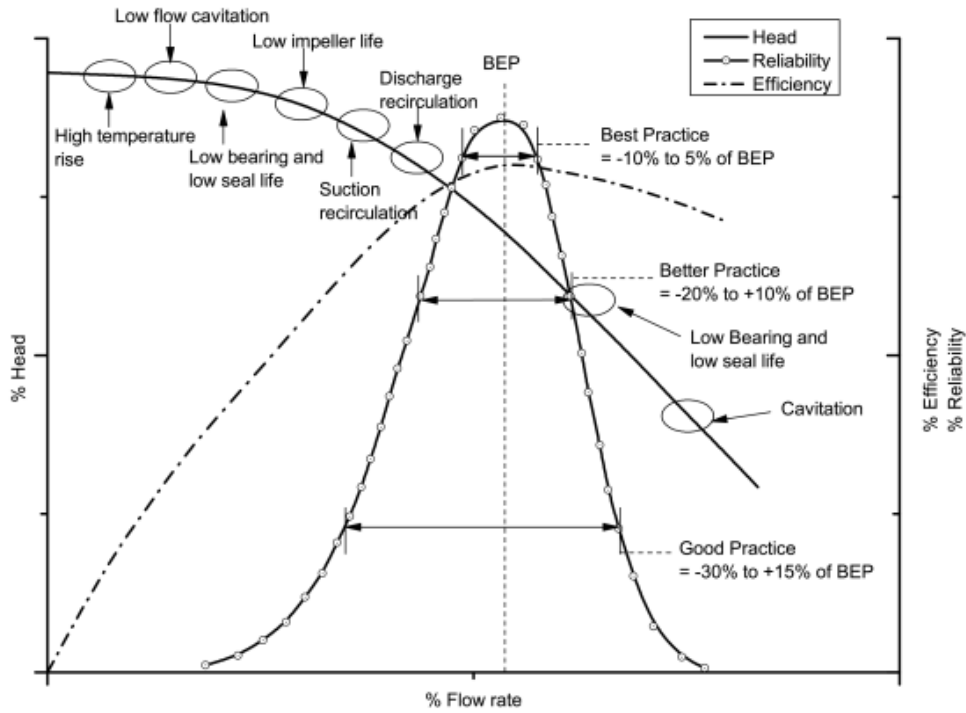


Figure 24: Pump curve reliability [101]

The pump efficiency is used to determine the BEP. The BEP is the point at which the reliability curve is maximum. In addition, it was previously stated that the POR is between 80-110% of the BEP. In the reliability curve, see Figure 25, this is reflected at the point where better practice is stated.

The AOR is determined by the ANSI/HI 9.6.1-2012 Guideline for NPSH Margin [95]. This is the point at which cavitation will occur.

Barringer and Nelson [97] [102] have linked the reliability curve to the Mean Time Between Failure (MTBF).

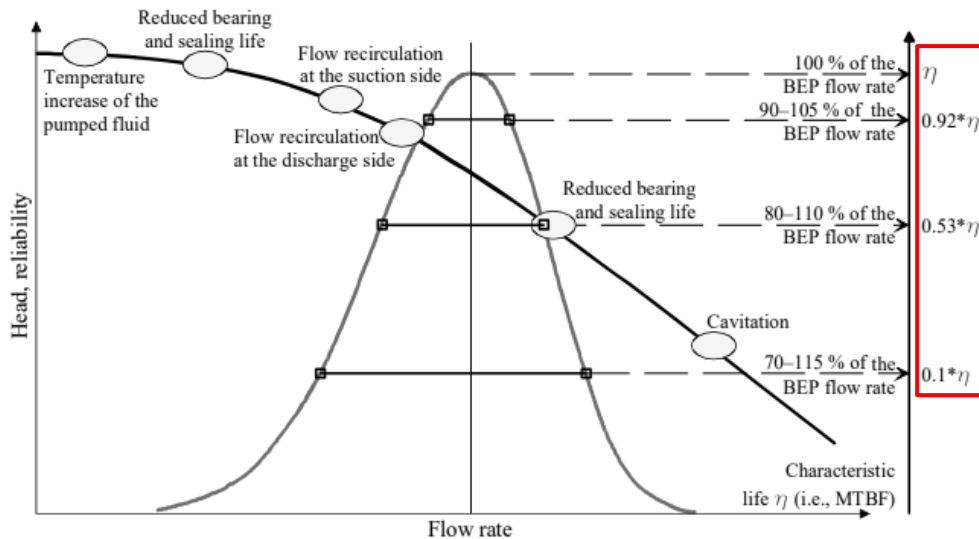


Figure 25: Pump curve reliability Barringer [100] [103]

The values in the red box in Figure 25 show how Barringer and Nelson [100] link reliability to MTBF. At the moment the pump is operating at its BEP the reliability equals 1. The MTBF then equals the values given by the manufacturer or experienced during operation. As soon as the pump operates between 90% and 105% of the BEP, the reliability of the pump decreases and the MTBF reduces by 8% (0.92η). Note that η here stands for the MTBF and not efficiency (η_{CP}).

Once the pump is operating at the edge of the POR then it can be seen in the curve that the first problems may arise. The MTBF is reduced by 47% (0.53η).

If the pump falls below the AOR value then cavitation will occur and the MTBF is further reduced by 90% (0.1η).

By relating the reliability to the MTBF, the probability of failure can be determined.

Figure 25 also shows the point at which cavitation occurs. By looking at which point on the reliability curve belongs to it, the influence of cavitation on the lifetime of the pump can be determined.

In short, the curves for the pumps in this research need to be established to determine the condition of the pump and how that affects the failure probability. But first it has to be determined how to find the BEP and related efficiency curve.

4.1.3.3 BEP

The best efficiency point (BEP) is the point at which the pump will perform best and have the longest lifespan. The pump will never be able to reach 100% efficiency. According to the API610 standard, the BEP of most single stage centrifugal pumps is between 80-85% of the shut-off head [104] [105] [106].

By establishing a formula for the head flow rate given by the manufacturer, the BEP can be determined.

It is assumed, in this research, that the BEP is at 85% of the shut-off head.

It is further assumed that a pump can achieve a maximum efficiency of 80%.

Based on the above assumptions and the manufacturer's data [90] [107], the H-q, NPSH_R-Q and efficiency curves can be prepared. Once these curves are established, the reliability curve of each pump can be determined.

Figure 26 shows an example of the pump curve of one of the pumps in the line. The pump charts also indicate the different operating regions. In Appendix D all the curves for all the different centrifugal pumps in the line can be found.

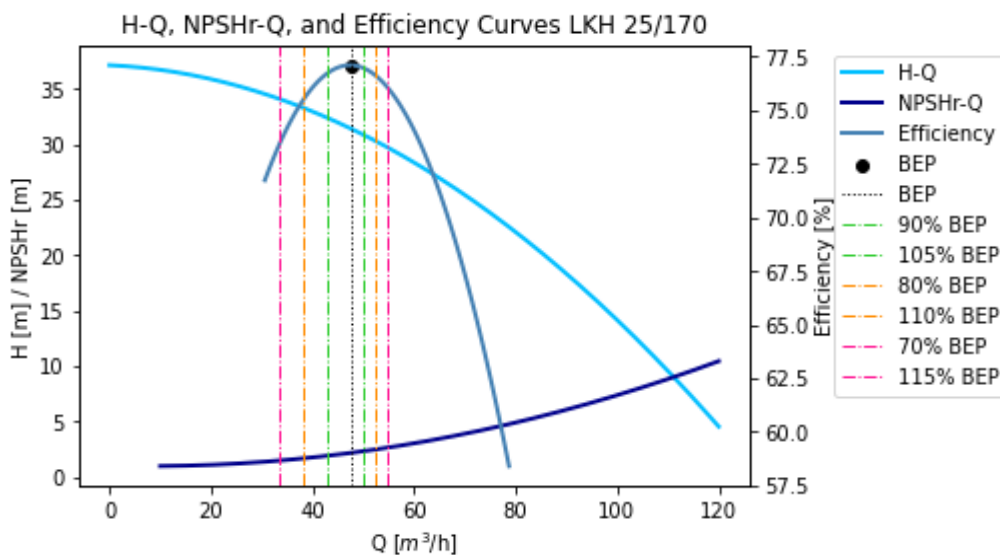


Figure 26: Pump curves of pumps P_{4.1} and P_{5.1}

4.1.4. Data conclusion

All in all, this research uses several functions that can be constructed from data. For the valves, a logistic function will be needed to relate the looptijd of the valve to the probability of failure. It will also look at how often each valve has failed in the given time frame.

For the pumps, too little data is available to find the right parameters that say something about the condition.

Therefore, the manufacturers' data and the theoretical way of determining pump curves and related reliability will be used. This data and curves can be used to determine the condition of each pump. For each unique pump, proprietary curves can be established that can be used to express the pump's behaviour and probability of failure.

For the data for the pumps, a synthetic dataset will have to be used. In this data, there should be data on pressure, density and flow velocity. This can then be used to determine whether the pump is operating within or outside the NPSH margin. In addition, this can be used to determine at what percentage of BEP the pump is performing. Once this is known, it can be related to Barringer's reliability curve.

By performing the analyses, a clear picture has emerged of what data is available and how that is related to the conditions of the components in the part of the production line. Based on this, a method should be chosen to develop the basis of a model for the improved maintenance strategy.

4.2 Method selection

As stated earlier in Chapter 3.4, data analysis can be used to see which method best fits the available data and the complexity of the system. In this research, the aim is to start predicting from the condition monitoring data when a failure will occur so that timely maintenance can be carried out. If unplanned downtime does occur, it would be desirable to be able to reason back what caused it.

Looking at the listed categories and corresponding explanation for prognostic models (knowledge-based, data-driven and physics-based) in Chapter 3.3.2, a data-driven model is the best fit. Within the data-driven models, an LEM or ANN should then still be chosen. Because there is not enough data and knowledge available to create a mathematical representation of each component, ANN is discarded. An LEM model will therefore be used in this research. As indicated earlier, a statistical or stochastic approach can also be chosen within the LEM. Sikorska et al [72] have investigated the different methods that fall under these categories and thus created the overview shown in Figure 27 based on [72]. These methods will be discussed further in the following sections, so that a choice can then be made as to which method of LEM best suits this research.

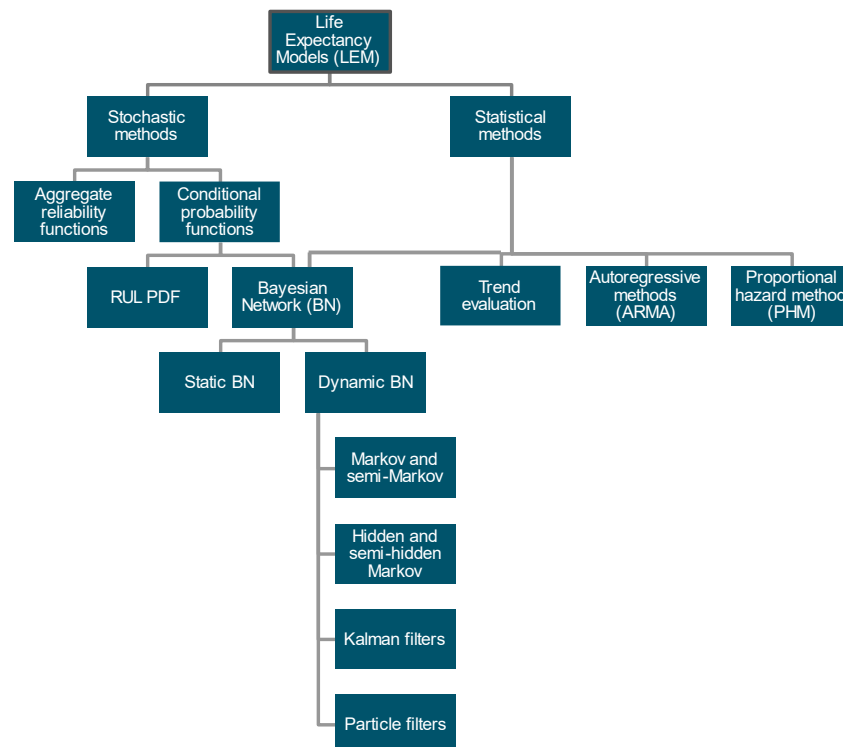


Figure 27: LEM overview [72]

4.2.1 LEM stochastic methods

Stochastic methods provide information about the reliability of a component as failure probabilities with respect to time, this includes the MTBF. Stochastic models work with the evolution of a random variable over time [108]. Historical data is used for this purpose [109]. Stochastic methods rely on the assumption that the time to failure of identical components can be considered statistically identical and independent random variables. Therefore they can be described by probability density functions (PDFs). Because it depends on the number of times failures occur in a dataset, censored data, such as data from times when there was no failure, can also be used to improve the accuracy of the method [110] [111]. There are numerous types of stochastic methods, some are highlighted below.

4.2.1.1 Aggregate reliability functions

Performing reliability analysis based on aggregated failure data is one of the ways most often used to analyse asset performance.

The method uses the analysis of failure times of a population of equipment. The analysis then determines a probability function and also a related hazard-rate function for this population. The probability function provides information on when a failure is expected to occur, it does not address the course of a single failure [72]. Several distributions can be used to model the failure data. Of these, Gaussian and Weibull functions are the most commonly used. Gaussian functions are employed when monotonic and gradual degradation are modelled [112]. Weibull functions are frequently used when multiple failure types need to be described. Even the well-known bathtub curve is composed of three Weibull functions, each describing a different dominant failure mode [72].

To shape the data to the different distributions, it is necessary to have a large dataset of the failures. In addition, reliability analysis is not enough to say anything about the RUL if the failure distribution is exponential.

The last point of interest is if there are too many different failure modes taken together then this will not give a reliable forecast about the RUL [72].

4.2.1.2 Conditional probability functions

Another subcategory of stochastic methods are the methods based on conditional probability functions. Here, the current state is described by a condition reliability function then the predicted expected behaviour is updated using Bayes theory.

A conditional reliability function is defined as the probability that a component continues to function without failure for a time frame t , given that the component has already survived to a given time, T [113]. The RUL function is defined as the conditional expected time to failure, given the current state [114]. Again, there are several methods that can be used, these differ in how the conditional reliability function is determined as what information is used in determining the current state.

4.2.1.2.1 RUL PDF

This method is actually an extension of the aggregate reliability function method. Namely, it uses the probability density function (PDF) which is constructed using the aggregate reliability method. Information is then retrieved to locate a specific point on the distribution. The distribution is then adjusted using Bayes theory to account for this information. With each new observation, this process is repeated. The final distribution is called the RUL PDF, here the confidence interval can also be derived [115]. The accuracy of the prognosis on the RUL increases as end-of-life is approached.

4.2.1.2.2 Bayesian networks (static)

Bayesian networks or Bayesian Belief networks (BN) are probabilistic graphical models that represent a random set of variables with associated probabilistic interdependencies [72]. BN models are a separate method because they can also be used as statistical methods, it depends on the type of information used to build the final model [66].

A BN consists of a collection of nodes that serve to represent the random variables that can take on different states. Lines can be drawn between the different nodes to indicate correlation. Conditional probabilities can be used to indicate how strong the relationship is between the nodes. This gives each node a conditional probability table that indicates the probability for each state of the node based on the states of the linked nodes [116].

BNs, because of the structure and conditional probabilities at each node, have the ability to determine the probability that a particular event may occur in the near future. The advantage of the BN method, because probabilities are used, is that the confidence interval is immediately determined [117].

Bayesian inferencing is used to update the states and probabilities of the nodes when new data is imported into the trained BN model [116] [118].

Dynamic Bayesian network

Dynamic BNs are used when time series data are modelled [119]. The arrows are then used to model a time step. Some methods of this are also known:

Markov methods

Markov methods assume that a component is in only one of a finite number of states at each moment in time. Probabilities are defined both for the states and for the transition state between different states. Based on this, probabilities for future failures can then also be determined. Markov methods further rely on the principle that future states are independent of previous past states [72]. In a Markov method, the time a component spends in a given state is assumed to be exponentially distributed. In addition, it is assumed that there is a constant failure rate [120]. Semi-Markov methods differ from Markov methods in that they do not assume that the time a component spends in a given state must necessarily be exponentially distributed. Semi-Markov methods also allow for other distributions [72]. One of the major drawbacks of Markov methods is that a separate model must be created for each potential failure mode. In addition, these types of methods and models are computationally intensive, even for a model with only a few states [66].

Kalman filters

Bayesian estimation with Kalman filters are used when the state of a dynamical system needs to be determined from a series of incomplete and noisy measurements in order to minimise the mean squared error. Kalman filters are defined by the state estimate and error covariance, this applies to any instant. The assumptions associated with Kalman filters are that process noise and measurement noises are Gaussian, white, independent of each other and additive [121]. Kalman filters determine the posterior PDF by extrapolating from the previous state. Kalman filters are very complex and require a lot of computing power, as both all covariance and model parameters have to be recalculated for each iteration [72].

Particle filters

Particle filters are used for estimating the posterior distributions in BN models, this type is not bound by the assumptions of Gaussian noise like the Kalman filters. This method is better known as Monte Carlo simulation and used when the posterior distribution is multivariate or non-standard [122]. Particle filters use Sequential Importance Sampling to simulate the next state in each iteration of the filter. This is done by taking a set of random samples "particles" from the theoretical density function and then adjusting the corresponding particle weights for each iteration. In the process, dynamic noise is also generated at each cycle [72]. Establishing a good working model is very complex, if too many iterations are done then the filter starts to distort resulting in the posterior PDF approaching zero [78].

4.2.2 LEM Statistical methods

Within life expectancy methods, there is also the subcategory of statistical methods. Statistical methods use comparisons based on inspection results of similar components to estimate the occurrence and progression of failures. By comparing with components that exhibit "healthy" behaviour, future deterioration can be determined. Statistical methods are used as an alternative to ANN if there is no suitable model of the physical process [72]. The models use condition or process monitoring data. There are several statistical methods that can be used, some of which will be discussed below.

4.2.2.1 Trend evaluation

The easiest way to predict the RUL is to use trend analysis of a single monotonic parameter correlated with the remaining lifetime. The parameter is plotted against time, in addition some alarm levels are determined that indicate that a component is heading towards the end of life. The trend can then be analysed using regression methods [72]. As soon as new data arrives, the moment in the graph can be checked to which it corresponds and the RUL can be calculated on that basis. The corresponding confidence levels can only be determined if interpolation has taken place. If extrapolation has been used, it is not possible to determine the confidence levels. In addition, this model struggles when it comes to the reliability of the expected RUL when multiple failure modes are used in the same model.

4.2.2.2 Autoregressive methods

ARMA, ARIMA and ARMAX methods are used to model and predict times series data [72]. This involves the assumption that the future value is a linear function of past observations and random errors [123]. The methods differ among themselves in how which linear function is used to relate inputs, outputs and noise.

ARMA and ARMAX methods should only be used for stationary data, this because they can remove temporal trends. In addition, the autocorrelation should also be time-independent [123]. To show that the stationarity assumption is correct, a trend analysis should be done before modelling. ARIMA methods use integration and therefore have no condition that it can only be used for stationary data.

ARMA methods are especially suitable for short-term forecasts. For the long term, there is too much influence of dynamic noise and sensitivity of initial conditions and accumulation of systematic errors [72].

4.2.2.3 Proportional hazards modelling

PHM base the model on the influence of covariates on the lifetime of the component [124]. Covariates are explanatory or additional variables. PHM assumes that there is a multiplicative relationship between the covariates. PHM models model component deterioration as the product of a baseline hazard rate and a positive function. The positive function gives the effect of the operating environment on the baseline hazard and is represented as a vector of covariates with the corresponding vector of unknown regression parameters. The RUL is derived from the corresponding survival function [124].

PHM is subject to a number of assumptions [72]:

- Times to failure are independent and identically distributed;
- Individual covariates are independent;
- The effect of the covariates is assumed to be time independent;
- All influential covariates are included in the model;
- Covariates have a multiplicative effect on the baseline hazard rate;
- The ratio of any two hazard rates is constant with respect to time.

4.3 LEM method selection

A model must be created in which one of the above methods is the basis for determining the probability of failure of the production line. Thus, the method is a part of the model. In the model, the method comes in handy to process the data to determine the probability of failure. Within the model, further work can be done based on this outcome to ultimately determine whether or not maintenance should be carried out before starting new production. The model is used to determine whether maintenance is needed, the method is there to serve as a basis for the model.

To choose from the above methods as base for the model of predictive maintenance, the first step is to determine what kind of data is available and what other requirements emerged from the earlier research.

The available data is historical data, consisting of condition monitoring data and process data from the past, there is also expert knowledge about the data and there is real-time condition monitoring data and information about the processes. The first data analysis shows that several variables influence the lifespan of the valves and pumps. In addition, there are some variables that indicate that a component is coming to the end of its lifetime. The failure data shows that there is not a single parameter responsible for failure, but that it is a combination of several factors. The method must therefore be able to take these relationships into account and display them clearly; in other words, the complexity must not become too great, otherwise it becomes unclear which variables have which influence. As soon as an unplanned downtime occurs, the condition monitoring data shows which component is responsible. Based on expert knowledge and historical data, it can also be determined which factors play what role in these kinds of downtimes.

In addition, the model must be trained on the historical data, so the method should also be capable of that. So that, when reading in the new condition monitoring data, it can first determine the current state of the component and then make a prognosis on the expected RUL, given that the future processes are known. Here, it is important that the method can make use of both healthy condition data and failure condition data. So that a complete picture emerges and all types of condition data can be used.

In short, the desired final model, with one of the methods described earlier as basis, must meet the following requirements:

- The model must be able to process both condition monitoring data and expert knowledge and be trainable based on the historical data.
- The relationships between the variables must be clearly expressed in the model; in other words, the model must represent the components in the line. The complexity should not become too high because then the relationships between the variables are no longer clearly visible.
- To make the model flexible for future adaptations, such as new relationships between variables or other types of data, it should not be computationally intensive.
- Because it works with different components within a line that each have their own failure mode, the model must be able to withstand multi-variate failure modes without sacrificing accuracy.

Using the descriptions from [Chapter 4.2.2](#) and [Appendix E: Table E1](#) [72], the methods can be evaluated against the criteria.

Criterion 1: The method must be able to work with expert knowledge.

Looking at [Appendix E: Table E1](#), it can be said that ARMA methods are then discarded, these are not suitable according to Sikorska et al [72] if expert data or historical data are present.

Criterion 2: The method should not become too complex with all kinds of underlying assumptions or other ways of bringing different variables together.

Looking at the complexity of the methods, using [Appendix E: Table E1](#) and the descriptions given earlier, it can be assumed that Kalman, Particle filters and PHM methods are too complex for this research. These work with all kinds of underlying functions where the variables affecting the failure modes have to be transformed into other kinds of data. In doing so, it is then no longer clear which factors have a direct influence on the failures of the various components.

Criterion 3: Calculation intensity must be low.

Because it should be a method in which it should be relatively easy to add relationships between variables, the method should not be computationally intensive. Markov methods then drop out because a separate model has to be created for each failure mode, making it very complex and these models and methods are therefore also computationally intensive, see [Appendix E: Table E1](#).

Criterion 4: The method should be able to work for multiple failure modes without sacrificing accuracy.

This allows trend extrapolation to be discarded, as this method works best when there is only one failure mode that can be derived from a single parameter. This is not the case in this research. In addition, the threshold for determining the RUL becomes unreliable if several failure modes have to be captured in the same model.

ARF is then also discarded. If too many failure modes have to be used, it is not possible to make reliable forecasts for the RUL.

Looking at the pros and cons and the descriptions, it can be concluded that the biggest difference in the methods is between using and not using condition monitoring data and clearly showing when what is used. RUL PDF does not use the times when failures have occurred but only looks at the times when no failures have occurred. Because the

failure data is essential in the desired model, it is better to choose BN. This is because the failure data contains a lot of information about which factors have an influence and can therefore be used well if a forecast has to be given based on the future processes.

Figure 28, which is a copy of Figure 27 [72], shows at which criterium which method was discarded.

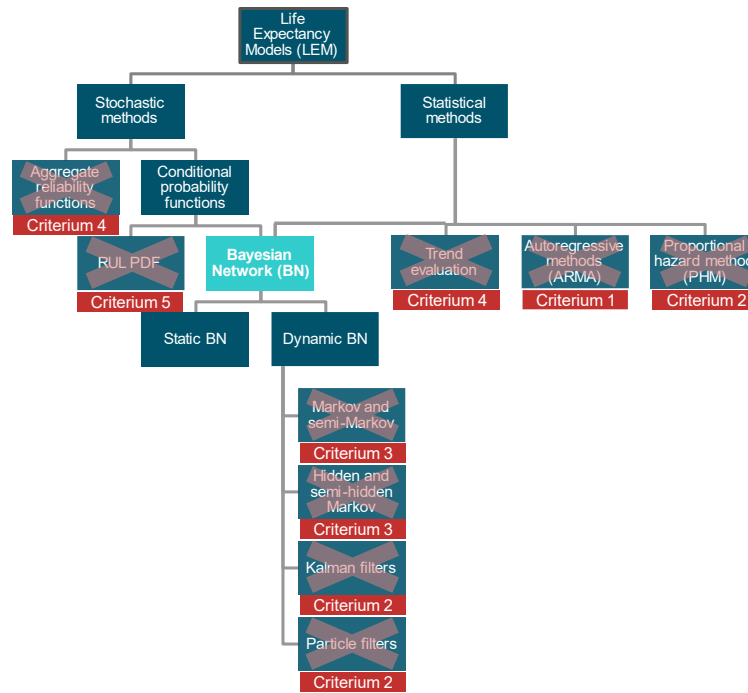


Figure 28: LEM overview [72], with criteria for selection

4.4 Bayesian Network

A BN is often used to simplify the representation of a complex model by visualising the relationships between different variables. A BN is a DAG (directed acyclic graph) that consists of nodes and arc. The nodes represent the variables within the system. The arcs are the connections between the nodes, indicating which nodes affect each other [125]. There are different types of nodes. Looking at the example in Figure 29 [126], nodes A and B are the parent nodes of child node C. Nodes A and B affect node C, which is shown by the arcs from A to C and B to C. Since there is no arc between node A and B, it can be assumed that these nodes are independent of each other. Node C in turn is the parent node of child node D, which can be seen by the arc between node C and D. Another name for nodes A and B can also be root nodes because the nodes mark the beginning of the BN. Node D can also be called leaf node because it is the last node in the BN [127]. Chapter 4.4.2 will go into more detail on the calculations for this BN, that serves as an example to give an idea of what calculations can be performed. This will later be used on a larger scale in the model.

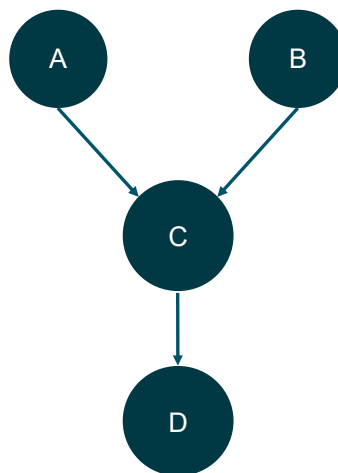


Figure 29: Example BN

A BN, defined by Jensen [128] as $\mathcal{N} = (\mathcal{X}, \mathcal{G}, \mathcal{P})$, consists of:

- A set of discrete random variables, \mathcal{X} , represented by the nodes of \mathcal{G}
- A DAG $\mathcal{G} = (V, E)$ with nodes $V = \{v_1, \dots, v_n\}$ and directed arcs E
- A set of conditional probability distributions, \mathcal{P} , containing one distribution, $P(X_v | X_{pa(v)})$, for each random variable $X_v \in \mathcal{X}$. The set of variables represented by the parents, $pa(v)$, of $v \in V$ in $\mathcal{G} = (V, E)$ is sometimes called the conditioning variables of X_v , the conditioned variable.

A BN encodes a joint probability distribution over a set of random variables, \mathcal{X} , of a problem domain. The set of conditional probability distributions, \mathcal{P} , specifies a multiplicative factorization of the joint probability distribution over \mathcal{X} as represented by the chain rule of BN [128], see Eq. 4.7:

$$P(\mathcal{X}) = \prod_{v \in V} P(X_v | X_{pa(v)})$$

Eq. 4.7

Looking at the example BN of Figure 29, then this BN can be written as followed:

Nodes: $V = \{A, B, C, D\}$

Set of directed arcs: $E = \{(A, C), (B, C), (C, D)\}$

The joint probability defined in Eq. 4.7 can be written as:

$$P(A, B, C, D) = P(A)P(B)P(C|A, B)P(D|C)$$

4.4.1 Probability theory

The underlying theory of probability will be further explained here, and the main assumptions and formulas will also be discussed.

Conditional probabilities

Conditional probabilities can be determined in addition to probabilities for individual events. This is noted as $P(A|B)$, which is the probability of A given that event B will occur [129]. The conditional probability can be calculated with Eq. 4.8.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Eq. 4.8

$P(B)$ is probability of event B
 $P(A|B)$ is probability of event A given that event B will occur
 $P(A \cap B)$ represents that both events A and B occur

Using the fundamental rule, it is possible to see how the probability of that both events A and B occur if the probability of A given B and the probability of B are known [129]. This will give the following Eq. 4.9:

$$P(A \cap B) = P(B|A)P(A)$$

Eq. 4.9

$P(A)$ is probability of event A
 $P(B|A)$ is probability of event B given that event A will occur
 $P(A \cap B)$ represents that both events A and B occur

Since it holds that $P(A \cap B) = P(B \cap A)$ it can be stated that, Eq. 4.10:

$$P(A|B)P(B) = P(A \cap B) = P(B|A)P(A)$$

Eq. 4.10

$P(A)$ is probability of event A
 $P(B)$ is probability of event B
 $P(A|B)$ is probability of event A given that event B will occur
 $P(B|A)$ is probability of event B given that event A will occur
 $P(A \cap B)$ represents that both events A and B occur

With this, Eq. 4.8 can then be converted to Eq. 4.11:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Eq. 4.11

$P(A)$	is prior probability of event A
$P(B)$	is marginal probability of event B
$P(A B)$	is posterior probability of event A given that event B will occur
$P(B A)$	is likelihood probability of event B given that event A will occur

Eq. 4.11 is known as the Bayes' rule [129] [130]. This is an important formula for the BN. Bayes' rule ensures that the belief about an event A can be updated if there is information about another event B.

In Eq. 4.11, $P(A)$ stands for the prior probability of event A. This is the probability formed about the probability of a possible outcome occurring without looking at data or other evidence.

$P(A|B)$ is the posterior probability of A given B. It is the updated probability of event A after new information has been considered. In this case, information about event B.

$P(B|A)$ is the likelihood probability. This is the probability of B given A.

$P(B)$ is the marginal probability of B. It is the probability of event B among all possible values of A. This can be done with the following Eq. 4.12:

$$P(B) = \sum_{i=0}^k P(B|A_i)P(A_i) \tag{Eq. 4.12}$$

$P(A_i)$	is probability of event A in state i
$P(B)$	is probability of event B
$P(B A_i)$	is probability of event B given that event A in state i will occur
i	indicates in which state A is

If the information about event B says nothing about the probability of A, then it is called independent events A and B. Eq. 4.13 should then apply:

$$P(A|B) = P(A) \tag{Eq. 4.13}$$

$P(A)$	is probability of event A
$P(A B)$	is probability of event A given that event B will occur

If Eq. 4.13 is put in the Bayes' rule [130], Eq. 4.11, it can be seen that, see Eq. 4.14:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B) \tag{Eq. 4.14}$$

If two events are independent then the fundamental rule is written as [129], Eq. 4.15:

$$P(A \cap B) = P(A|B)P(B) = P(A) \cdot P(B) \tag{Eq. 4.15}$$

However, events A and B can be conditional independent over a given event C. That is, the if information comes in about, say, event B, nothing changes about the belief of event A if there is already knowledge about event C.

If A and B are conditional independent given event C then Eq. 4.16 and Eq. 4.17 apply:

$$P(A|B \cap C) = P(A|C) \tag{Eq. 4.16}$$

$P(A B \cap C)$	is probability of event A given that event B and event C occur
$P(A C)$	is probability of event A given that event C occur

$$P(B|A \cap C) = P(B|C) \tag{Eq. 4.17}$$

$P(B A \cap C)$	is probability of event B given that event A and event C occur
$P(B C)$	is probability of event B given that event C occur

For the multiplication rule it will follow that, see Eq. 4.18:

$$P(A \cap B|C) = P(A|C) \cdot P(B|C) \tag{Eq. 4.18}$$

To apply the above explanation and formulas in a BN for this research, some conversions need to be made. Instead of events, it will be called variables. For each outcome, the variable has an associated state. Suppose there is a variable A then the corresponding set of states of A is; $sp(A) = (a_1, a_2, \dots, a_n)$. These states must be mutually exclusive and

exhaustive. By exhaustive is meant that the variable must be in one of the states (without knowing in which one). Each variable has a finite number of states [130]. In this research, uppercase letters (e.g. A) are used to indicate variables and lowercase letters (e.g. a) for states. [128]. In addition, there is a probability distribution of A, given as $P(A)$. This can be used to express the probability that node A is in a given state. The following Eq. 4.19 and Eq. 4.20 must hold:

$$P(A) = (a_1, \dots, a_n) \quad a_i \geq 0 \tag{Eq. 4.19}$$

$$\sum_{i=1}^n a_i = a_1 + \dots + a_n = 1 \tag{Eq. 4.20}$$

Eq.4.20 states that the probabilities associated with the possible states of node A must add up to 1. Suppose node A has two states and $P(a_1)$ is 0.65, then it follows from Eq. 4.20 that $P(a_2)$ must equal 0.35. It follows that $P(a_1) = 1 - P(a_2)$. If node A has only two states.

For each node in the BN, it is necessary to determine its states and the probabilities associated with them. Suppose node D in the example BN of Figure 29 has two possible states: d (true) and $\neg d$ (not true). Then $P(d)$ means the probability of node D being in the true state. $P(\neg d)$ is the probability that node D is in the not true state. These probabilities are complementary to each other because node D has only 2 possible states. Since the arcs in the BN indicate which nodes are connected, it is also possible to look at which nodes are independent and which nodes are dependent. In the example, see Figure 29, nodes A and B are conditional independent of node C, see Eq. 4.18. In a BN, each node must have a conditional probability table. This is the probability of a given state given the states of the parent nodes. For this, the formulas described earlier can be used. For root nodes A and B, the conditional probability will contain the marginal probabilities of the states, this is because there is no influence of parent nodes [130].

With a BN, it is possible to calculate the other probabilities based on evidence, i.e. information about a particular node being in a particular state. This can be done with the joint distribution of the entire BN [129]. In addition, it is possible to do forward or backward inference with a BN [131]. Forward inference is that there is evidence for certain causes and then it can be calculated what then are the expected consequences. This is also known as predictive or causal inference [132]. Backward inference has evidence that a particular consequence has been observed and then reasoning back to what the possible cause may have been. This is also known as diagnostic inference [133].

In the example, see Figure 29, then forward inference is when there is information about node A, for example, and it looks at what then impacts node D. Backward inference is when there is information about node D and it is reasoned back what might be the cause of that [133].

4.4.2 Example calculations

In this section, a generic example will be used to show how the above equations can be applied to a BN. The BN shown in Figure 29 will be used for this purpose. As indicated earlier, the lowercase letter serves to indicate that a node is in a particular state. In this example, each node can have 2 states; true and not true. True is indicated by the lowercase letter and not true is indicated by \neg in front of the lowercase letter. Values are assumed for the probabilities [134], see Table 4.1, this can be based on knowledge or on a dataset.

Table 4.1: Values of nodes for calculations example BN [134]

$P(d c)$	0.4
$P(d \neg c)$	0.1
$P(c a,b)$	0.3
$P(c \neg a,b)$	0.5
$P(c a, \neg b)$	0.7
$P(c \neg a, \neg b)$	0.9
$P(a)$	0.6
$P(b)$	0.2

As stated earlier, based on these probabilities, it is possible to determine the marginal probabilities of the nodes. Suppose $P(d)$, probability of node D is in the true (d) state, needs to be determined.

As shown in [Figure 29](#), node D depends only on node C. To determine P(d), it is necessary to marginalise over node D:

$$P(d) = P(d, c) + P(d, \neg c)$$

Next, node C must be eliminated to determine the marginal probability of P(d). This can be done by conditioning over C, see the next step in the calculation. Since node C is eliminated and both states (c and $\neg c$) of node C are used to do this, calculating P(d) can be written as [Eq. 4.21](#):

$$\begin{aligned} P(d) &= P(d, c) + P(d, \neg c) \\ &= P(d|c)P(c) + P(d|\neg c)P(\neg c) \\ &= \sum_c P(d|C)P(C) \end{aligned}$$

[Eq. 4.21](#)

To eliminate C and fill in [Eq. 4.21](#), P(c) and P($\neg c$) must be determined. However, only one of the two probabilities needs to be calculated because the probabilities are complementary to each other so that $P(\neg c) = 1 - P(c)$. Node C has parent nodes A and B. Here again, the uppercase letters indicate that both states of node A and node B are used. Node A has states a and $\neg a$. Node B has states b and $\neg b$. Both nodes must be eliminated to ensure that P(c) can be calculated. Once P(c) is calculated, P($\neg c$) can also be calculated. By entering this in [Eq. 4.21](#), it can be determined what P(d) is. But first [Eq. 4.22](#) applies:

$$P(c) = \sum_{A,B} P(c, A, B) = \sum_{A,B} P(c|A, B)P(A, B) = \sum_{A,B} P(c|A, B)P(A)P(B)$$

[Eq. 4.22](#)

P(A,B) may be written as P(A)P(B) if nodes A and B are independent.

Filling in [Eq. 4.22](#) with the values of [Table 4.1](#) gives:

$$\begin{aligned} P(c) &= \sum_{A,B} P(c|A, B)P(A)P(B) \\ &= P(c|a, b)P(a)P(b) + P(c|\neg a, b)P(\neg a)P(b) + P(c|a, \neg b)P(a)P(\neg b) + P(c|\neg a, \neg b)P(\neg a)P(\neg b) \\ &= (0.3 * 0.6 * 0.2) + (0.5 * (1 - 0.6) * 0.2) + (0.7 * 0.6 * (1 - 0.2)) + (0.9 * (1 - 0.6) * (1 - 0.2)) = 0.70 \end{aligned}$$

Now that it is known what P(c) is, it is also known what P($\neg c$) is, namely 1-P(c).

This can be used to calculate P(d) by filling in [Eq. 4.21](#) with the values of [Table 4.1](#):

$$P(d) = \sum_c P(d|C)P(C) = 0.4 * 0.7 + 0.1 * (1 - 0.7) = 0.31$$

Inference example

Now that the probabilities of P(d) and P(c) are known, it can be looked at how to work with inferences. Suppose it is given that P(d) is true, this is the evidence, and now, based on this information, the probability that P(b) is true must be found out.

For this, Bayes' rule, like [Eq. 4.11](#), is then used to determine the probability of P(b|d), or the posterior probability of b, see [Eq. 4.23](#).

$$P(b|d) = \frac{P(d|b)P(b)}{P(d)} \quad \text{Bayes' rule}$$

$$P(b|d) = \frac{P(d|b)P(b)}{P(d)} = \frac{\sum_c P(d|C) \sum_A P(C|A, b)P(A)P(b)}{\sum_c P(d|C) \sum_{A,B} P(C|A, B)P(A)P(B)}$$

[Eq. 4.23](#)

Below the partial line is actually P(d) as calculated earlier, $P(d) = 0.31$.

Above the partial line, the following is calculated, by using the values of [Table 4.1](#):

$$\sum_A P(C|A, b)P(A) = (P(C|a, b)P(a) + P(C|\neg a, b)P(\neg a)) = ((0.3 * 0.6) + (0.5 * (1 - 0.6))) = 0.38$$

$$\sum_C P(d|C) \sum_A P(C|A, b) P(A) P(b) = (0.4 * 0.38 * 0.2) + (0.1 * (1 - 0.38) * 0.2) = 0.0428$$

Entering these values in Eq. 4.23 then gives:

$$P(b|d) = \frac{P(d|b)P(b)}{P(d)} = \frac{\sum_C P(d|C) \sum_A P(C|A, b) P(A) P(b)}{\sum_C P(d|C) \sum_{A,B} P(C|A, B) P(A) P(B)} = \frac{0.0428}{0.31} = 0.138$$

So the probability that b is true given d is true ($P(b|d)$) is 0.138.

This is how a BN can be used to update probabilities when information about another node is known. The above example gives an idea of how a BN should be constructed and what calculations can be performed with it. This is a general example, in Chapter 5 the BN specifically for this research will be discussed.

4.5 Conclusion

This chapter considered the sub-research question: *Based on the available data and knowledge, what kind of model is most appropriate to develop a model with for this problem?*

To answer this question, a data analysis was performed, complemented by expert knowledge on what the data say. For valves, the behaviour of the looptijd parameter and how it affects the state of the valve was looked at. Based on the behaviour, it can be determined that a logistic function is needed to relate the failure probability to the looptijd. In addition, for each valve, the number of times each valve failed in the given period was considered.

Insufficient data were available for the pumps. The choice was made to construct the pump curves based on the data from the manufacturers of the pumps and underlying pump theories. With these curves, it is possible to determine the BEP of the pump and then, based on assumptions, determine the reliability curve of each pump. By relating this curve to the MTBF of the pump, the failure probability can be determined.

Next, it examined which method best suited the available data and knowledge to create a model. This showed that a BN best suited the combination of data and expert knowledge in this study. The terms and calculations associated with a BN have been briefly discussed. Chapter 5 elaborates on how the method contributes to the model for this research.

5. Method for the model

This chapter looks at the sub-research question: *What are the steps to develop and verify the method of the model?* To answer this sub-research question, the structure of the BN is first considered. It then looks at the nodes to be included in the BN and the relationships between them. This is followed by the corresponding CPTs. Finally, the model needs to be verified.

5.1 Multi-layer structure

Within this research, a BN will be made in which there will be several layers in the network. In the literature, this is called a multi-layer structure model with a bottom-up approach by Verbert et al [135], as shown in Figure 30. It is an approach where at the component level, for each individual component, the output, in this case the probability of failure of the component, is considered. In the next layer, one can determine what the forecast is of the potential failure of a part of the line given the component's probabilities of failure. Finally, in the highest layer, all forecasts come together to arrive at a system output.

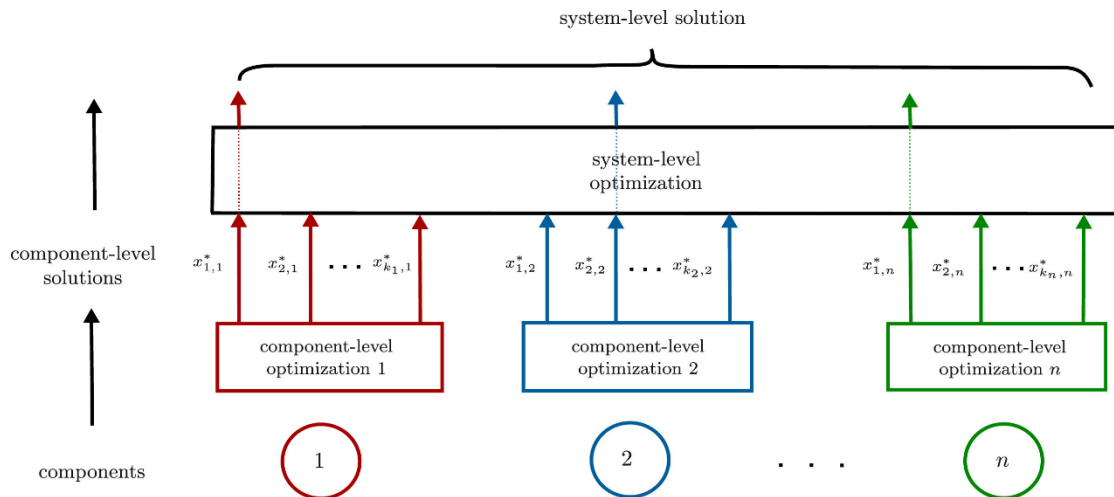


Figure 30: Multi-layer structure [135]

5.2 Bayesian Network

Looking at the roadmap for drafting a BN, the variables that belong in the BN were considered. Then, using the knowledge present in the company, it was determined which variables affect each other. This allows the arcs between the nodes to be created. Next, there is the step of setting up the conditional probabilities of each node. To do this, it is necessary to know how many states each node can assume and what the independent nodes are.

For this research, it was decided to divide the line into 5 line segments. This makes the calculations and indications clearer. Line segment 1 (LSI) has two valves (V1.1 and V1.2) and one pump (P1.1). Line segment 2 (LSII) has two valves (V2.1 and V2.2) and two pumps (P2.1 and P2.2). Line segment 3 (LSIII) has 8 valves (V3.1, V3.2, V3.3, V3.4, V3.5, V3.6, V3.7 and V3.8) and 1 pump (P3.1). Line segment 4 (LSIV) has 5 valves (V4.1, V4.2, V4.3, V4.4 and V4.5) and 1 pump (P5.1). Line segment 5 (LSV) has 4 valves (V5.1, V5.2, V5.3 and V5.4) and 1 pump (P5.1).

The first layer contains the parameters that have a relation to the failure probability of the valves and pumps in the line. In the case of the valves, this is the "looptijd" (LT) and in the case of the pumps, it is the parameter B and parameter G.

The second layer then contains the nodes that represent the condition of the valves and pumps, this is the probability of failure given the parameters of the parent nodes.

Then there is a layer where all component nodes come together in the child node of line segment. Here the probability of failure of the line segment given the probability of failure of the individual components can be determined.

Finally, there is a child node in which all line segments come together with which the probability of failure of the entire line can be determined. This can then also be used to determine the probability of downtime.

Figure 31 shows the schematic overview of the BN to be used in this research. Appendix F shows the complete BN and also a list of all the symbols used in the BN.

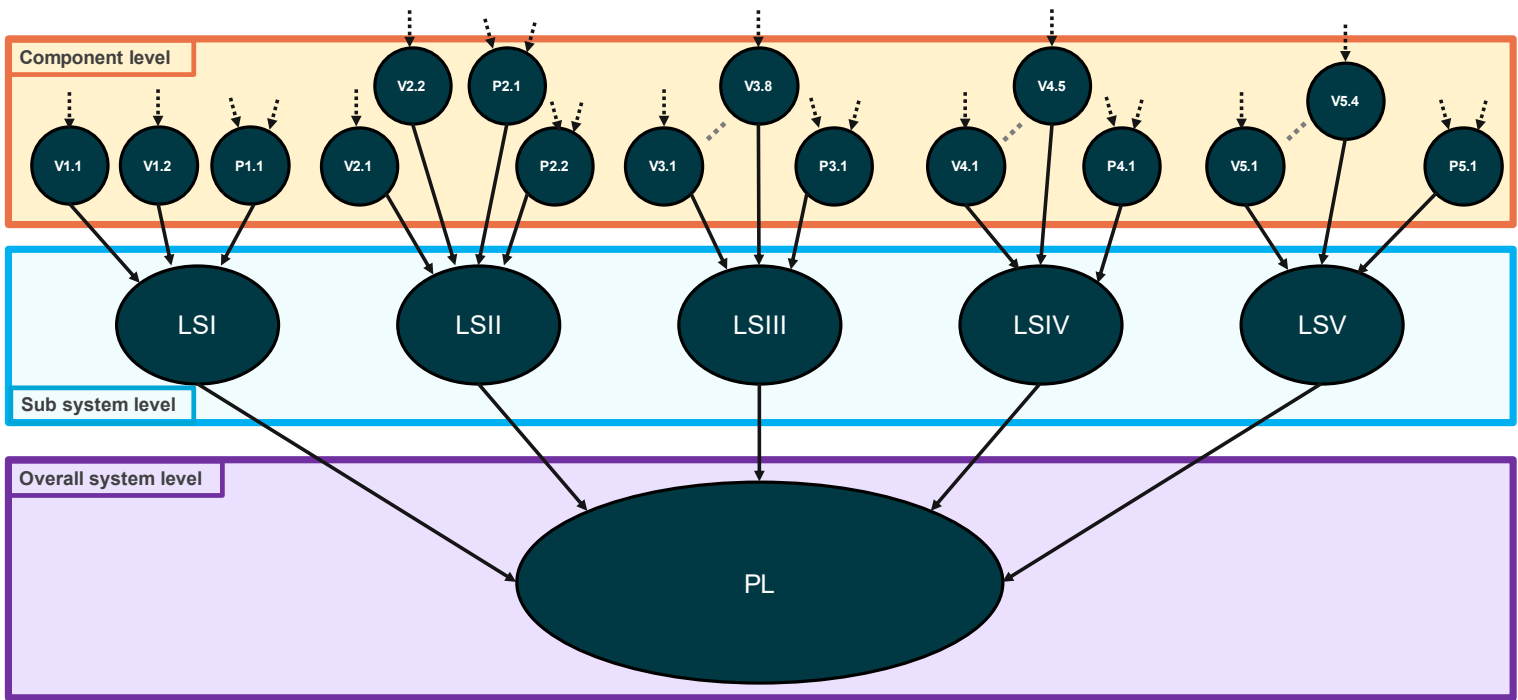


Figure 31: BN with multi-layer structure

The BN can thus be divided into the component level, subsystem level and overall system level.

The number of states of each node and its relationship with the other nodes is determined.

Now that it is known which nodes are linked together, the mathematical underpinnings and how the conditional probability tables are set up can be looked at. The equations of Chapter 4.4 will be used.

The part of the BN of line segment 1 will be used to explain the mathematical background, see Figure 32. In line segment there are two valves (V1.1 and V1.2) and one pump (P1.1).

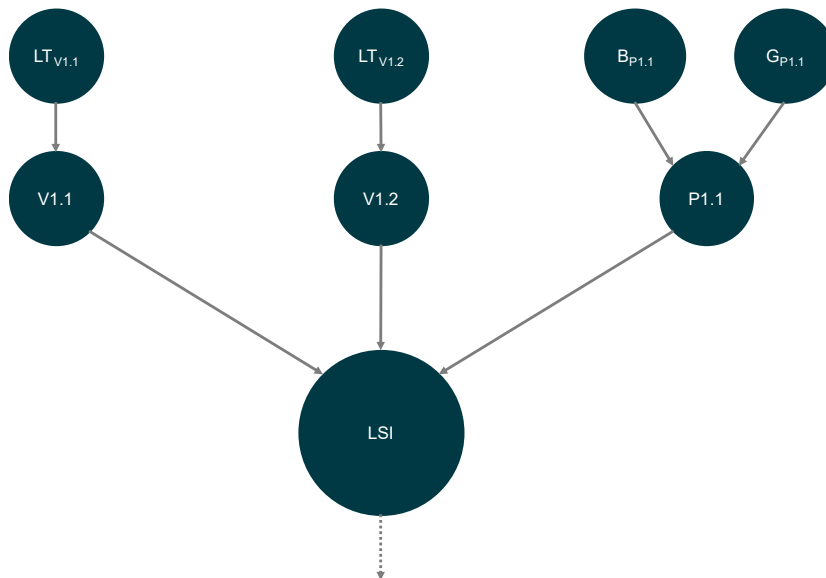


Figure 32: Line segment 1, part of BN

Starting with the upper nodes. These are the variables that affect the probability of valve and pump failure. In the case of the valves, this is the looptijd, the time how long it takes for a valve to open or close. For the pump, these are two variables, B and G, related to reliability and cavitation.

Valves

The variable "Looptijd" ($LT_{Vx,x}$) has 15, 20, 25, 30 or 50 possible states, it depends on the specific valve [130][136]. The states can be written as, if "looptijd" has 15 states:

$$\Omega_{LT_{Vx,x}} = \{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15\}$$

Or if "looptijd" has 20 possible states:

$$\Omega_{LT_{Vx,x}} = \{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20\}$$

The same thing can be done for all the different looptijd nodes (so it can also be done for the valves with looptijd 25 or 30 or 50 states, it depends on the maximum looptijd of the specific valve) to give the specific amount of states for each unique valve.

x,x should be replaced, the first x indicates the number of the specific line segment and the second x is the number of the valve in that specific line segment. For example $LT_{V1,2}$ indicated the looptijd of the second valve in line segment 1.

In this BN, the LT is a root node and there is no parent node to influence it.

It must apply that all individual probabilities of a given state of $LT_{Vx,x}$ added together must equal 1, as earlier explained with Eq. 4.19 and Eq. 4.20 [129], see Eq. 5.1 for the $LT_{Vx,x}$ nodes:

$$\sum_{i=1}^k P(LT_{Vx,x} = i) = P(LT_{Vx,x} = 1) + \dots + P(LT_{Vx,x} = k) = 1$$

Eq. 5.1

k depends on the amount of possible states of looptijd, this varies between 15, 20, 25, 30 or 50 states. The amount of states is defined for every valve in the network.

Pumps

For the pumps, two parameters are defined that indicate the probability of failure of the pump. The parameter B is related to the BEP. This is the deviation from the BEP at which the pump operates. Parameter G is for determining whether there is a chance of cavitation. Both parameters can be determined as described earlier in Chapter 4.1.3.3.

For parameter $B_{Px,x}$, 5 states can be determined, see Figure 33 [96]:

State 1: Pump performs at BEP

State 2: Pump performs at -10% of BEP or +5% of BEP

State 3: Pump performs between -20% and -10% or between +10% and +5% of BEP

State 4: Pump performs between -30% and -20% or between +15% and +10% of BEP

State 5: Pump performs at more than -30% or +15% of BEP

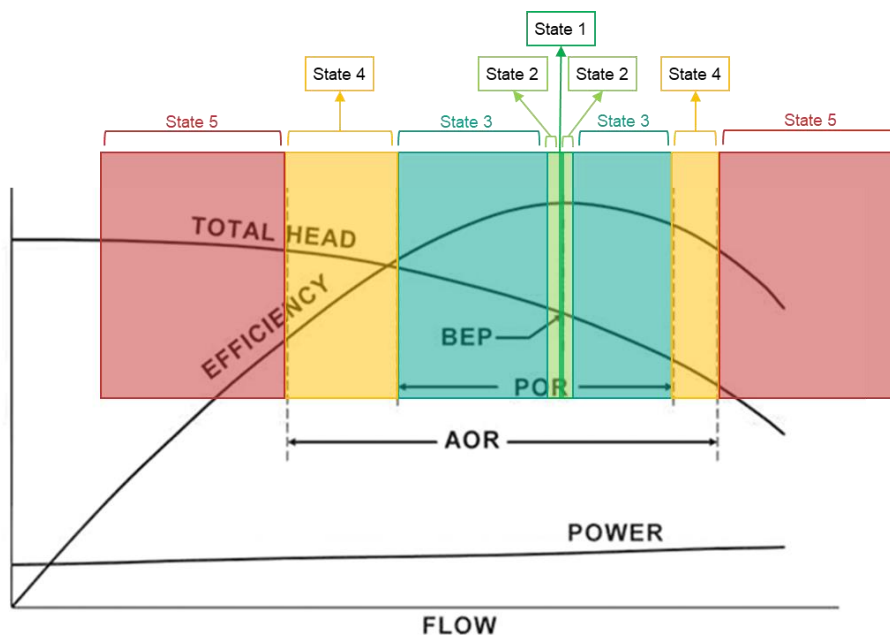


Figure 33: States of $B_{Px,x}$

Variable $B_{P_{x,x}}$ has 5 possible states.

$$\Omega_{B_{P_{x,x}}} = \{1,2,3,4,5\}$$

The x,x can be changed; the first x is for the line segment number and the second x is for the number of the pump in that specific line part. For example, $B_{P_{3,1}}$ stands for the first pump in line segment 3.

Following on from [Eq. 4.19](#) and [Eq. 4.20](#), there must be stated that, see [Eq. 5.2](#):

$$\sum_{i=1}^5 P(B_{P_{x,x}} = i) = P(B_{P_{x,x}} = 1) + \dots + P(B_{P_{x,x}} = 5) = 1$$

[Eq. 5.2](#)

The second parameter, $G_{P_{x,x}}$, is to determine whether cavitation occurs. This can be expressed in 2 states.

State 1: No cavitation, NPSH margin ≥ 1.1 (see [Eq. 4.4](#) and [Figure 22](#))

State 2: Cavitation, NPSH margin < 1.1 (see [Eq. 4.4](#) and [Figure 22](#))

Variable $G_{P_{x,x}}$ has 2 possible states.

$$\Omega_{G_{P_{x,x}}} = \{1,2\}$$

The x,x can be changed; the first x is for the line segment number and the second x is for the number of the pump in that specific line part. For example, $G_{P_{3,1}}$ stands for the first pump in line segment 3.

Following on from [Eq. 4.19](#) and [Eq. 4.20](#), there must be stated that, see [Eq. 5.3](#).

$$\sum_{i=1}^2 P(G_{P_{x,x}} = i) = P(G_{P_{x,x}} = 1) + P(G_{P_{x,x}} = 2) = 1$$

[Eq. 5.3](#)

5.2.1 Component nodes

Then, the nodes representing the probability of failure of each component given the parent nodes can now be looked at. In the example, these are nodes $V_{1.1}$, $V_{1.2}$ and $P_{1.1}$, see [Figure 32](#).

V1.1

Node $V_{1.1}$ has 1 parent node $LT_{V_{1.1}}$. Node $V_{1.1}$ represent the probability of failure of valve 1.1, it has 2 possible states:

$$\Omega_{V_{1.1}} = \{low, high\} = \{0,1\}$$

As stated earlier in [Chapter 4.1.2](#), a logistic function can be constructed for each valve to determine the probability of failure based on the looptijd of the valve. For this purpose, see [Eq. 4.1](#).

The value for m can be calculated by adding up all times of maturity and then taking the average.

The value for k can be calculated using expert knowledge. It has been explained that at the time the looptijd has a value of 15 seconds that it is assumed 99% of the valve fails. Since the value of m has already been calculated earlier in [Chapter 4.1.2](#), by combining this knowledge, the value of k can be determined. It is assumed for this example that $m=8$ seconds. With this, [Eq. 4.1](#) can be filled in:

$$0.99 = \frac{1}{(1 + e^{(-k*(15-8)})}$$

This rewrite gives:

$$e^{-7k} = \frac{1}{0.99} - 1$$

$$-7k = \ln(0.010101)$$

$$k = \frac{\ln(0.010101)}{-7} = 0.6564$$

Filling the k into [Eq. 4.1](#) gives:

$$probability\ of\ failure = \frac{1}{(1 + e^{(-0.6564*(looptijd-m)})}$$

[Eq. 5.4](#)

[Eq. 5.4](#) will provide a curve, as can be seen in [Figure 34](#).

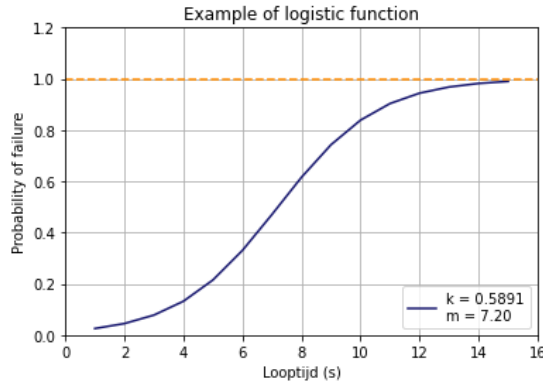


Figure 34: Example of logistic function

From the curve in Figure 34, it can be determined what the probability of failure is given a given looptijd. Since node V1.1 has only 2 possible states, this can also be used to determine the probability of not failing.

For the local conditional probability of the V1.1 node, using Eq. 4.7, it can be stated that Eq. 5.5:

$$P(LT_{V1.1}, V_{1.1}) = P(LT_{V1.1})P(V_{1.1}|LT_{V1.1})$$

Eq. 5.5

V1.2

The same thing can be done for node V1.2, see Eq. 5.6, keep in mind to change the average looptijd (Eq. 5.4).

$$P(LT_{V1.2}, V_{1.2}) = P(LT_{V1.2})P(V_{1.2}|LT_{V1.2})$$

Eq. 5.6

P1.1

Node P1.1 represents the probability of failure of pump 1.1 also this node has 2 possible states.

$$\Omega_{P_{x,x}} = \{low, high\} = \{0,1\}$$

Node P1.1 has two parent nodes: B_{P1.1} and G_{P1.1}.

Depending on the values obtained from parameters B and G, the reliability curve as explained in Chapter 4.1.3 can be used. This determines the probability of failure of the pump given the condition of the parent nodes.

The local conditional probability if the P1.1 node, using Eq. 4.7, can be written as Eq. 5.7:

$$P(P_{1.1}, B_{P1.1}, G_{P1.1}) = P(G_{P1.1})P(B_{P1.1})P(P_{1.1}|B_{P1.1}, G_{P1.1})$$

Eq. 5.7

5.2.2 Subsystem level nodes

Line segment 1 (LS_i)

Next, the node LS_i (line segment 1) can be looked at. This is the child node of V1.1, V1.2 and P1.1, see Figure 32.

The node LS_i represents the probability of failure of line segment 1. The joint probability can be written using the local conditional probabilities of the other nodes, just like Eq. 4.7 [130]. This will give Eq. 5.8.

$$\begin{aligned} P(LS_i, V_{1.1}, V_{1.2}, P_{1.1}, LT_{V1.1}, LT_{V1.2}, B_{P1.1}, G_{P1.1}) \\ = P(G_{P1.1}) \cdot P(B_{P1.1}) \cdot P(LT_{V1.1}) \cdot P(LT_{V1.2}) \cdot P(P_{1.1}|B_{P1.1}, G_{P1.1}) \cdot P(V_{1.1}|LT_{V1.1}) \cdot P(V_{1.2}|LT_{V1.2}) \\ \cdot P(LS_i|V_{1.1}, V_{1.2}, P_{1.1}) \end{aligned}$$

Eq. 5.8

Hereafter, the notation JP₁ will be used to refer to Eq. 5.8.

This can be done for each defined line segment of the production line. This determines the probability of failure of each line segment separately.

Line segment 2 (LS_{II}):

Line segment 2 consists of 2 valves and 2 pumps. The local conditional probabilities are determined in the same way as for line segment 1.

This gives the following for the joint probability of line segment 2, by filling in Eq. 4.7, see Eq. 5.9:

$$\begin{aligned} P(LS_{II}, V_{2.1}, V_{2.2}, P_{2.1}, P_{2.2}, LT_{V_{2.1}}, LT_{V_{2.2}}, B_{P_{2.1}}, G_{P_{2.1}}, B_{P_{2.2}}, G_{P_{2.2}}) \\ = P(G_{P_{2.1}}) \cdot P(B_{P_{2.1}}) \cdot P(G_{P_{2.2}}) \cdot P(B_{P_{2.2}}) \cdot P(LT_{V_{2.1}}) \cdot P(LT_{V_{2.2}}) \cdot P(P_{2.1} | B_{P_{2.1}}, G_{P_{2.1}}) \\ \cdot P(P_{2.2} | B_{P_{2.2}}, G_{P_{2.2}}) \cdot P(V_{2.1} | LT_{V_{2.1}}) \cdot P(V_{2.2} | LT_{V_{2.2}}) \cdot P(LS_{II} | V_{2.1}, V_{2.2}, P_{2.1}, P_{2.2}) \end{aligned}$$

Eq. 5.9

Hereafter, the notation JP₂ will be used to refer to Eq. 5.9.

Line segment 3 (LS_{III}):

Line segment 3 consists of 8 valves and 1 pump. The local conditional probabilities are determined in the same way as for line segment 1.

This gives the following for the joint probability of line segment 3, by filling in Eq. 4.7, see Eq. 5.10:

$$\begin{aligned} P(LS_{III}, V_{3.1}, V_{3.2}, V_{3.3}, V_{3.4}, V_{3.5}, V_{3.6}, V_{3.7}, V_{3.8}, \\ P_{3.1}, LT_{V_{3.1}}, LT_{V_{3.2}}, LT_{V_{3.3}}, LT_{V_{3.4}}, LT_{V_{3.5}}, LT_{V_{3.6}}, LT_{V_{3.7}}, LT_{V_{3.8}}, B_{P_{3.1}}, G_{P_{3.1}}) \\ = P(G_{P_{3.1}}) \cdot P(B_{P_{3.1}}) \cdot P(LT_{V_{3.1}}) \cdot P(LT_{V_{3.2}}) \cdot P(LT_{V_{3.3}}) \cdot P(LT_{V_{3.4}}) \cdot P(LT_{V_{3.5}}) \cdot P(LT_{V_{3.6}}) \cdot P(LT_{V_{3.7}}) \\ \cdot P(LT_{V_{3.8}}) \cdot P(P_{3.1} | B_{P_{3.1}}, G_{P_{3.1}}) \cdot P(V_{3.1} | LT_{V_{3.1}}) \cdot P(V_{3.2} | LT_{V_{3.2}}) \cdot P(V_{3.3} | LT_{V_{3.3}}) \cdot P(V_{3.4} | LT_{V_{3.4}}) \\ \cdot P(V_{3.5} | LT_{V_{3.5}}) \cdot P(V_{3.6} | LT_{V_{3.6}}) \cdot P(V_{3.7} | LT_{V_{3.7}}) \cdot P(V_{3.8} | LT_{V_{3.8}}) \\ \cdot P(LS_{III} | V_{3.1}, V_{3.2}, V_{3.3}, V_{3.4}, V_{3.5}, V_{3.6}, V_{3.7}, V_{3.8}, P_{3.1}) \end{aligned}$$

Eq. 5.10

Hereafter, the notation JP₃ will be used to refer to Eq. 5.10.

Line segment 4 (LS_{IV}):

Line segment 4 consists of 5 valves and 1 pump. The local conditional probabilities are determined in the same way as for line segment 1.

This gives the following for the joint probability of line segment 4, by filling in Eq. 4.7, see Eq. 5.11:

$$\begin{aligned} P(LS_{IV}, V_{4.1}, V_{4.2}, V_{4.3}, V_{4.4}, V_{4.5}, P_{4.1}, LT_{V_{4.1}}, LT_{V_{4.2}}, LT_{V_{4.3}}, LT_{V_{4.4}}, LT_{V_{4.5}}, B_{P_{4.1}}, G_{P_{4.1}}) \\ = P(G_{P_{4.1}}) \cdot P(B_{P_{4.1}}) \cdot P(LT_{V_{4.1}}) \cdot P(LT_{V_{4.2}}) \cdot P(LT_{V_{4.3}}) \cdot P(LT_{V_{4.4}}) \cdot P(LT_{V_{4.5}}) \cdot P(P_{4.1} | B_{P_{4.1}}, G_{P_{4.1}}) \\ \cdot P(V_{4.1} | LT_{V_{4.1}}) \cdot P(V_{4.2} | LT_{V_{4.2}}) \cdot P(V_{4.3} | LT_{V_{4.3}}) \cdot P(V_{4.4} | LT_{V_{4.4}}) \cdot P(V_{4.5} | LT_{V_{4.5}}) \\ \cdot P(LS_{IV} | V_{4.1}, V_{4.2}, V_{4.3}, V_{4.4}, V_{4.5}, P_{4.1}) \end{aligned}$$

Eq. 5.11

Hereafter, the notation JP₄ will be used to refer to Eq. 5.11.

Line segment 5 (LS_V):

Line segment 5 consists of 4 valves and 1 pump. The local conditional probabilities are determined in the same way as for line segment 1.

This gives the following for the joint probability of line segment 5, by filling in Eq. 4.7, see Eq. 5.12:

$$\begin{aligned} P(LS_V, V_{5.1}, V_{5.2}, V_{5.3}, V_{5.4}, P_{5.1}, LT_{V_{5.1}}, LT_{V_{5.2}}, LT_{V_{5.3}}, LT_{V_{5.4}}, B_{P_{5.1}}, G_{P_{5.1}}) \\ = P(G_{P_{5.1}}) \cdot P(B_{P_{5.1}}) \cdot P(LT_{V_{5.1}}) \cdot P(LT_{V_{5.2}}) \cdot P(LT_{V_{5.3}}) \cdot P(LT_{V_{5.4}}) \cdot P(P_{5.1} | B_{P_{5.1}}, G_{P_{5.1}}) \cdot P(V_{5.1} | LT_{V_{5.1}}) \\ \cdot P(V_{5.2} | LT_{V_{5.2}}) \cdot P(V_{5.3} | LT_{V_{5.3}}) \cdot P(V_{5.4} | LT_{V_{5.4}}) \cdot P(LS_V | V_{5.1}, V_{5.2}, V_{5.3}, V_{5.4}, P_{5.1}) \end{aligned}$$

Eq. 5.12

Hereafter, the notation JP₅ will be used to refer to Eq. 5.12.

5.2.3 Overall system level node

PL

Finally, there is the leaf node. This is the node PL that indicates the probability of failure of the entire production line. This is a child node of LS_I, LS_{II}, LS_{III}, LS_{IV} and LS_V, see Figure 35.

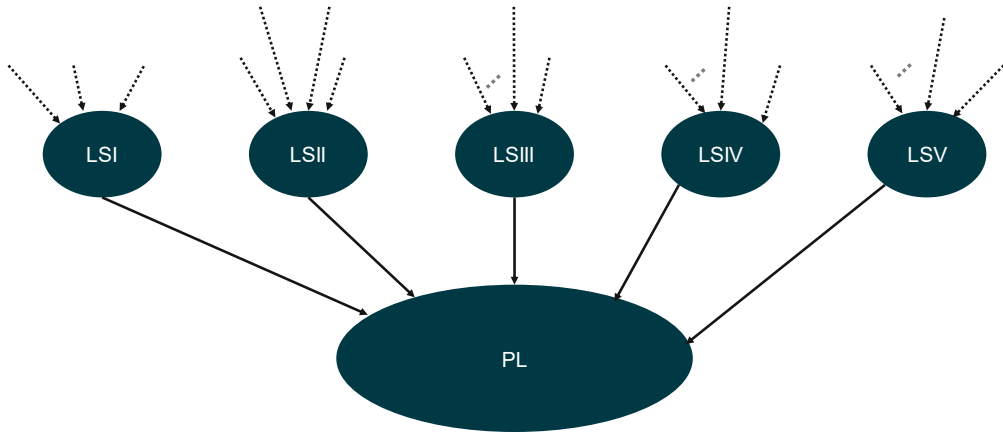


Figure 35: Last section of BN

The joint probability of the whole BN can be written as, Eq. 5.13:

$$P(PL, LS_I, LS_{II}, LS_{III}, LS_{IV}, LS_V, \dots \dots \dots, G_{P5.1}) = P(PL | LS_I, LS_{II}, LS_{III}, LS_{IV}, LS_V) \cdot JP_1 \cdot JP_2 \cdot JP_3 \cdot JP_4 \cdot JP_5$$

Eq. 5.13

Eq. 5.13 can be written as Eq. 5.14:

LS_i	$i = I, II, III, IV, V$
V_j	$j = 1.1, 1.2, 2.1, 2.2, 3.1, \dots, 3.8, 4.1, \dots, 4.5, 5.1, \dots, 5.4$
LT_{Vj}	$j = 1.1, 1.2, 2.1, 2.2, 3.1, \dots, 3.8, 4.1, \dots, 4.5, 5.1, \dots, 5.4$
P_k	$k = 1.1, 2.1, 2.2, 3.1, 4.1, 5.1$
B_{Pk}	$k = 1.1, 2.1, 2.2, 3.1, 4.1, 5.1$
G_{Pk}	$k = 1.1, 2.1, 2.2, 3.1, 4.1, 5.1$

$$\begin{aligned}
 P(PL, \{LS_i\}, \{V_j\}, \{P_k\}, \{LT_{Vj}\}, \{B_{Pk}\}, \{G_{Pk}\}) = & \\
 & P(PL | \{LS_i\}) \cdot \\
 & P(LS_I | V_{1.1}, V_{1.2}, P_{1.1}) \cdot P(P_{1.1} | B_{P1.1}, G_{P1.1}) \cdot \left(\prod_{m=1}^2 P(V_{1.m} | LT_{V1.m}) \right) \cdot \\
 & P(LS_{II} | V_{2.1}, V_{2.2}, P_{2.1}, P_{2.2}) \cdot \left(\prod_{n=1}^2 P(P_{2.n} | B_{P2.n}, G_{P2.n}) \right) \cdot \left(\prod_{q=1}^2 P(V_{2.q} | LT_{V2.q}) \right) \cdot \\
 & P(LS_{III} | V_{3.1}, \dots, V_{3.8}, P_{3.1}) \cdot P(P_{3.1} | B_{P3.1}, G_{P3.1}) \cdot \left(\prod_{r=1}^8 P(V_{3.r} | LT_{V3.r}) \right) \cdot \\
 & P(LS_{IV} | V_{4.1}, \dots, V_{4.5}, P_{4.1}) \cdot P(P_{4.1} | B_{P4.1}, G_{P4.1}) \cdot \left(\prod_{t=1}^5 P(V_{4.t} | LT_{V4.t}) \right) \cdot \\
 & P(LS_V | V_{5.1}, \dots, V_{5.4}, P_{5.1}) \cdot P(P_{5.1} | B_{P5.1}, G_{P5.1}) \cdot \left(\prod_{w=1}^4 P(V_{5.w} | LT_{V5.w}) \right) \cdot \\
 & \left(\prod_{k=1.1}^{5.1} (P(B_{Pk})P(G_{Pk})) \right) \cdot \left(\prod_{j=1.1}^{5.4} (P(LT_{Vj})) \right)
 \end{aligned}$$

Eq. 5.14

5.3 Conditional probability tables

Before inferences can be performed with the model, probabilities have to be determined. These must be included in the model which is programmed in Python using the PGMPY package [137].

For the looptijd ($LT_{V_{x,x}}$) nodes, it was assumed that there is a uniform distribution on the probability of occurrence of a given looptijd. All the CPTs can be found in [Appendix G](#).

For every valve ($V_{x,x}$ nodes), as shown in [Chapter 5.2](#), a unique logistic function representing the probability of failure of the valve was determined. Based on this, the CPT was constructed for each $V_{x,x}$ node. All the CPTs can be found in [Appendix G](#).

For the $G_{P_{x,x}}$ nodes, a uniform distribution was assumed because there is no knowledge of whether a particular state occurs more frequently.

For the $B_{P_{x,x}}$ nodes, consideration was given to what is logical for the occurrence of certain states. [Chapter 4.1.3](#) described that most pumps operate in state 3 or 4, hence the probabilities are higher. It is further assumed that the pumps hardly ever operate in state 1.

For the $P_{x,x}$ nodes, in terms of the influence of the $B_{P_{x,x}}$ nodes, the reliability function as explained in [Chapter 4.1.3](#) has been considered. It is assumed to be the probability of failure given the state of $B_{P_{x,x}}$ and that the state of node $G_{P_{x,x}}$ is in the lowest state. For the influence of node $G_{P_{x,x}}$ on node $P_{x,x}$, it is assumed that the probability of failure increases as $B_{P_{x,x}}$ also deteriorates. The highest probability of failure is when both node $G_{P_{x,x}}$ and $B_{P_{x,x}}$ are in their highest state.

For the LS nodes, assumptions have been made about what the probability of failure is given the condition of a particular component. It was assumed here that if nothing is in bad condition then there is a 1% chance of failure of the line segment. If 1 component fails then the probability of failure is already higher than the probability of not failing. This increases to the point where all components are in a poor condition, here the probability of failure then equals 99%. A 1% margin was chosen because there can always be a failure in the PLCs that results in wrong values being issued.

For the PL node, assumptions have been made about the probability of failure of the production line given the states of the different line segments. Here, it has been assumed that there is a 10% probability of failure if all line segments are in state 0.

Further, the probabilities of failure increase if several line segments are in state 1.

The above BN including all nodes and tables is programmed in Python with the PGMPY package [137]. This makes it possible to perform inferences and other calculations. However, validation and verification of the BN should be considered first.

5.4 Forward inference

As stated earlier, the BN can be used to perform inferences. In this research, only forward inference is used. This involves entering values for the $LT_{V_{x,x}}$, $B_{P_{x,x}}$ and $G_{P_{x,x}}$ nodes as evidence to ultimately determine the probability of failure of the PL node. In other words, it is determined: $P(PL | \{LT_{V_j}\}, \{B_{P_k}\}, \{G_{P_k}\})$ with $k = 1.1, 2.1, 2.2, 3.1, 4.1, 5.1$ and $j = 1.1, 1.2, 2.1, 2.2, 3.1, \dots, 3.8, 4.1, \dots, 4.5, 5.1, \dots, 5.4$.

Here, the previously defined relations between all nodes in [Chapter 5.2.1](#), [Chapter 5.2.2](#) and [Chapter 5.2.3](#) are used in addition to the Bayes rule, as explained in [Chapter 4.4](#) with [Eq. 4.11](#). Furthermore, [Eq. 5.14](#) is used.

As explained earlier in the example of [Chapter 4.4.2](#), node elimination is also used here, this is performed on the component level nodes ($V_{x,x}$ and $P_{x,x}$ nodes) and the sub system nodes (LS_x nodes). This makes it possible to perform forward inference to the PL node from the $LT_{V_{x,x}}$, $B_{P_{x,x}}$ and $G_{P_{x,x}}$ nodes. All is done by using the PGMPY package, which includes a module to perform inference using the Bayes rule.

5.5 BN verification

Now that the BN has been built, it is necessary to look at the verification of the BN. Verification is determining whether the BN accurately reflects the specifications and descriptions previously given [138] [139].

In this research, two different verification methods are used. First, some checks are performed to determine whether the nodes and the relationships between them are correctly defined. The second way to verify is with a sensitivity analysis. It also examines whether the outcomes of the BN in known situations match expectations.

5.5.1 Verification of nodes and relations

To determine whether the nodes are defined correctly, it is possible to look at the CPTs, see [Appendix G](#).

For verification, the values in the CPTs were checked to see if they match the model described earlier.

The $LT_{Vx.x}$ nodes and $G_{Px.x}$ nodes should therefore have a uniform distribution, looking at the CPTs in [Appendix G](#), this is met.

Furthermore, the $V_{x.x}$ nodes should have increasing failure probabilities given a higher $LT_{Vx.x}$. Looking at the CPTs in [Appendix G](#), this is met.

For the $B_{Px.x}$ nodes, there must be a clear difference between the 5 states. State 1 and state 2 should additionally have a lower probability because it is less likely that a pump often operates in one of these two states. This is also met according to the CPTs in [Appendix G](#).

For the LS_x nodes, there should be a clear distinction in the probability of failure if more components fail at the same time each time. So suppose 3 components fail simultaneously then a higher probability of failure should be given than if only 2 components fail simultaneously. This is also in order for this model according to the CPTs.

Finally, there is the PL node. For this, the same applies as for the LS_x nodes. If 3 line segments fail simultaneously then this should give a higher failure probability than if only 2 line segments fail simultaneously.

All in all, it can be seen from the CPTs in [Appendix G](#) that the nodes are correctly defined in the model.

To check whether the relationships between the nodes are well defined, tests are performed to determine independence. For instance, the LS_x nodes must depend on the valves and pumps in that particular line segment. Furthermore, the PL node must depend on the LS nodes but not directly on the other nodes.

The PL is correctly defined if it depends only on the LS_x nodes and is independent of the other nodes in the BN. The independence check for the PL node gives the following result:

(PL ⊥ G2.1, LT1.2, P1.1, P5.1, LT3.7, B2.2, LT3.1, LT5.3, LT5.4, V3.6, LT3.2, V5.2, V4.5, LT3.3, B3.1, B4.1, V3.2, G1.1, V3.8, V4.3, LT5.1, V2.1, LT3.6, G3.1, V3.5, LT4.4, V5.4, B2.1, LT4.5, V3.7, V2.2, V3.3, LT2.2, V4.4, P2.2, LT4.3, G4.1, LT2.1, P2.1, LT4.2, B5.1, G2.2, P4.1, P3.1, V4.1, LT3.8, LT3.4, G5.1, LT3.5, B1.1, LT5.2, V3.1, V3.4, V5.1, LT1.1, V4.2, V1.1, LT4.1, V5.3, V1.2 | LS1, LS2, LS3, LS4, LS5)

This shows the PL node given the LS nodes is independent of all other nodes, these are shown in light grey.

The independence checks for the LS nodes can be found in [Appendix H](#).

From the results of the checks, it can be seen that the values and relationships are well constructed in the BN.

5.5.2 Verification: Sensitivity analysis

Verification can be performed using sensitivity analysis [140]. This involves looking at how a small change in the values of the variables affects the BN. This can provide insight into which nodes have a large influence on the outcome and which nodes have less influence.

Moreover, known scenarios can be used to determine whether the BN is a good representation of the proposed system [138]. In this research, verification is used to determine which nodes to use as evidence, due to the limitation of the PGMPY package used [137]. It also considers the situation where the all evidence nodes are set to the worst state and set to the best state. The results can be used to determine whether the BN meets expectations.

To determine which nodes can be temporarily excluded for this research, it can be considered how each node affects the outcomes of the PL node. To determine this with the constraints in the BN, 2 nodes were each fixed to exclude and then varied across the remaining nodes. By running the BN, it is possible to determine how the combination of eliminated nodes affects the value of the PL node. The aim is to determine which nodes have the least impact on the PL value.

Looking at the number of parent nodes of the different line segments, it is noticeable that line segment 3 has the most parent nodes. Due to the limitations of the package, the choice was made to see which 3 nodes of line segment 3 have the least influence on the outcome of the PL node. To determine sensitivity, the first scenario considered was where all the remaining LT, G and B nodes are in the worst state. Worst case $PL(0)$ gives the probability of non-failure at the time when all LT, G and B nodes are in their worst state. For the LT nodes, this can vary in 15, 20, 25, 30 or 50. For the G nodes it is state 2 and for the B nodes it is state 5.

The results from the BN for the $PL(0)$ can be found in [Table 5.1](#). The top rows of this table show for which nodes no evidence was entered. Forward inference was then performed with the remaining LT, B and G nodes to determine the value of $PL(0)$. $PL(0)$ is the probability of non-failure of the production line.

If this is averaged then the deviation from the mean can be looked at to determine which combination has the least and most impact. From the values of [Table 5.1](#) for worst case $PL(0)$, the average is taken. It can then be determined what the deviation from the mean is for each combination. The average for $PL(0)$, in the case where all the remaining

LT, G and B nodes are at their worst state, is equal to 0.0179. The deviation of each combination from this mean is shown in [Figure 36](#).

Next, the scenario in which all LT, except excluded LT nodes, B and G nodes are in the best state was also considered. In the best state means state 1.

The average for PL(0), based on the values of [Table 5.1](#), in the case that all the remaining LT, G and B nodes are at the best state, is equal to 0.5147. The deviation of each combination from this average for PL(0) is shown in [Figure 37](#).

Table 5.1: Values best and worst case scenarios for all different combinations of LT nodes of LS3

	LT3.1 LT3.2 LT3.3	LT3.1 LT3.2 LT3.4	LT3.1 LT3.2 LT3.5	LT3.1 LT3.2 LT3.6	LT3.1 LT3.2 LT3.7	LT3.1 LT3.2 LT3.8	LT3.1 LT3.3 LT3.4	LT3.1 LT3.3 LT3.5	LT3.1 LT3.3 LT3.6	LT3.1 LT3.3 LT3.7	LT3.1 LT3.3 LT3.8	LT3.1 LT3.4 LT3.5	LT3.1 LT3.4 LT3.6	LT3.1 LT3.4 LT3.7
Worst case PL(0)	0.0166	0.017	0.018	0.0182	0.017	0.017	0.0177	0.0172	0.0173	0.0176	0.0174	0.0176	0.0177	0.0178
Best case PL(0)	0.5241	0.5174	0.5234	0.5164	0.5262	0.5172	0.5190	0.5237	0.5240	0.5277	0.5179	0.5193	0.5172	0.5255
	LT3.1 LT3.4 LT3.8	LT3.1 LT3.5 LT3.6	LT3.1 LT3.5 LT3.7	LT3.1 LT3.5 LT3.8	LT3.1 LT3.6 LT3.7	LT3.1 LT3.6 LT3.8	LT3.1 LT3.7 LT3.8	LT3.2 LT3.3 LT3.4	LT3.2 LT3.3 LT3.5	LT3.2 LT3.3 LT3.6	LT3.2 LT3.3 LT3.7	LT3.2 LT3.3 LT3.8	LT3.2 LT3.4 LT3.5	LT3.2 LT3.4 LT3.6
Worst case PL(0)	0.017	0.0182	0.0182	0.0173	0.0185	0.0176	0.0179	0.0177	0.0183	0.0177	0.0175	0.0174	0.0185	0.0179
Best case PL(0)	0.5172	0.5246	0.5259	0.5167	0.5181	0.5138	0.5179	0.5214	0.5183	0.5165	0.5219	0.5252	0.5088	0.5109
	LT3.2 LT3.4 LT3.7	LT3.2 LT3.4 LT3.8	LT3.2 LT3.5 LT3.6	LT3.2 LT3.5 LT3.7	LT3.2 LT3.5 LT3.8	LT3.2 LT3.6 LT3.7	LT3.2 LT3.6 LT3.8	LT3.2 LT3.7 LT3.8	LT3.3 LT3.4 LT3.5	LT3.3 LT3.4 LT3.6	LT3.3 LT3.4 LT3.7	LT3.3 LT3.4 LT3.8	LT3.3 LT3.5 LT3.6	LT3.3 LT3.5 LT3.7
Worst case PL(0)	0.0179	0.017	0.0189	0.018	0.0182	0.0188	0.0182	0.0174	0.0183	0.018	0.0192	0.0183	0.0171	0.0184
Best case PL(0)	0.5195	0.5134	0.5109	0.5159	0.5106	0.5084	0.5014	0.5084	0.5134	0.5109	0.5056	0.5096	0.5191	0.5191
	LT3.3 LT3.5 LT3.8	LT3.3 LT3.6 LT3.7	LT3.3 LT3.6 LT3.8	LT3.3 LT3.7 LT3.8	LT3.4 LT3.5 LT3.6	LT3.4 LT3.5 LT3.7	LT3.4 LT3.5 LT3.8	LT3.4 LT3.6 LT3.7	LT3.4 LT3.6 LT3.8	LT3.4 LT3.7 LT3.8	LT3.5 LT3.6 LT3.7	LT3.5 LT3.6 LT3.8	LT3.5 LT3.7 LT3.8	LT3.6 LT3.7 LT3.8
Worst case PL(0)	0.0179	0.0187	0.0182	0.0191	0.018	0.0179	0.0177	0.0178	0.0181	0.0181	0.0186	0.0175	0.0176	0.0179
Best case PL(0)	0.5093	0.5121	0.5071	0.5083	0.5100	0.5099	0.5072	0.5028	0.5014	0.5080	0.5130	0.5090	0.5097	0.4908

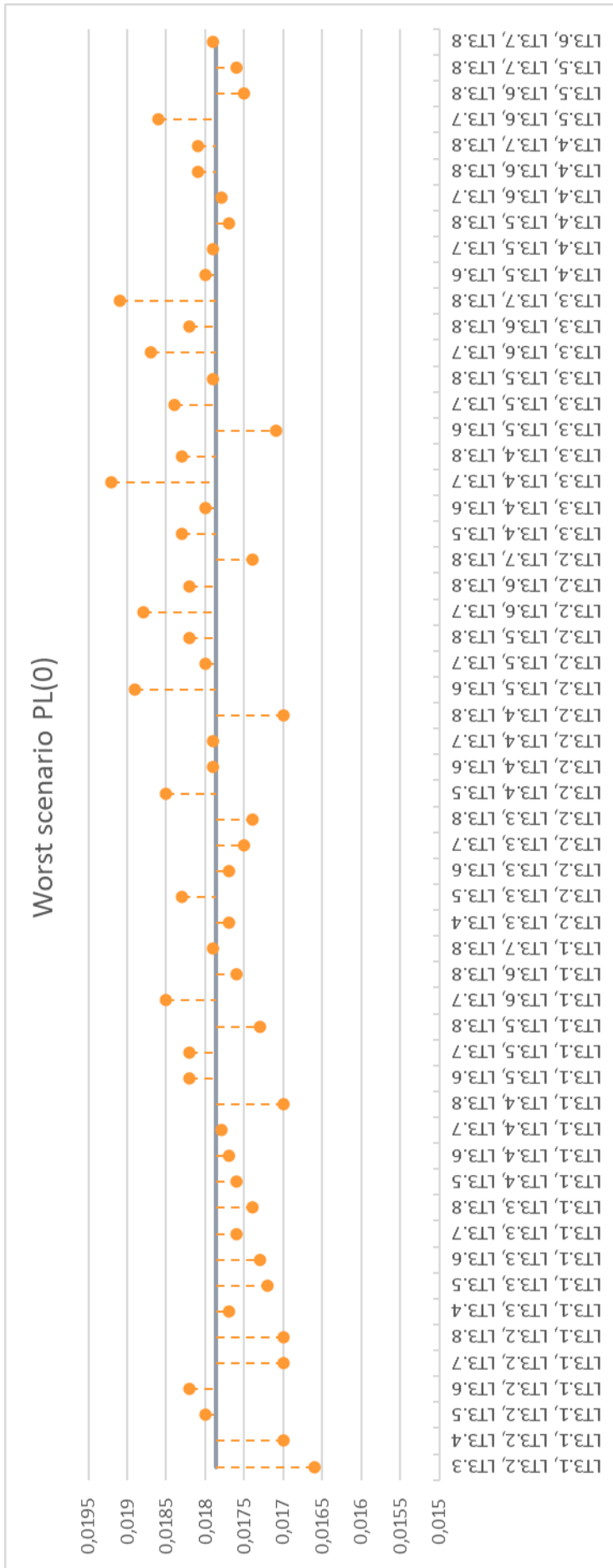


Figure 36: Worst case scenario PL(0)

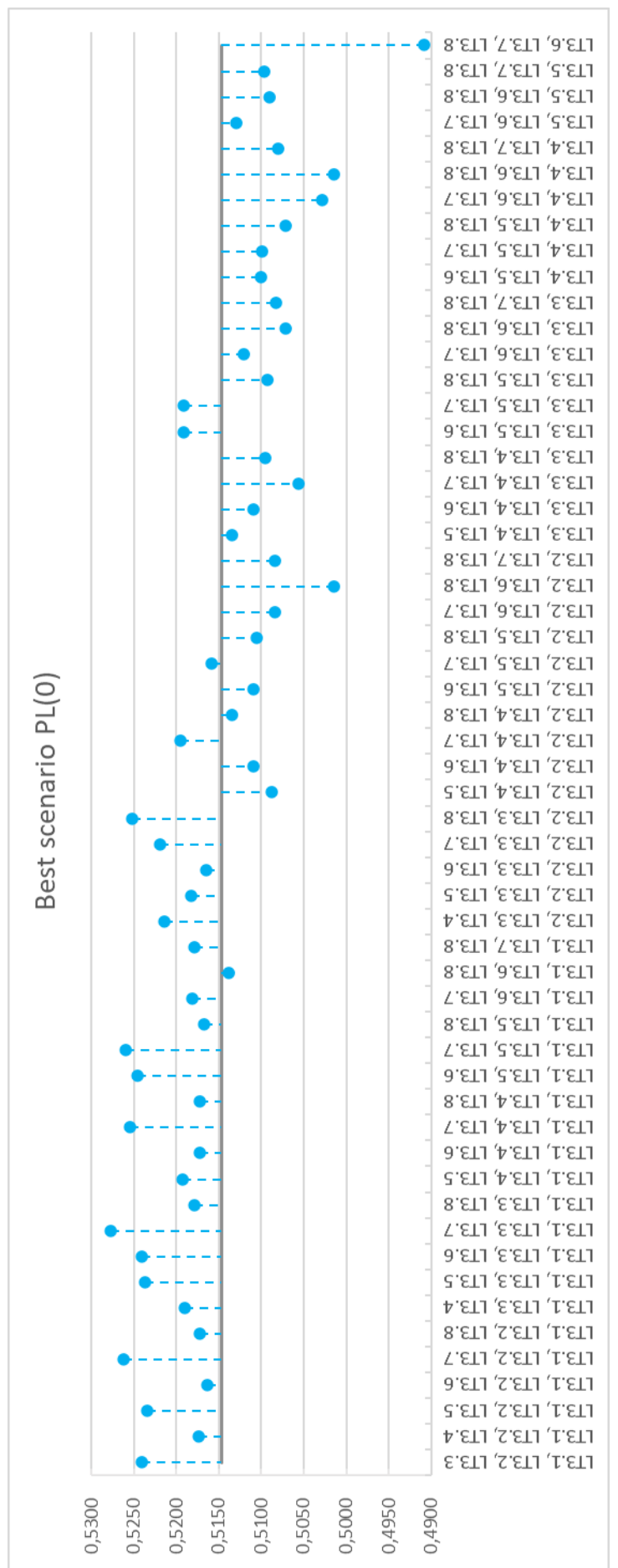


Figure 37: Best case scenario PL(0)

To determine which combinations of LT nodes have the least impact on the outcomes for PL(0), Figure 36 and Figure 37 are combined in Figure 38.

It can be seen from Figure 38 that there are some combinations that have little deviation from the mean. These are the combinations:

- LT3.1, LT3.2, LT3.6
- LT3.1, LT3.6, LT3.8
- LT3.2, LT3.3, LT3.6
- LT3.2, LT3.4, LT3.8
- LT3.2, LT3.5, LT3.7
- LT3.3, LT3.4, LT3.5

Thus, if evidence is used to determine the output of the BN, one of these combinations can be chosen to provide no evidence for. After all, verification has shown that these combinations have little impact on the average value of PL(0) in the best and worst case scenarios.

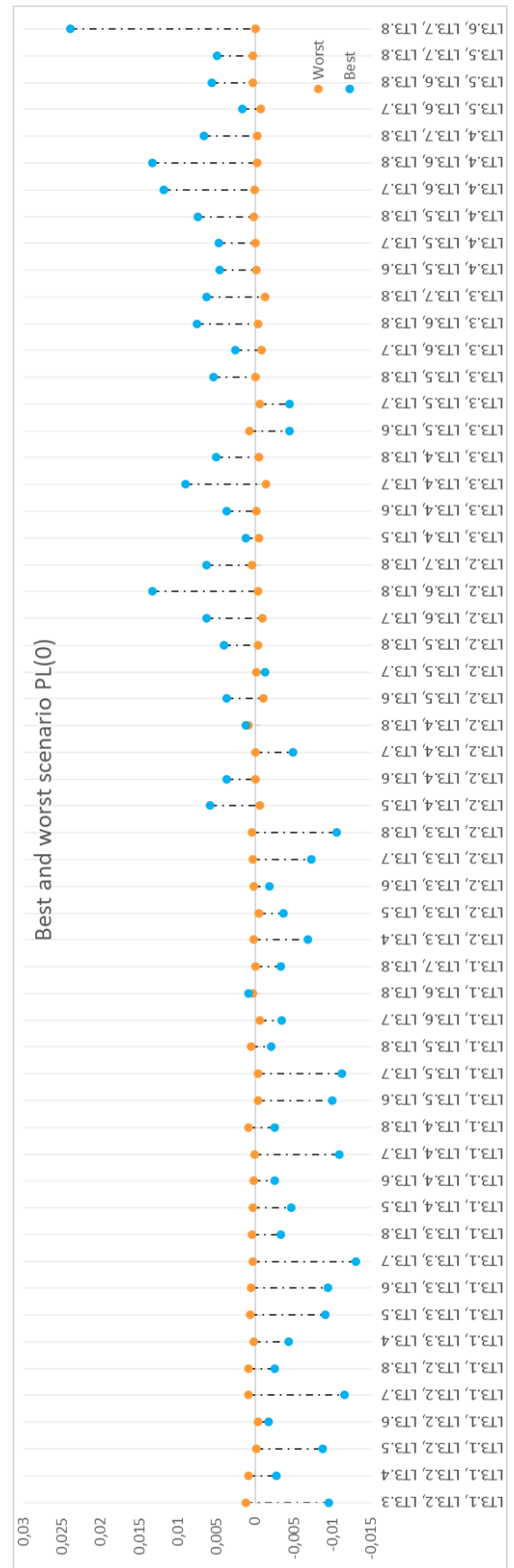


Figure 38: Combination of Figure 36 and Figure 37

Now the research will continue to see if the BN performs as expected.

For this part of the verification of the BN, some inferences can be made to check whether the BN does what was thought of beforehand with the data. Here, it is necessary to see how the BN works with data. This can be real data or synthetic data.

It needs to be established whether the BN gives the outputs that can be expected when certain data is entered as evidence.

It looks at what happens when all component nodes are set to the lowest and highest positions. The BN should be constructed based on the assumption that if all component nodes are in the lowest state, there is a low probability of failure and hence low probability of downtime. The opposite is true for the expectation if all component nodes are in the highest state. The BN should then give a high probability of failure and hence downtime.

First, the failure probability of the production line is examined if all component-level nodes (nodes $V_{x,x}$ and $P_{x,x}$) are set to their lowest values (in this case, all set to state 0). If all component-level nodes are in a particular state, this is also called introductory evidence. That is the obtained proof of whether a node will definitely be in a particular state. Based on this, exact inference can be used to calculate the failure probability of the node PL. Variable elimination was used in this case. This involves eliminating variables that are not needed to arrive at an outcome for the desired variable [141]. If for the component-level nodes (nodes $V_{x,x}$ and $P_{x,x}$), see [Figure 31](#), it is assumed that they are all in state 0, this is the proof that is introduced to determine what happens to the outcome of node PL.

It is assumed that if all component-level nodes are in the best state, state 0, then there is a low probability of failure. In other words, a low value should come out of the inference for PL(1).

For this case, subsystem-level nodes (LS_x) are eliminated according to BN's variable elimination techniques [141].

At the point when for all component-level nodes state 0 is entered as evidence then it follows from the inference with the BN that:

$$\begin{aligned} PL(0) &= 0.8741 \\ PL(1) &= 0.1259 \end{aligned}$$

Thus, there is 12.6% probability of failure of the production line at this time, given that all component-level nodes are in the lowest state. This corresponds to the assumption given earlier that there is a low probability of failure of the production line if all components are in a healthy state.

It was then assumed that if all component-level nodes are in the worst state, state 1, then there is a high probability of failure of the PL node. In other words, the value of PL(1) must be high.

If all component-level nodes (nodes $V_{x,x}$ and $P_{x,x}$) are in the worst state (state 1), then the output of the BN with inference is:

$$\begin{aligned} PL(0) &= 0.0123 \\ PL(1) &= 0.9877 \end{aligned}$$

This shows that the probability of failure of the production line is 98.8% at the time when all component-level nodes are in the worst condition. This matches the earlier assumption. The failure probability of the production line increases dramatically the moment all components are in a bad state.

If more data becomes available for the looptijd of the valves, the logistic function for each valve has to be adjusted. This is because the logistic function, parameter m , see [Chapter 4.1.2](#), is based on the average looptijd in the historical data. As a result, the CPTs of the $V_{x,x}$ nodes will change.

Currently, theoretical data from the pumps are used, but when data from sensors at the pumps become available, the BN has to be validated again.

5.5.3. Verification conclusion

All in all, it can be concluded that based on the verification in several areas, the BN meets the specifications and descriptions described earlier. The first verification looked at definition of the nodes and the relationships between them. This showed that the BN created in PGMPY corresponds to the BN as stated in [Chapter 5.2](#). All nodes are present, have a CPT and the interrelationships are in order. Next, sensitivity analysis was used to see which LT nodes from line segment 3 have the least impact on the outcomes of the PL node. This was done because there are limitations in PGMPY for the number of evidence values that can be entered to perform inference. This showed that there are a number of combinations of different LT nodes that have little impact on the outcomes for the PL nodes. In further use of the BN, one can choose between these combinations which LT nodes are not assigned evidence. Furthermore, it has been examined whether the BN responds as expected. For this, the component-level nodes (nodes $V_{x,x}$ and $P_{x,x}$), see [Figure 31](#), were first set to the best state, state 0. The BN works properly if a low probability

of failure for the production line then comes out of the inference. In other words, a low value for the PL(1). The inference shows that this is indeed the case. At the time when all component-level nodes are in the healthy state, it follows from the inference that $PL(1)=0.1259$. Thus, there is a low probability of failure of the production line. Next, the scenario where all component-level nodes are in the worst state, state 1, was considered. The BN works properly if, in this case, a high value comes out of the inference for the PL(1). After entering the evidence and performing inference, it follows that: $PL(1)=0.9877$. This establishes that the BN is verified and that it works as previously described and specified.

5.6 Conclusion

This chapter answers the sub-research question: *What are the steps to develop and verify the method of the model?* In order to create the BN, it is first necessary to determine which nodes to include and the relationships between them. In this research, there should be nodes representing the parameters that affect the failure probability of the valves, in this research the looptijd ($LT_{V_{x,x}}$), and pumps, in this research the cavitation ($G_{P_{x,x}}$) and BEP ($B_{P_{x,x}}$). There must also be nodes to represent the valves ($V_{x,x}$) and pumps ($P_{x,x}$) in the production line. The production line in this research is divided into five different line segments. So there should be nodes representing the line segments (LS_x). Finally, there should be a node (PL) representing the probability of failure of the entire production line. After that, all the different probabilities and CPTs need to be set up for the BN. All nodes and probabilities are then programmed in Python using the PGMPY package. This makes it possible to perform forward inferences and determine whether the BN satisfies the previously described functionalities. The BN is then subjected to some verification to check if the BN is then still satisfactory. Now that all this has been done, it is possible to move on to determining the BN performance. The BN forms the basis of the final model. However, some other steps are needed to arrive at a complete model with which to determine whether reliability improves. In the final model MTBF will also be used. In [Chapter 6](#) a more detailed explanation will be given about the complete model, the role of the BN and how to determine whether the reliability improves.

6. Model performance

In this chapter the model, with BN, and the sub-research question: *In what way is the implementation of the model contributing to the reliability of the production line?*, will be discussed in more detail.

As mentioned earlier, the BN is part of the model. The model contains more functionalities that can ultimately be used to determine whether maintenance is needed or not.

To determine the contribution of the model to improving production line reliability, the definition of reliability is first considered. Then some KPIs are set up to test the influence of the model on reliability compared to the situation if the model were not used. Following this, the flowchart of the model, with the BN, is discussed and also some assumptions made.

6.1 Reliability

It is important to establish the definition of reliability for this research so that this can be used to determine what the model contributes to improving reliability. In this research, the following definition for reliability is used: "The ability of a system or component to perform its required functions under stated conditions for a specified period of time." [142].

The reliability of the line and its components is linked to the duration during which it can perform its function correctly [143]. In equation form, reliability $R(t)$ is given as [144] [Eq. 6.1](#).

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

Eq. 6.1

λ failure rate
 t time

The failure rate can be determined using the mean time between failure (MTBF) [144], see [Eq. 6.2](#).

$$MTBF = \frac{1}{\lambda}$$

Eq. 6.2

λ failure rate

This is the time between failures. It is possible to determine this both at component level and system level.

The probability of failure of a system can be linked to reliability with [Eq. 6.3](#).

$$R(t) = 1 - F(t)$$

Eq. 6.3

Here, $F(t)$ is the formula for the cumulative failure probability [143].

From the above equations and the definition of reliability, it can be inferred that reliability improves when a component can perform its function correctly for longer. In other words, if the MTBF improves, reliability also improves. By performing timely maintenance, it is possible to improve component reliability. Performing the maintenance ensures that the component does not have the opportunity to fail. This increases the time between failures. Looking at [Eq. 6.1](#), this indicates that component reliability improves. Implicitly, this will also improve the reliability of the line [145].

To achieve higher reliability within a certain time interval, it is necessary to look at how failures can be predicted and prevented [146]. Predicting and preventing failures improves reliability as the MTBF becomes longer. Improved reliability results in fewer unplanned downtimes and lower maintenance costs [147]. Implicitly, it is indicated that predicting failures allows timely intervention and leads to fewer unplanned downtimes and lower maintenance costs. This is consistent with a predictive maintenance strategy, see [Chapter 3.2](#).

This research looks at how a predictive maintenance strategy can help improve the reliability of a soft drink production line. For this purpose, a model was created that can be used to determine the condition of the components in the production line and the condition of the entire production line.

To determine how a predictive maintenance strategy improves reliability, it is compared with the current corrective maintenance strategy. According to Barringer [147], increased reliability results in fewer unplanned downtimes and lower maintenance costs. For this research, KPIs are used to determine what the improvement is with the PdM versus the corrective maintenance strategy.

6.2 Key Performance Indicators

In order to determine what the reliability of the line is and how it changes when a different maintenance strategy is applied, a Key Performance Indicator (KPI) for reliability has to be identified. A KPI is defined as: "the critical (key) quantifiable indicator(s) of progress towards an intended result." [148].

6.2.1 Maintenance downtime index

The first KPI that will be used is the Maintenance Downtime Index (MDI) [149]. This is the ratio between the number of hours of downtime used for scheduled maintenance and the total number of hours of downtime, see Eq. 6.4.

$$MDI = \frac{DT \text{ hours for scheduled maintenance}}{\text{total DT hours}} = \frac{DT \text{ hours for scheduled maintenance}}{\text{Scheduled} + \text{unscheduled DT hours}}$$

Eq. 6.4

This KPI provides insight into how many hours of the total downtime (DT) are used for planned maintenance. From this, it can be deduced how many hours of the total DT were used for unscheduled maintenance. The lower the MDI, the more hours go into unscheduled maintenance [149].

If components receive more frequent maintenance at scheduled maintenance times, this benefits the MTBF. As a result, a component will be able to perform its function for longer and the MTBF increases. Implicitly, a higher MDI is an indication of improved reliability and fewer unplanned downtimes.

6.2.2 Cost key performance indicators

The second, third and fourth KPIs look at costs. According to Peng et al [150], reliability and maintenance costs are linked, see Figure 39.

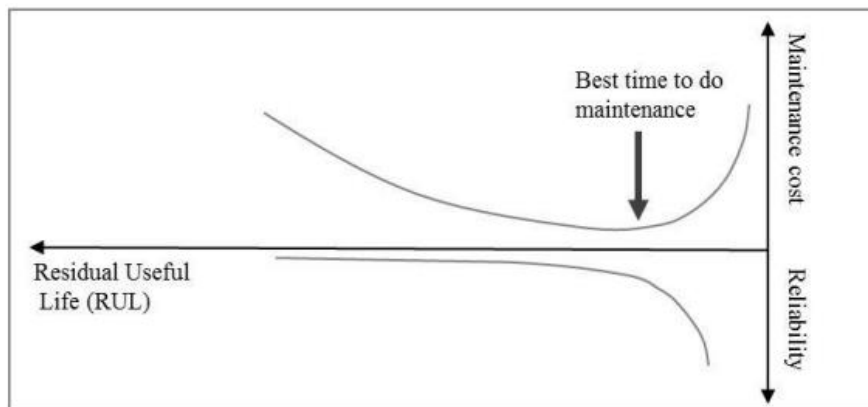


Figure 39: Reliability and maintenance costs [150]

If a component reaches the end of its service life, reliability drops and maintenance costs rise [150] [151].

Furthermore, it is known that there are differences between the maintenance costs of different maintenance strategies [152], see Figure 40.

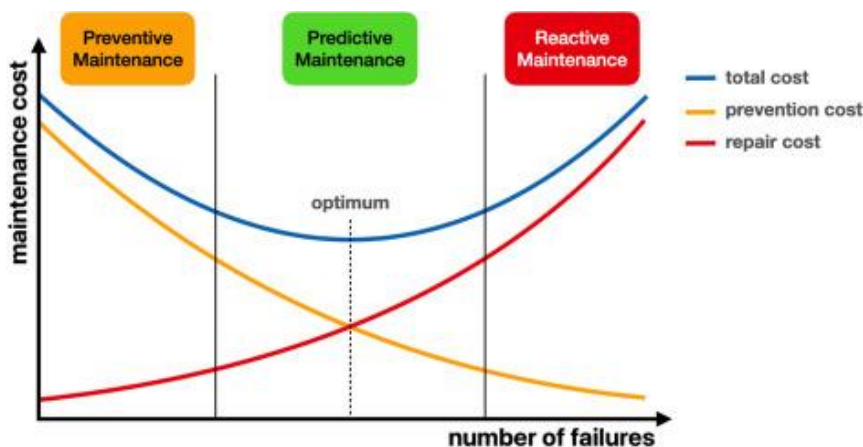


Figure 40: Costs for different maintenance strategies [152]

In this research, maintenance costs and reliability are linked, see [Figure 39](#). Maintenance costs are lowest when maintenance is performed just before the end of a component's lifetime. [Figure 40](#) shows that the lowest maintenance costs are observed when predictive maintenance is correctly used. Predicting when maintenance is needed can improve the reliability of a component. This is because a component then receives maintenance at the right time, increasing the mean time between failure. The MTBF is linked to reliability, see [Eq. 6.1](#) and [Eq. 6.2](#). All in all, it can be observed that if there are low total maintenance costs that this indicates that components have received maintenance at the right time. In this research, this indicates that the reliability of the components is increasing.

A brief overview of the costs considered in this research will be given for each maintenance strategy.

6.2.2.1 Corrective maintenance costs

For corrective maintenance, the costs are made up of replacement cost of (parts of) component(s), labour cost and cost of production loss [153].

The replacement cost is made up of the number of components and the cost per component, see [Eq. 6.5](#). A component consists of parts. As an example, a valve has at least a stem or disc and an O-ring. Pumps, for example, have an impeller and bearings. If too many parts of a component are broken, then a choice can be made to replace the entire component (valve or pump).

$$C_R = \sum_{i=1}^n (P_i * C_i)$$

Eq. 6.5

- i stands for unique parts/components.
- P_i represents the number of part or component i to be replaced.
- C_i is the price of the part or component i , in €.
- C_R is the total cost of replacing part(s) or component(s), in €.

The cost for the man-hours required consists of the the number of hours worked (H_j) times the cost per hour for wages (S_j). A summation sign is used in [Eq. 6.6](#) because multiple mechanics may be needed who may be paid different wages.

$$C_L = \sum_{j=1}^n (S_j * H_j)$$

Eq. 6.6

- j is the number of mechanics.
- S_j is the hourly wage of mechanic j , in €/Hr.
- H_j is the number of hours mechanic j worked, in Hr.
- C_L is the total wage cost, in €.

Finally, there is the cost of production loss, [Eq. 6.7](#). This consists of two types of cost. The first kind are costs incurred because the batch has to be rejected. The second kind is the cost of production loss. The cost of rejecting the batch depends on how much soft drink is left in the line/tank and should no longer be used (according to food and commodity authority rules). The cost of production loss is the number of hours the line cannot be used times the loss in revenue per hour. This calculates the time the line is out of use for both repair and cleaning.

$$C_{LP} = (B * C_B) + (LPH * CLPH)$$

Eq. 6.7

- b is the number of litres to be discarded from the current production batch, in L.
- C_B are the costs involved, in €/L.
- LPH are the number of production loss hours, in Hr.
- $CLPH$ are the average cost per hour of production loss, in €/Hr.
- C_{LP} are the cost of lost production, in €.

This gives the overall equation for the cost of corrective maintenance (C_{Corr}), see [Eq. 6.8](#).

$$C_{Corr} = C_R + C_L + C_{LP}$$

Eq. 6.8

6.2.2.2 Preventive maintenance costs

For preventive maintenance, costs can be split into different cost items [153], [Eq. 6.9](#). There are the costs of corrective maintenance the moment a component does fail earlier than planned. There are also the costs of replacing a component as scheduled.

$$C_{prev} = C_{corr} + C_{prev,R} + C_{prev,L}$$

Eq. 6.9

$C_{prev,R}$ is the cost of replacing one or more parts or component(s) based on a defined schedule. This consists of the cost per part or component times the number, see Eq. 6.10.

$$C_{prev,R} = \sum_{l=1}^n (P_l * C_l)$$

Eq. 6.10

- l represents the number of unique parts or components to be replaced.
- P_l are the parts or components l to be replaced.
- C_l are the cost per part or component l , in €.
- $C_{prev,R}$ are the total replacement cost for preventive maintenance, in €.

Besides the cost of the parts or components, there are also costs for the mechanics who have to carry out the maintenance. This consists of the wages per mechanic and the number of hours they worked, see Eq. 6.11.

$$C_{prev,L} = \sum_{m=1}^n (S_m * H_m)$$

Eq. 6.11

- m is the number of mechanics.
- S_m is the hourly wage of mechanic m , in €/Hr.
- H_m is the number of hours mechanic m worked, in Hr.
- $C_{prev,L}$ is the total wage cost for preventive maintenance, in €.

For preventive maintenance, this research assumes that maintenance is carried out at times when the line is not in use. As a result, there is no loss of production hours.

6.2.2.3 Predictive maintenance costs

For predictive maintenance, the costs can be made up of costs for inspection and replacing components that are at the end of their useful life according to condition monitoring, Eq. 6.12 [153].

$$C_{pDM} = C_{insp} + C_{CR} + C_{LR}$$

Eq. 6.12

In this, C_{insp} is the cost of the additional checks to be carried out. This can be represented by the Eq. 6.13.

$$C_{insp} = \sum_{q=1}^n (IH_q * Cl_q)$$

Eq. 6.13

- q is the number of mechanics.
- IH_q is the number of hours inspections are done by mechanic q , in Hr.
- Cl_q is the labour cost of mechanic q , in €/Hr.
- C_{insp} is the total cost incurred for inspections, in €.

C_{CR} are the costs incurred for replacing the components or parts that are due for replacement according to condition monitoring. This can be determined using Eq. 6.14:

$$C_{CR} = \sum_{s=1}^n (P_s * C_s)$$

Eq. 6.14

- s is the number of unique parts or components.
- P_s is the number of parts or component s to be replaced based on condition monitoring.
- C_s is the price per part or component s to be replaced, in €.
- C_{CR} is the total cost of replacing part(s) or component(s), in €.

The cost of man-hours required consists of the number of hours worked (H_r) times the cost per hour for wages (S_r). A summation sign, in Eq. 6.15, is used because multiple mechanics may be needed who may be paid different wages.

$$C_{LR} = \sum_{r=1}^n (S_r * H_r)$$

- r is the number of mechanics. Eq. 6.15
 S_r is the hourly wage of mechanic r , in €/Hr.
 H_r is the number of hours worked by mechanic r , in Hr.
 C_{LR} is the total wage cost, in €.

6.2.2.4 Cost KPIs

From the above, the following KPIs can be derived based on the maintenance standard EN15341 [154]. For the second KPI, Eq. 6.16, this research looks at the ratio between the cost of corrective maintenance to the total maintenance cost.

$$E15 = \frac{\text{Corrective maintenance cost}}{\text{Total maintenance cost}} * 100$$

Eq. 6.16

The third KPI, Eq. 6.17, looks at the cost of preventive maintenance relative to total maintenance costs.

$$E16 = \frac{\text{Preventive maintenance cost}}{\text{Total maintenance cost}} * 100$$

Eq. 6.17

The fourth KPI, Eq. 6.18, looks at the cost of condition-based maintenance costs relative to total maintenance costs. In this research, condition-based maintenance costs are assumed to be the same as predictive maintenance costs.

$$E17 = \frac{\text{Condition based maintenance cost}}{\text{Total maintenance cost}} * 100$$

Eq. 6.18

In this research, the total maintenance cost is determined by adding the corrective, preventive and predictive maintenance costs.

It is assumed in this research that the reliability of the components, and hence the reliability of the production line, increases at the time when there are lower total maintenance costs in the situation where the model is used. In addition, it should be noted that the cost item for predictive maintenance is higher in the case where the model is used.

6.3 Overview of the model

Before looking at a comparison between the current situation and the situation where the predictive maintenance model is used, there are some steps that need to be done.

The model consists of a part that works based on the method chosen earlier in Chapter 4.2, namely the BN. In addition, there is another part of the model that works with the MTBF. Both parts are needed to eventually use the model. First, the BN used in the model will be looked at. Then the use of the MTBF part will be covered. The overall model is shown with a flow chart in Chapter 6.3.4.

6.3.1 BN

The BN is first used to determine what the outputs are for the PL and LS nodes based on historical data. This can be used to determine whether clear trends are visible. As input, the values are used from the process data. Next, forward inference can be used to determine what the outputs are based on the given input evidence.

This shows that there are already indications of potential problems about 2 hours in advance. By determining the values of both PL and LS nodes associated with the indications on failure and the failure itself, this can be used later for prediction. Table 6.1 shows the values of PL and the LS nodes at which there is an indication that something is going to fail soon and the value at which something might fail. So this does not necessarily mean that something will fail immediately, but there is that chance.

Table 6.1: Warning and potential failure values of PL and LS nodes

	PL(1)	LS1(1)	LS2(1)	LS3(1)	LS4(1)	LS5(1)
Warning	0.6025	0.5073	0.5270	0.5799	0.5111	0.5421
Potential failure	0.6031	0.5167	0.5327	0.5828	0.5115	0.5485

These values were arrived at by performing inference with evidence values. For this, 1 node from the LS was set to the worst state each time and the other nodes were all set to the state that occurs most often in the historical process data.

6.3.2 MTBF

The historical process data was also used to determine the MTBF of the various valves. For the pumps it is assumed that, based on the manufacturers data, the pump requires maintenance once a year on average.

Eq. 6.19 is used to determine the MTBF for the valves and pumps [155].

$$MTBF = \frac{\text{Total time in use}}{\text{Number of failures}}$$

Eq. 6.19

For total time in use, the number of hours the line was used to produce plus the number of hours the line was cleaned has been chosen. See Appendix I: Table I.1 for the MTBF for every component.

To determine how far into its life a component is, a piece of code has been written. This can be used to track how many hours are left of the MTBF of each individual component.

It is assumed that there is a critical zone of 2% of the MTBF. As soon as a component enters the 2% zone then a signal is returned with the time remaining.

6.3.3 Cost analysis

It is assumed that the average cost for the parts of a valve is €160.-. Once a valve needs maintenance then a sum of €160.- is charged. For pumps, it is assumed that the average cost for parts is €1,000.-. Furthermore, it is assumed that at least two mechanics are always needed to carry out the repairs. If several components need maintenance at the same time, then two mechanics are charged extra per component. The idea behind this is that this allows for the most efficient work possible. In this research, the hourly wage of a mechanic was set at €75,- per hour. Then there is the cost of flushing away the soft drinks in case of an unplanned downtime. As no data on this is available, it is assumed in this research that this is a fixed amount of €1,000.-. Table 6.2 shows the summary of all costs.

Table 6.2: Summary of all the costs

	Cost [€]
Repair maintenance per valve	160,-
Repair maintenance per pump	1000,-
Wages per hour per mechanic	75,-
Rinse away soft drinks	1000,-

Since the service life of a valve and pump is between 10 and 20 years [33] [34], it is assumed that in the time frame considered in this research, only repairs are done and not complete replacements.

Furthermore, it is assumed that inspecting a valve takes 30 minutes on average. Repairing a valve on average 1 hour. Inspecting a pump averages 1 hour and repairing a pump averages 2 hours. Inspecting a valve requires 1 mechanics and inspecting a pump requires 2 mechanics. Only labour costs are charged for inspections where no parts are replaced. See Table 6.3 for the overview.

Table 6.3: Summary of the hours for inspection and maintenance

	Time [hr]	Number of mechanics
Inspection of valve	0.5	1
Inspection of pump	1	2
Repair maintenance of valve	1	2
Repair maintenance of pump	2	2

6.3.4 Flow chart of the model

For the PdM model, the previously discovered times between processes are used. Indeed, historical data has shown that there is often a period of time between processes. Using this time to carry out inspections and minor maintenance can potentially reduce the number of unplanned downtimes.

It is assumed that the different types of soft drinks that have crossed the line in the past are the same as those that should cross the line in the near future. As a result, the process data can be used to create a dataset for the PdM model.

Pump dataset

To arrive at a dataset for the pumps, assumptions were made about the type of product that crossed the line. These include whether they are, for example, sugary soft drinks or just flavoured water. In addition, assumptions were made about the temperatures and speeds at which these products crossed the line. This applies to both the production processes and the CIP processes. As indicated earlier in [Chapter 4.1.3](#), [Eq. 4.2](#), the occurrence of cavitation depends on the velocity, density and vapour pressure of the liquid passing through the pumps. Chemically, this can be further derived as shown in [Appendix C](#). the $NPSH_R$ was determined based on assumptions about the pump curves, as explained in [Chapter 4.1.3](#). Based on this and the pump characteristics, a random dataset was then created for the pumps using the Scikit package in Python. This took into account the expectations in which states the pumps will operate most often.

To show how the BN method and the MTBF are used in the model to decide when maintenance should be carried out, a flow chart has been created, see [Figure 41](#).

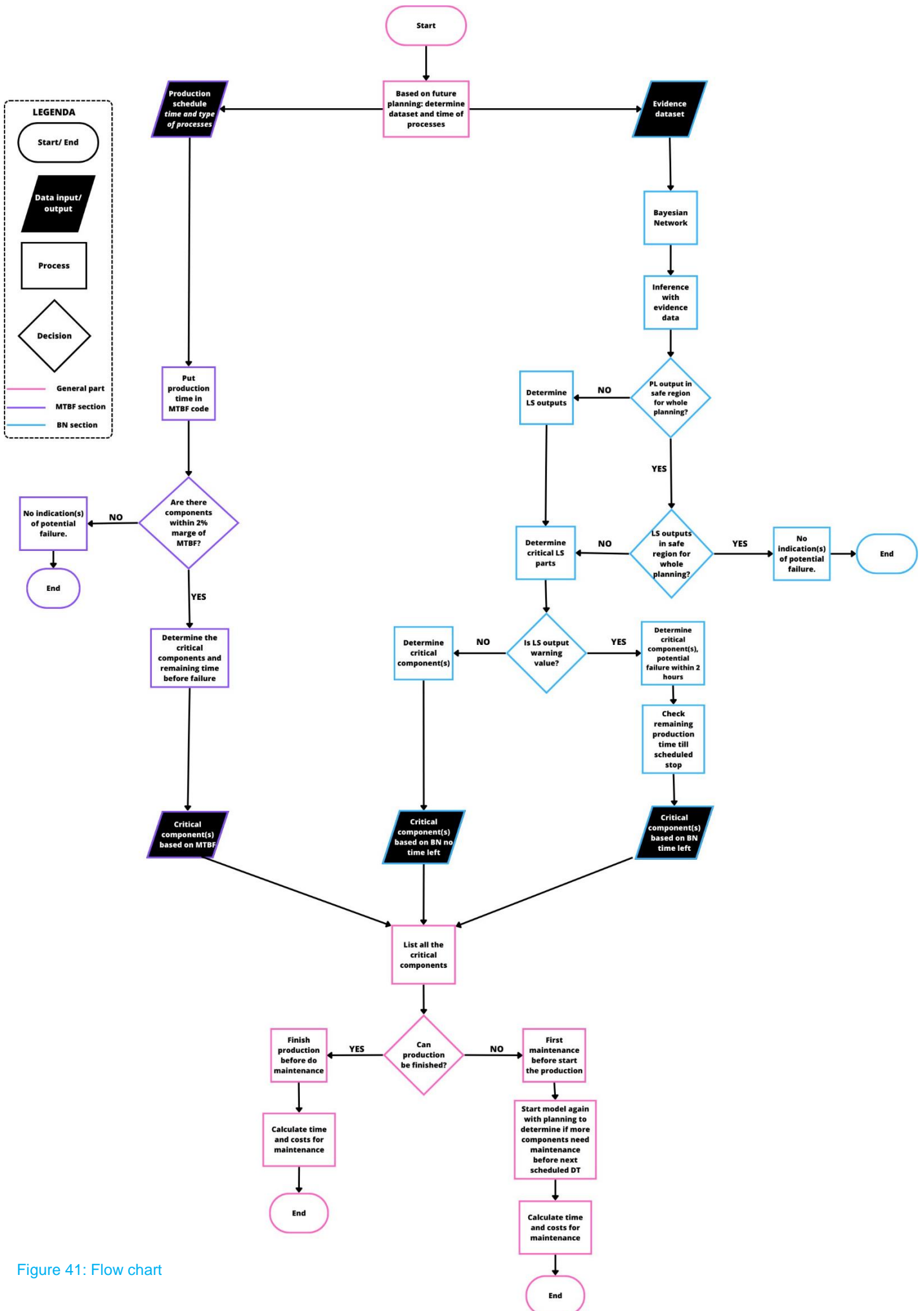


Figure 41: Flow chart

The flow chart distinguishes between the BN section (sections framed in light blue), the MTBF section (sections framed in purple) and the general section (sections framed in pink).

Starting from start, it is first necessary to determine how long and which processes are sent across the line based on the schedule. Then, based on the historical process data, a dataset can be generated that matches the data for the future process. (See previous assumption about the future processes matching the historical processes). The dataset with how long and which process are used to determine the MTBF (the purple part). The dataset created is used in the BN (the blue part).

MTBF section (purple)

The time to planned downtime should be entered in the code for MTBF. The code uses a time ticker to determine how long each component has been in use. Once the time is added of future production then the code checks if there are any components that are coming or are within 2% of end of life. If not then from this section there is no indication that any components will fail until the next scheduled downtime. If there are components within the 2% margin, then it is necessary to determine which components they are and the remaining life of those components. As output data from this section, a list of the critical components and their expected remaining life is given.

BN section (blue)

The dataset created from the historical data contains values for LT, G and B nodes for every 15 minutes. This was chosen because the analysis showed that valves are often open or closed for longer periods of time. No data was available for the pumps, here it is assumed that little changes in 15 minutes.

The data is used as evidence to then do inferences to determine the output values for the PL and LS nodes. As indicated earlier, values have been identified for the PL and LS nodes that indicate whether a potential failure is approaching. Hence, after doing the inferences, it is necessary to check whether the values of the PL node fall within safe region. In this research, the PL node should have as output a value $PL(1) \leq 0.6025$. Anything above this may indicate that a component is failing in the short term.

If the value of PL(1) exceeds 0.6025, then the output values of LS nodes should be looked at. For these nodes too, there are indicator values that indicate that there may be a component failure in the near future. See [Table 6.1](#) for the values of LS nodes at the warnings.

If all LS values are also within the safe range, there is no indication from the BN that a potential failure will occur during the process. It is then assumed that production can be carried out without any problems.

If there are alarming LS values, it must be determined whether it is a warning value or a value that indicates that more is already going on. In the case that it is a warning value, then it must be taken into account that a potential failure may occur in the next two hours. The component(s) responsible for this must then be determined. Then the time until the next scheduled downtime has to be considered. As output data, it is given which components have a critical value and how much time is left until the next stop.

If the value of the LS nodes already exceeds the warning value(s) then it is only necessary to determine which component(s) are causing this. This is collected and used as output data.

General section (pink part at the bottom)

A list is made of the critical components based on the output data, along with how much time is left from MTBF and from the BN. It then needs to be determined whether there is enough time to complete production without unplanned downtime or not. In case there is enough time left then production can be completed before maintenance is carried out. It can then be calculated how much time is needed for maintenance and then what the cost will be.

If there is not enough time to fully complete production then maintenance/inspections must be carried out before any production can be started at all. Then it has to be redetermined what the schedule is then and whether more components need maintenance. Finally, again the time for maintenance and the costs must be calculated.

Once maintenance has been performed on a component then the usage time must be reset in the MTBF code.

To calculate costs, the formulas given earlier for the KPIs in [Chapter 6.2](#) can be used.

A choice can be made to carry out an inspection and decide on the basis of this whether a component needs maintenance. If the time a component is used is longer than the MTBF then after inspection (i.e. when no maintenance is required), the usage time is not reset but counts up until maintenance is required.

6.3.4.1 Preparations before use of the model

Before the model can be used, planning needs to be determined. For this research, a synthetic schedule was created based on historical production schedules. This determines how long each process takes and when there are planned downtimes. This takes into account the findings from the data analysis of [Chapter 4.1](#). In short, it amounts to about 5,000 hours of line operation. It also takes into account the two stop weeks per year when preventive maintenance can be performed.

Once the schedule is known, the data needed for the input of the BN can be considered. Based on the processes scheduled, a dataset can be created using the historical process data. Here, the values to be entered as evidence are determined for each 15 minutes.

In the case of pump data, no historical data is available. To solve this, it was decided to create a random dataset. This is done based on previous knowledge that a pump is most likely to be in state 3 or 4. The dataset will be based on this with some deviations to the other states.

In the MTBF part, all that remains to be determined is how much time of life the component is at before the schedule kicks in. In this research, it was chosen to set these start times to random, the values used for this research can be found in [Appendix I: Table I.1](#).

6.3.4.2 Verification of the dataset

Looking at the synthetic schedule, it can be seen that according to this schedule, the line will be in operation for 5059 hours. Of this, 4746.5 hours are for production and 312.5 hours for cleaning. This is in line with the hours determined earlier based on the historical schedules. Further, the synthetic schedule includes 2 stop weeks. This also corresponds to the historical schedules. In total, there are 518.5 hours for scheduled downtime. This includes the 200 hours counted for the stop weeks (100 hours per stop week).

In addition, the number of failures of each component can be looked at. Based on historical data, it is known of each component how often they fail per year.

If this is compared with the number of failures per component from the synthetic dataset, it can be seen that it almost matches. See [Appendix I Table I.1](#) column name: "Number of failures in synthetic dataset". It should be noted that a number of components receive maintenance during stop weeks.

6.3.4.3 Assumptions

There are also some assumptions that are made. These are listed below.

Assumption 6.1:

That x number of hours of maintenance can be done by bringing forward a scheduled downtime.

This could be because there would be a planned downtime after the specific process in which the line stopped. By bringing it forward, maintenance can then be done without delay.

Assumption 6.2:

Planned downtime can be used to perform maintenance.

Assumption 6.3:

If the planned downtime time is too short for full maintenance then it is assumed that the remaining time needed for maintenance is counted as unplanned downtime time.

Assumption 6.4:

It is assumed that all parts for the valves and pumps are always in stock. Therefore, there is no waiting time for the part to be replaced. As a result, there is no delay in production planning caused by waiting for parts.

Now the model is ready to be used.

Using the synthetic dataset, the impact of the model can be determined. This is done by using the dataset for the case where the model is used and the case where the model is not used.

6.4 Conclusion

This chapter looked at the sub-research question: *In what way is the implementation of the model contributing to the reliability of the production line?*

Reliability in this research is related to the time a component can perform its expected function under known conditions. To determine how reliability can improve using the model, working with a BN and with the MTBF, some KPIs have been established. Using a synthetic dataset, two situations can be compared with the KPIs. In the first situation, the model will not be used, this corresponds to the current situation where corrective and a small part of preventive maintenance is used. The second situation uses the model as described in the flow chart of [Figure 41](#). The first KPI relates to the number of hours of downtime used to perform planned maintenance compared to the total number of hours of downtime, this is defined as the MDI. The more components receive maintenance during planned downtime hours then this ensures that fewer hours of unplanned downtime are needed. Performing maintenance on time ensures that components can perform their required function for longer. This in turn is related to reliability, which is all about being able to perform the requested function for as long as possible without failure. MDI is therefore seen in this research as an indicator that there are fewer unplanned downtime hours.

KPI two, three and four are cost-related. These KPIs provide insight into the percentage of corrective, preventive and predictive maintenance costs related to the total maintenance costs. These KPIs can be used to determine how the model affects different cost items. As stated earlier in [Chapter 6.2.2](#), maintenance costs and reliability are linked. Lower total maintenance costs and highest percentage of predictive maintenance costs would indicate that reliability is increasing in this research.

A flow chart is then used to explain how the model works and what preparations and assumptions need to be made. Furthermore, how a synthetic dataset was created based on historical production schedules was discussed in detail. [Chapter 7](#) discusses the results obtained and the comparison between the situation with and without the model. [Chapter 7](#) will therefore continue on the sub-research question of [Chapter 6](#), but with the results.

7. Results

This chapter continues with the sub-research question: *In what way is the implementation of the model contributing to the reliability of the production line?*

In [Chapter 6.2](#) and [Chapter 6.3](#), the KPIs and the flow chart were explained. In this chapter the difference between using and not using the model is examined. First, the results of not using the model are presented. Then the results when the model is used. Next, a comparison is made between using and not using the model. The chapter concludes with the results arising from the KPIs defined earlier.

To determine the impact of the predictive maintenance strategy compared to the current corrective maintenance, some KPIs were set up in [Chapter 6.2](#). To calculate these, results are needed. This relates to the number of hours of downtime, both planned and unplanned, as well as for calculating costs.

This will be calculated on an annual basis, as the synthetic dataset is for one year. See [Appendix J](#) for a part of the synthetic production planning.

7.1 Without model

The moment the model is not used and corrective maintenance is used, the following results are obtained. First, the overview of when there was failure of one or more components in the line. This is plotted against the number of hours the line is in use. [Figure 42](#) shows the timeline of failures.

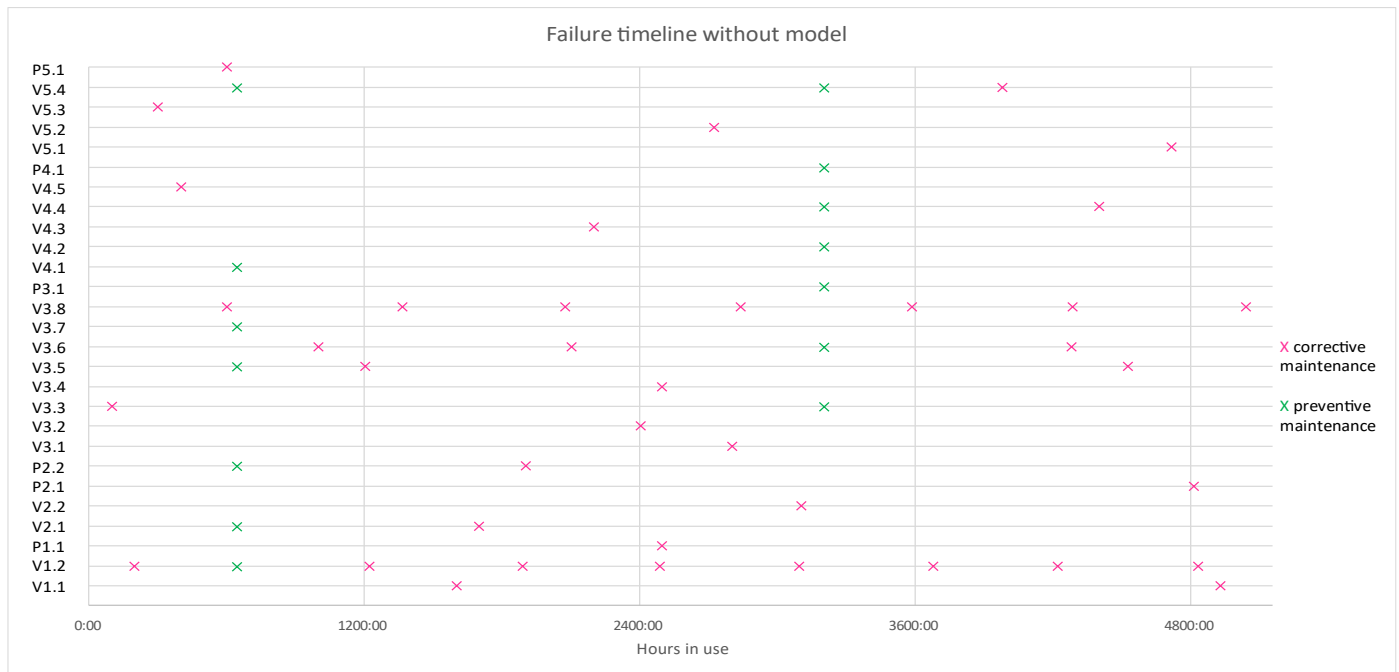


Figure 42: Overview of failure and maintenance without model

In [Figure 42](#), the pink crosses indicate the times when a component failed. The green crosses indicate preventive maintenance performed during the stop weeks.

Furthermore, there are 38 unscheduled downtimes.

14 components receive maintenance during stop weeks.

A total of 57 hours of maintenance were carried out. Of these, 22 hours were performed during scheduled downtime hours. 14 hours of this during stop weeks during which preventive maintenance is performed and 8 hours of maintenance performed during a planned stop, see [Assumption 6.2](#). The remaining 35 hours of maintenance were performed during unscheduled downtimes, see [Figure 43](#) for the distribution of the maintenance hours.

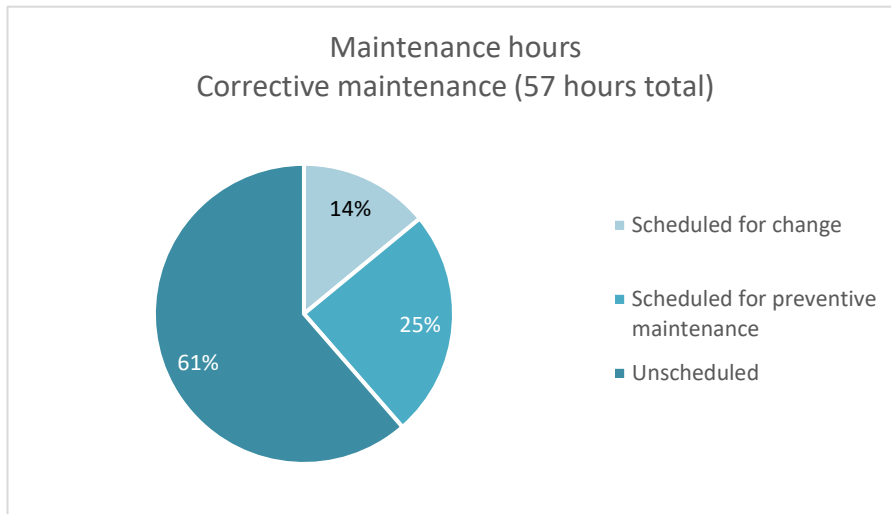


Figure 43: Pie chart of maintenance hours, without model

The cost analysis shows that the total maintenance cost comes to €49,120.-. This consists of three cost items, see [Table 7.1](#) for the cost items and associated amounts.

Table 7.1: Maintenance costs without model

Cost of spare parts	€ 12,720.00
Salary	€ 8,400.00
Cost of unplanned DT	€ 28,000.00
Total	€ 49,120.00

7.2 With model

The model has been run for the entire synthetic schedule. Based on this, it can be seen that indeed, there is often an indication from the PL and LS values an hour or 2 before failure. In [Figure 44](#) and [Figure 45](#), this is illustrated.

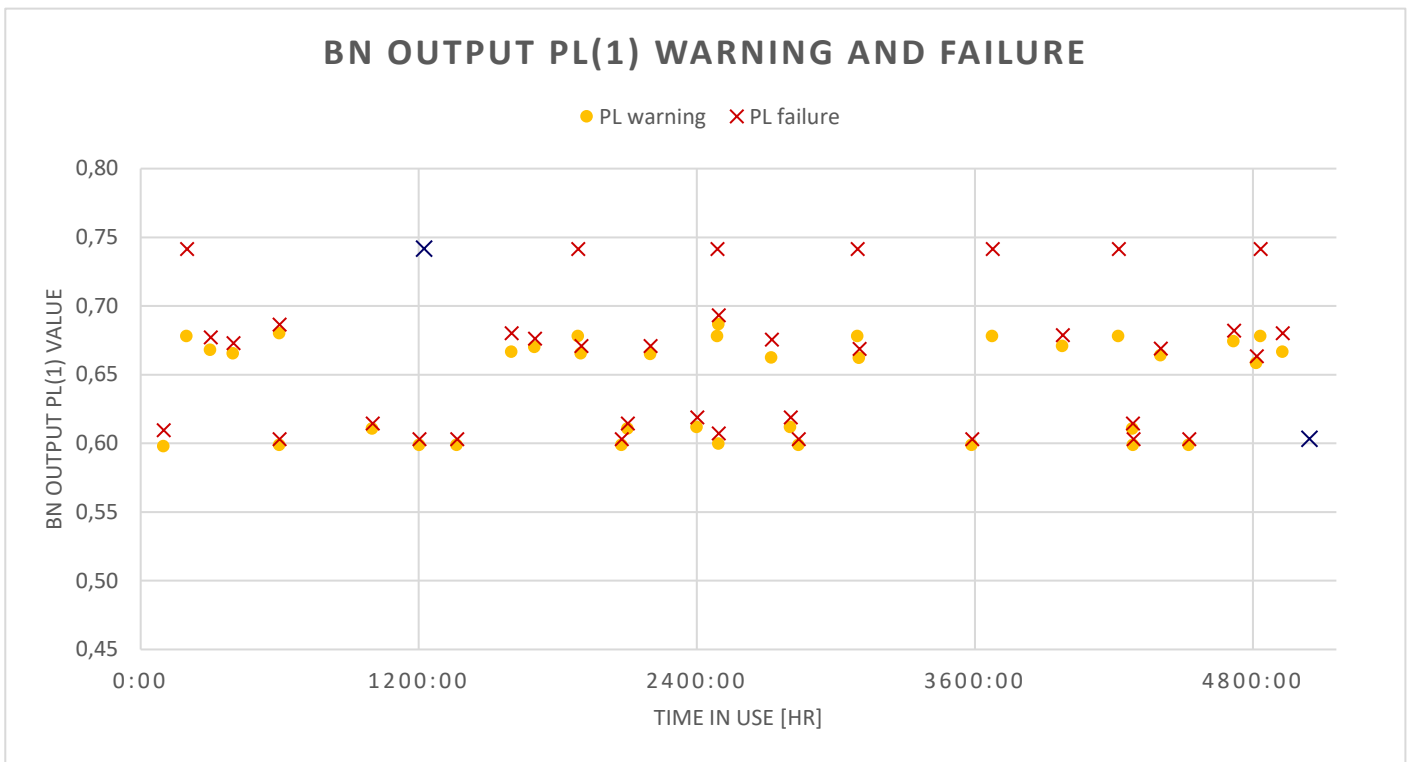


Figure 44: BN output PL(1) for warning and failure values

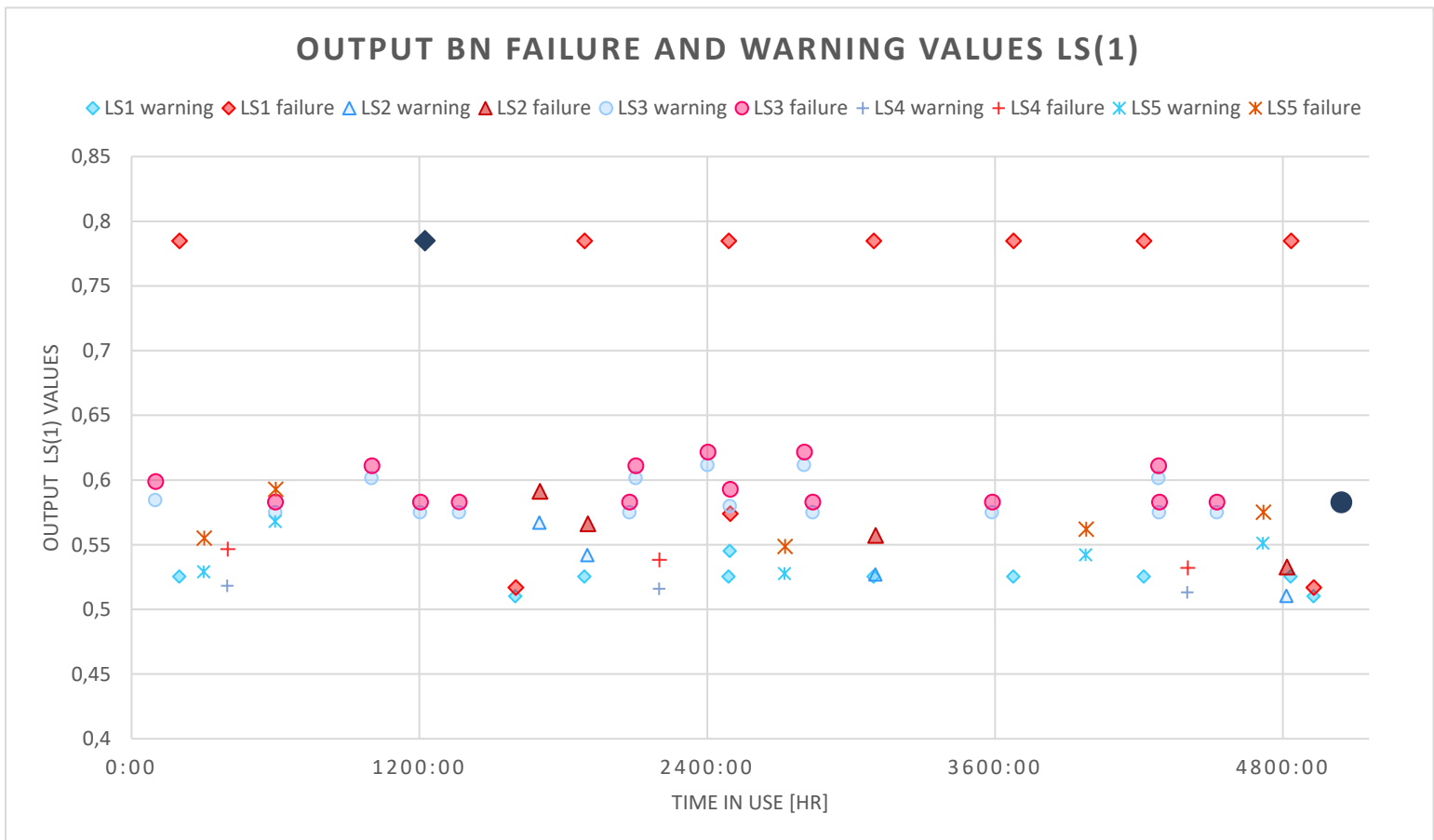


Figure 45: BN output LS(1) for warning and failure values

Figure 44 and Figure 45 shows that in most cases a warning value was detected before there was a failure. However, on the dark blue markings, there was no warning but still a failure. Corrective maintenance was therefore carried out here.

Furthermore, an overview was given of how much time of the MTBF had elapsed at the time maintenance was performed. This overview can be found in Figure 46. It is in order of when which component had maintenance.

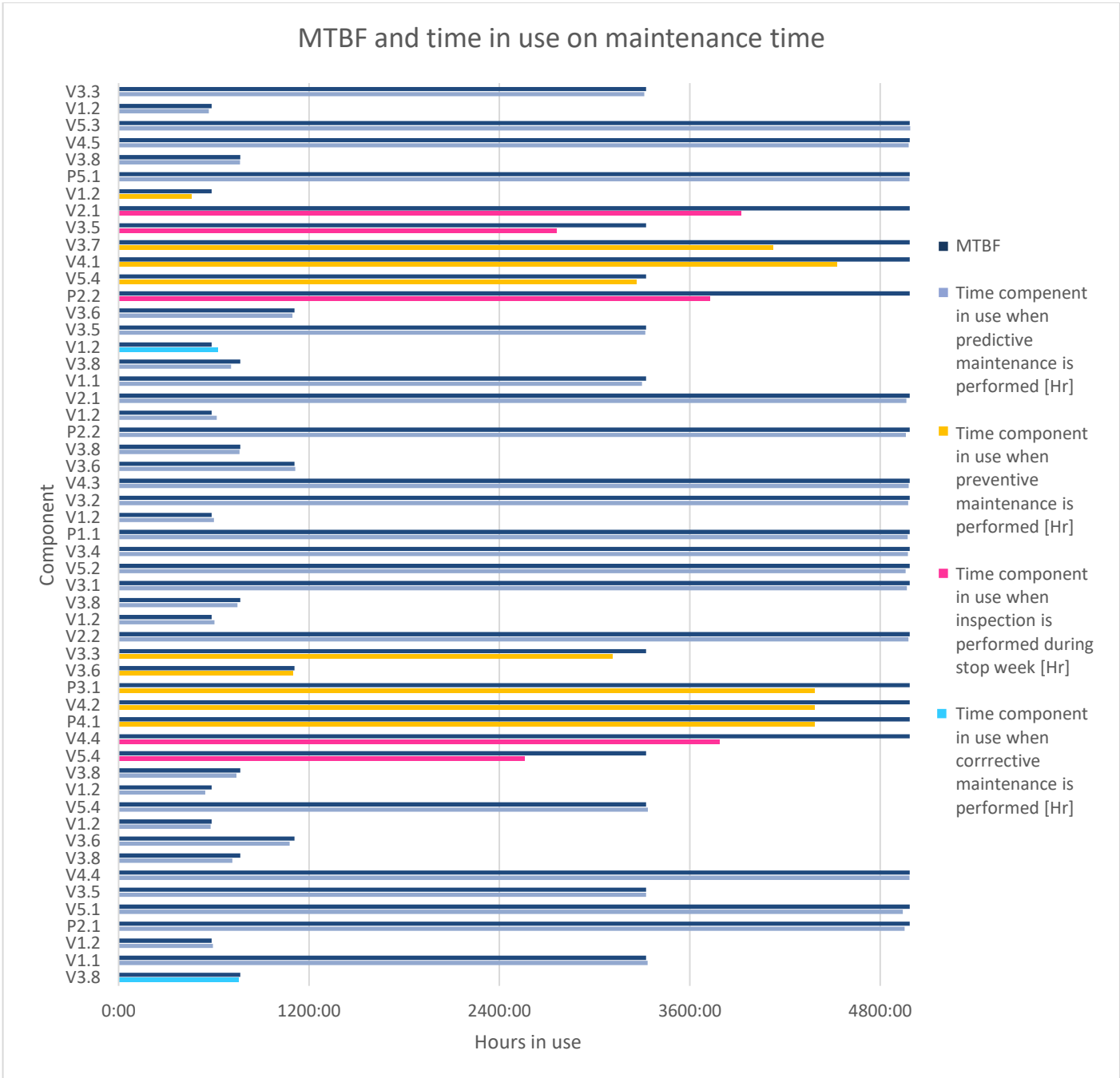


Figure 46: Overview of time used by every component on moment of maintenance

The orange/ yellow bars indicate that preventive maintenance was performed on that component and then the time in use was reset. The pink bars indicate that a component had an inspection at that point only because 75% of the MTBF had elapsed at the time of the scheduled stop week. The bright blue bars indicate the time of components that have had corrective maintenance.

Furthermore, using the model gives the following results. There are 2 unplanned downtimes.

14 components receive maintenance during stop weeks. The rest of the maintenance takes place during scheduled downtime hours as much as possible, otherwise Assumption 6.2 and Assumption 6.3 apply. A total of 57 hours and 50 minutes of maintenance will be performed. This is 37 hours during scheduled downtimes. 14 hours during stop weeks during which preventive maintenance is performed and 6 hours and 50 minutes during unplanned downtimes, see Figure 47.

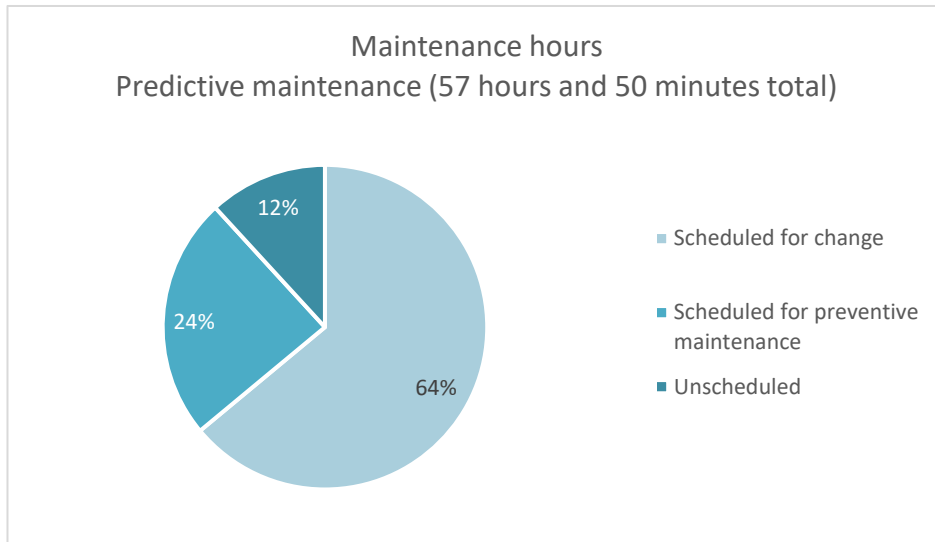


Figure 47: Pie chart of maintenance hours with model

The cost analysis shows that the total maintenance cost comes to €23,120.-. This consists of three cost items, see [Table 7.2](#) for the cost items and associated amounts.

Table 7.2: Maintenance costs with model

Cost of spare parts	€ 12,720.00
Salary	€ 8,400.00
Cost of unplanned DT	€ 2,000.00
Total	€ 23,120.00

7.3 Comparison

Comparing the use or non-use of the model shows that almost the same number of hours were used for maintenance. There is a big difference between the number of hours used for maintenance during unplanned downtimes. If the model is not used, 35 hours of unplanned downtimes are needed for maintenance, compared to 6:50 hours if the model is used. The difference is clearly visible in [Figure 48](#).

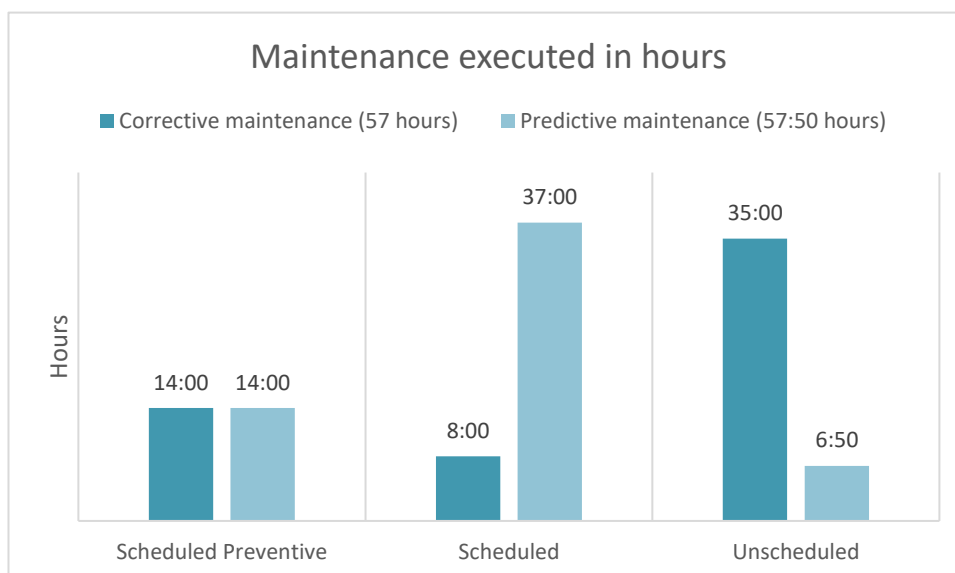


Figure 48: Comparison of maintenance hours with and without model

When looking at costs, component and labour cost items are the same in both cases. The difference is in the costs that cause unplanned downtimes, see [Figure 49](#).

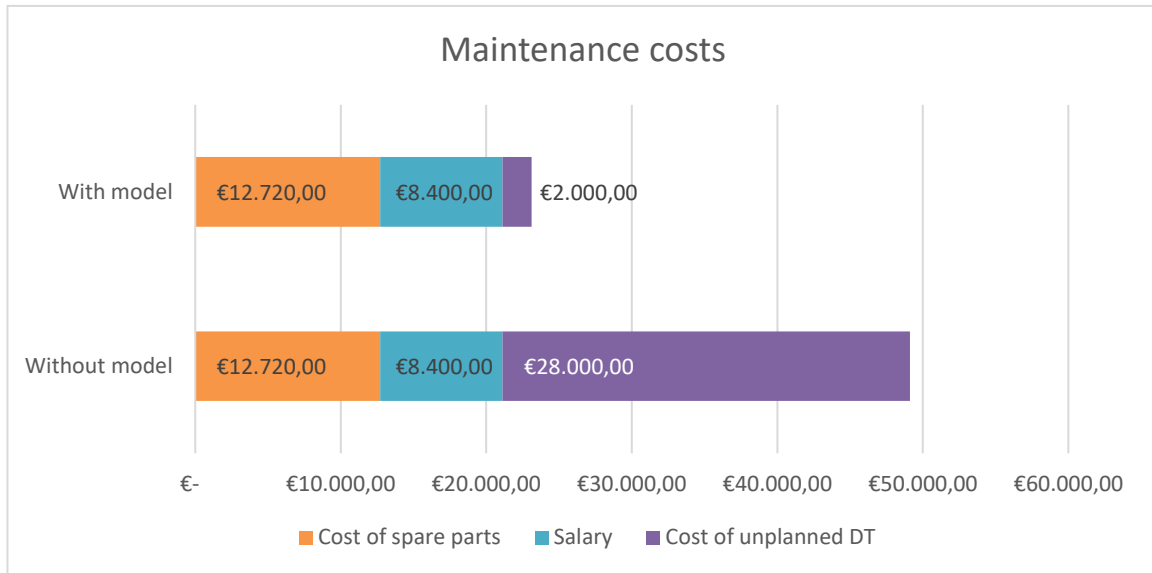


Figure 49: Comparison of maintenance costs with and without model

Instead of looking at the specific cost items, as in Figure 49, it is possible to look at how costs relate to different maintenance strategies. For this, maintenance costs are split into costs for corrective maintenance, preventive maintenance and predictive maintenance. This then gives the result as visible in Figure 50. The costs are then for replacement parts, wages and unplanned downtime costs, but divided by the type of maintenance performed.

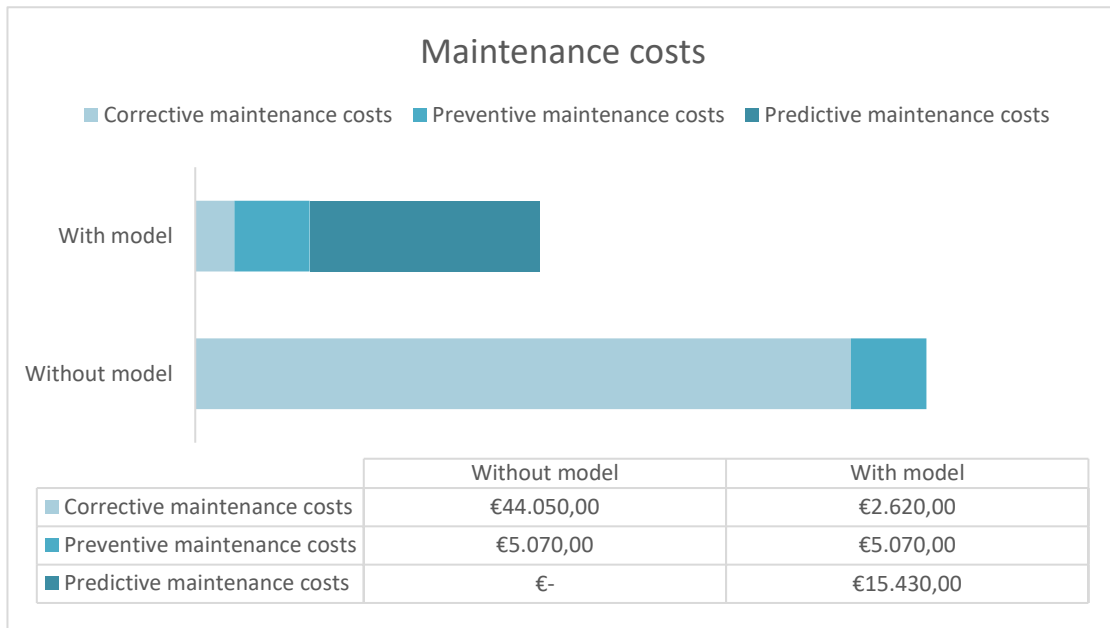


Figure 50: Comparison of type of maintenance costs with and without model

Calculating the KPIs further requires the total number of hours of downtimes. This involves adding the number of hours of planned downtime to the total number of hours of unplanned downtime. As indicated earlier when validating the synthetic data, in this research it is that there are 518:25 hours of planned downtime.

In the case of not using a model, 35 hours of unplanned downtime should be added to this. The total downtime in the case of no model comes to 553:25 hours.

In case the model is used, 6:50 should be added. This brings the total downtime using the model to 525:15 hours.

7.4 Results KPIs

The results found earlier can be used when completing the KPIs.

7.4.1 Results MDI KPI

For the KPI where the MDI is determined, it is important to know how much time of the total downtime was used for planned maintenance. Total downtime should be calculated by adding planned downtime to unplanned downtime.

Next, the number of hours of the downtimes used for planned maintenance should be determined. After that, [Eq. 6.4](#) can be entered. This will show the results in [Table 7.3](#).

Table 7.3: Comparison of maintenance hours with and without model

	Without model	Time used for maintenance	With model	Time used for maintenance
Scheduled for change and PdM	318:25:00	8:00:00	318:25:00	37:00:00
Scheduled for preventive maintenance	200:00:00	14:00:00	200:00:00	14:00:00
Unscheduled	35:00:00	35:00:00	6:50:00	6:50:00
Total downtime	553:25:00		525:15:00	
Maintenance during planned DT		22:00:00		51:00:00
MDI		3.98%		9.71%

The MDI when the model is not used is calculated with [Eq. 6.4](#). Entering the hours here gives:

$$MDI_{without\ model} = \frac{22:00\ hours}{553:25\ hours} * 100 = 3.98\%$$

This can also be done when the model is used. [Eq. 6.4](#) will then be like:

$$MDI_{with\ model} = \frac{51:00\ hours}{525:15\ hours} * 100 = 9.71\%$$

Looking at the relationship given earlier, see [Chapter 6.2.1](#), between MDI and reliability, it can be seen that the reliability in the case that the model is used is better than when the model is not used. This is due to the fact that the number of hours of downtime used to perform planned maintenance in the case the model is used is 29 hours more than the situation without the model, see [Table 7.3](#). In addition, the total number of hours of downtime is 28 hours and 10 minutes less if the model is used. The difference can be seen in [Table 7.3](#) in the number of hours of unscheduled maintenance. The results of the MDI show that at the time the model is used, more scheduled downtime hours are used to perform maintenance. This results in fewer hours being required for unplanned maintenance and hence fewer hours of unplanned DT. Since there are fewer hours for unplanned downtime, it can be concluded that the reliability of the production line improves when the model is used compared to the situation when the model is not used.

7.4.2 Results cost KPIs

Then there are the three KPIs related to costs. For this, the results given earlier can be used from [Figure 50](#). Furthermore, it can be extracted from [Table 7.1](#) that the total maintenance cost without model comes down to €49,120.-.

From [Table 7.2](#) it can be extracted that the total maintenance cost with model comes down to €23,120.-.

Then, if [Eq. 6.16](#), [Eq. 6.17](#) and [Eq. 6.18](#) are filled in, it gives the results as shown in [Table 7.4](#).

Table 7.4: Results cost KPIs

KPI	Without model	With model
E15 (corrective maintenance)	89.68%	11.33%
E16 (preventive maintenance)	10.32%	21.93%
E17 (predictive maintenance)	0.00%	66.74%

[Chapter 6.2.2](#) describes the relationship between reliability and maintenance costs. Building on that, several things can be concluded from the results. First, the total maintenance costs are the lowest when using the model. There is a difference of €26,000.- in total maintenance costs. Furthermore, it can be seen from [Figure 50](#) that there are no predictive maintenance costs in the situation without the model, this is logical as the situation without the model only uses preventive and corrective maintenance. It can also be seen from [Figure 50](#) that the preventive maintenance costs are the same in both situations. However, because the cost KPIs are divided by the total maintenance costs, this is not reflected in [Table 7.4 KPI E16](#). Looking at the relationship between maintenance costs and reliability, in the situation where the model is used it can be seen that the low total maintenance costs ([Figure 50](#)) and that the highest cost item is in predictive maintenance costs, see [Table 7.4 E17](#), contribute to improved reliability. More components receive maintenance before they can fail. In addition, the cost item of preventive maintenance ([E16, Table 7.4](#)) is lower than predictive maintenance ([E17, Table 7.4](#)) so it can be observed that few components get maintenance too early. There

is also low corrective maintenance cost percentage (E15, Table 7.4) indicating that only a few components have caused unplanned downtime. All this implies that the components get a better MTBF and can perform their function for longer, resulting in improved reliability of the production line in case the model is used.

7.5 Conclusion

In this chapter, following on from Chapter 6, the sub-research question : *In what way is the implementation of the model contributing to the reliability of the production line?* has been explored in more detail.

To obtain the results between using and not using the model, a synthetic dataset was created based on the historical production schedules. Next, the results were examined to see what the results would be if the model was not used. This showed, for the results without the model, that 57 hours of maintenance were performed, 61% of which caused unscheduled downtime hours and cost €49,120.-. If the model is used then a total of 57 hours and 50 minutes of maintenance was performed, of which 12% caused unplanned downtime hours. The maintenance cost comes out to €23,120.-. During the comparison, it emerged that the difference is made in the unplanned maintenance costs. After that, results of the KPIs were looked at. This shows that the MDI KPI is better if the model is used. $MDI_{with\ model}$ comes out at 9.71%, versus $MDI_{without\ model}$ at 3.98%.

The cost KPIs reveal that there are large differences in the outcomes of E15 and E17. As stated earlier in Chapter 6.2.1, a higher MDI is a good indication that the reliability of the line is improving. In this case it will mean that the model does a good job of improving reliability compared to the situation when the model is not used.

Looking at costs, Chapter 6.2.2 had stated that an indication of improved reliability is that there are lower total maintenance costs and there should be a higher percentage of predictive maintenance costs. In addition, there should be a low percentage of corrective maintenance costs. Looking at total maintenance costs, there is a difference of €26,000.- here, see Figure 49. In the situation without the model, the largest cost item is corrective maintenance costs, which in this case results in a high percentage for the E15 KPI. In the situation where the model is used, there is a clear shift from corrective to predictive maintenance costs. The percentage of corrective maintenance costs (E15 KPI) is still only 11.3%, see Table 7.4. The predictive maintenance cost (E17 KPI) is 66.7%, see Table 7.4. The lower maintenance costs and the higher percentage of predictive maintenance costs imply that line reliability improves when the model is used.

In short, in answer to the sub-research question, it can be stated that implementing the model contributes to the improvement of reliability because fewer unplanned downtime hours are needed to perform maintenance. In addition, in the situation where the model is used, there are lower total maintenance costs and the costs that are there come from predictive maintenance actions. Earlier it had been stated that these are indications that the reliability of the components and hence the reliability of the production line is improving.

8. Conclusion

This research focused on the main question:

How can a predictive maintenance strategy contribute to improving the reliability of a soft drink production line?

The soft drink production line in this research consists of a collection of valves, pumps, tanks and pipes to mix syrup and water. To make this into soft drinks, carbonation is often required. In this research, the focus is on the butterfly valves, double seat valves and centrifugal pumps within the production line. These are the components that cause many unplanned downtimes to be experienced in a Dutch soft drink factory. One of the causes of the unplanned downtimes can be found in the corrective maintenance strategy, falls under the reactive maintenance strategies, that is applied. This involves replacing or repairing components only when they have failed and often caused unplanned downtimes. A solution that has emerged in recent years is a predictive maintenance strategy, this falls under proactive maintenance strategies. This involves looking at how condition monitoring data and historical data can be used to create a model that predicts whether a component needs maintenance or not. There are several methods that can be used for this purpose. Based on established criteria for this research, it was determined that a Bayesian Network (BN) is the most appropriate method for the model to be developed in this research. The model should provide insight into how a predictive maintenance strategy can contribute to improving the reliability of the soft drink production line. Reliability in this research is defined as the time a component can perform its requested function under known conditions.

The predictive maintenance strategy used for this research was created based on historical data. This created a BN that can be used to determine the probability of unplanned production line downtime. In this BN, valves and pumps are represented as components in the production line. To determine the condition of the valves, the running time of each valve is used. This is coupled with a logistic function that can be used to determine the probability of failure of each individual valve. For the pumps, the probability of cavitation is used, which can be determined with the NPSH margin. It also looks at what percentage of the best efficiency point (BEP) the centrifugal pumps are operating at. Indeed, this can be linked to the failure probability for the pumps.

The production line in the BN, for this research, is divided into five line segments. The looptijd, probability of cavitation and at what percentage of BEP are the parameters that indicate the condition of the valves and pumps. This can then be compiled into the different line segments. In each line segment, there are some valves and at least 1 pump. Based on the conditions of the valves and pumps, the probability of failure of each line segment can be determined. If this is known then the probability of failure of the entire production line can be determined. All this is done by entering evidence for looptijd, cavitation probability and percentage of BEP. Forward inference can then be used to determine the failure probability of the node for the entire production line.

Besides the BN, the mean time between failure (MTBF) is also used in the model. The MTBF is determined based on the historical failures of each individual component. The historical production schedule was used as the basis for the synthetic production schedule. Then, based on that planning, a synthetic dataset was also created based on the historical process data. With these datasets, the flowchart can be followed. For predictive maintenance, the model is used to run the synthetic data through the model and flowchart before a process starts. From this, it then follows whether maintenance is needed or not. Based on this, the operator can make a decision on whether to start production or not, or whether maintenance should be performed first. By using the outcomes based on the flowchart to perform timely maintenance, reliability can be improved. To demonstrate this, four key performance indicators (KPIs) are looked at and calculated for two situations. The first KPI looks at the ratio of downtime hours used for scheduled maintenance to total downtime hours. This KPI is known by the maintenance downtime index (MDI). The second, third and fourth KPIs look at the cost of corrective, preventive and predictive maintenance relative to total maintenance costs. These KPIs can be qualitatively related to reliability. The higher the MDI the better the reliability of the line. More downtime hours are then used for planned maintenance, relative to the total number of downtime hours. For the cost KPIs, the reliability of the line improves if there is less cost for corrective maintenance and more for predictive maintenance.

In the first situation, the synthetic datasets are used to determine what happens when the model is not used. This is also called the current situation. Here, only the corrective maintenance and preventive maintenance are used, which takes place during the scheduled stop weeks. For this, the outcomes for the various KPIs are determined. The second situation is when the model is used. With this, a predictive maintenance strategy is then applied. The KPIs are also calculated for this.

After this, the differences between the two situations were examined. What stands out here is that in situation with the model, 50 minutes more time is needed to perform maintenance compared to if the model is not used. However, only

24% of DT time is needed for unplanned DTs. This contrasts with the situation where the model is not used, where 61% goes on unplanned DTs. Looking at the first KPI, the MDI, there is a difference of 5.73% in favour of the situation where the model is used. This is the first indication that the reliability of the production line improves if the model is used, as more hours of total DT are used for planned maintenance.

Next, the costs were looked at. Here, the first thing that stands out is the difference in total maintenance costs. In the situation that the model is not used, the total maintenance costs are €49,120.-. This is €26,000.- more than in the situation where the model is used, where the total maintenance costs are €23,120.-. In other words, in the situation where the model is used, the total costs are 47% of the total maintenance costs of the situation without the model.

The gain achieved by the model here can be traced back to the reduction in corrective maintenance costs, the second KPI. In addition, in the situation where the model is used, the value of the fourth KPI, namely the proportion of total maintenance costs used for predictive maintenance, increases. This KPI points in favour of the situation where the model is used. With this, the final indications are in indicating that the reliability of the line is being improved.

In short, it can be concluded that based on this research, the predictive maintenance strategy contributes to improving the reliability of the production line. This is based on that the MDI is better in the case where the model is used compared to the situation where the model is not used. In addition, there is a big difference in total maintenance costs and cost KPIs pointing in favour of the situation with the model. All this suggests that the number of unplanned DTs can be reduced if the model is used. Because of this reduction, it can be qualitatively inferred that the reliability of the line improves.

In short, in answer to the main research question, a model for a predictive maintenance strategy was created in this study. A Bayesian Network is one of the methods used in this model. The model is able to determine in advance of a production based on the type of product and the duration of the production whether a component in the production line could possibly fail. Based on the results from the model, a decision can be made on whether or not to execute maintenance. Two situations were compared, one where the model is not used and one where the model is used, to determine how the model with the predictive maintenance strategy contributes to improving the reliability of the soft drink production line. It emerged in this study that when the model is used, fewer unscheduled maintenance hours are required, resulting in fewer unplanned downtime hours. This provides the first indication that the model contributes to improving reliability. In addition, total maintenance costs are much lower when the model is used. Furthermore, most of the maintenance costs are incurred for predictive maintenance. Thereby, the costs for corrective maintenance have decreased tremendously when the model is used, fewer components have needed corrective maintenance. This indicates that the reliability of the components and thus the production line improves when the model is used with the predictive maintenance strategy. The reduction in unplanned downtime hours and lower maintenance costs reflect this.

9. Discussion

Within this research, an accessible way to increase the reliability of components in a production line based on condition monitoring data was investigated. In the specific case of the soft drink industry.

Based on theory and available data, a Bayesian network was chosen as method for the model in this research. This method makes it easy to add nodes if more parameters are needed. Moreover, it gives a graphical representation of the system and which nodes affect each other.

There are other techniques such as an ANN that can give more reliable results. However, more knowledge must then be available about the components, influences, behaviour and lifetime. ANN can only be used if there is a large dataset to build, validate and verify the model.

Chapter 2.4 mentioned some techniques for monitoring the condition of valves and pumps. During the case study it was found that there was not enough data available to use most of these techniques. This resulted in the condition of valves now being dependent on only one parameter, namely looptijd. Despite the use of different types of valves, mixproof and butterfly valves, it was chosen to approach all valves in the same way in this research. This was done with the idea that both types should both open and close, albeit in different ways.

Even less data was available for the pumps. It was chosen to generate the pump data with a random data generator based on theory and the manufacturer's pump curves. As a result, much of the model is based on assumptions. To get a more accurate and better picture of the condition of the components in the pipe, more and different parameters need to be added to the model. These parameters need to be trained with data collected on the specific line.

Furthermore, more research needs to be done on the lifetime of the pumps and valves to give an accurate value to this. This would give a better understanding of the specific reliability of the components and the production line.

Maintenance records are also essential. In this research, it was assumed that valves and pumps only receive maintenance where parts are replaced. No valve or pump was completely replaced. In addition, it was assumed that after performing maintenance, the operating time is restored to 0 and 100% of the MTBF can be used again. However, it is not entirely realistic. If a valve or pump only gets certain replacement parts, it is more plausible that there is less than 100% of the MTBF left until the next time a part will fail. Although, due to the lack of data on this, the 100% MTBF rule after maintenance has been carried out was still used for this research. When more data is available on the failure of components and the influence of certain components on the MTBF, this can be adjusted.

For the BN, only the valves and pumps are included in the line through which liquid flows. The valves through which gas flows were not included in this research. Moreover, the pipeline section to be analysed was divided into 5 sections. It was taken into account that there are at least 1 pump and some valves in each pipe segment. Nevertheless, the differences in the number of components are large, this can be clearly seen in line segment 3. Here there are 8 valves and 1 pump. Here there are 8 valves and 1 pump. It is therefore difficult to define appropriate prior probabilities for this in the CPT of LS3. This is because there are 512 possible combinations of the valves and pump states.

In the BN, all components are assumed to be independent of each other. If V1.1 fails, it does not necessarily mean that V1.2 also fails. However, it happens that multiple components fail simultaneously, often caused by water hammer. The model currently does not consider interrelationships between the different components and the different line segments.

The BN is modelled in Python with the package PGMPY. Here, it is important to define all nodes correctly and to make assumptions about the prior probabilities of the nodes. However, this has the consequence that, as mentioned earlier, for line segment 3, a very large number of values have to be entered, namely 512. This makes it difficult to check all these values.

The model also has a limitation in the number of values that can be entered as proofs. In this case, 32 evidence values can be entered, but in the model for this research, there are 35 nodes for which it is desirable to enter evidence. For this research, it was chosen to see which nodes have the least impact on the failure probability of a line. The 3 nodes that emerge from this analysis with the least influence are considered nodes for which no evidence is added. This affects the overall failure probability given the nodes for which evidence is introduced.

Otherwise, only verification of the BN was done. It was not possible to perform a validation. Validation should be done with existing data coming directly from the system. In this study, there was no condition monitoring data available from the production line in the case study on the parameters in the BN to do the validation.

To arrive at the results, synthetic datasets are used. The first is the synthetic production planning. This is created based on the historical production schedules. For this, it has been assumed that the processes in the past are the same processes that should cross the line in the near future. The synthetic process dataset is based on the historical process data. Here, the assumption is that the past processes produce almost the same data for the future processes. However if a new product or other type of CIP is sent across the line then data has to be collected before it can be included in the model.

For the pumps, the data were derived from the pump curves and randomized. As more data becomes available, the model has to be adjusted to the new values. This is especially true for the pumps, but also for the logistic functions of the valves. If the mean changes, the values of k and m become different. This affects the failure probability of the valves.

Looking at the results in [Chapter 7.1](#), situation where the model is not used, a number of things stand out. These include when comparing the failure moments in [Figure 42](#) with the previously defined synthetic dataset based on the MTBF, see [Appendix I](#). As soon as no model is used, there are a number of components that operate longer than the MTBF found earlier. This is the case for V3.3, V3.6, V3.7, V4.1, P4.1 and V5.4. The reason for this is that these components were given timely MTBF. One reason is that these components received timely maintenance during the stop weeks when preventive maintenance was performed. This is consistent with the explanation in [Chapter 6.1](#) that reliability increases as the MTBF of components gets longer, in other words as the time between failures of a component gets longer.

However, there are also components that fail more often, such as V3.8. In addition, there are also components that continue to fail as often despite performing preventive maintenance. This is the case for components V1.2 and P2.2. For these components, preventive maintenance does not improve reliability. If a component subsequently fails despite preventive maintenance, double costs arise. First the cost of preventive maintenance and then the cost of corrective maintenance. However, this will only show up in the cost KPIs and total maintenance costs. In the MDI, this will have no impact. After all, if fewer hours of preventive maintenance had been done, the total number of hours of downtime would also have been lower, the percentage remains the same.

When looking at the results in [Chapter 7.2](#), situation where the model does get used, there are by some striking things. First of all, it can be seen from [Figure 46](#) that few components complete the entire MTBF time. Most of the components receive maintenance earlier. There are some components that could have undergone many more hours of production during the stop week when preventive maintenance is performed. Furthermore, [Figure 46](#) shows that component V1.2 lasted longer than its MTBF once but caused unplanned downtime as a result. Then again, the latest V1.2 shows that maintenance was only performed after the MTBF had expired without causing any unplanned downtime. This may be due to the length of productions. In the former case, it could be that the production took a few hours too long and the choice was made to risk it, with the result that one time it went well and another time it did not.

Looking at the results of that come from the KPIs, a few things stand out. First of all, there is a difference of 50 minutes in the total number of maintenance hours. This is because in the situation where the model is used, some planned DT hours were just not long enough to perform the maintenance. The time needed for maintenance outside the scheduled DT hours were added to the unscheduled DT hours.

When looking at the total DT hours for both situations, there is a difference of 28 hours and 10 minutes. The reason for this is due to the difference in hours of unscheduled DT.

The total number of DT hours is important for calculating the MDI. Since there are very many more hours of DT scheduled than maintenance, the percentages that come from the MDI calculations are very low. If the MDI calculation is approached slightly differently and only the ratio between the number of hours of DT used for planned maintenance versus the total number of hours used for maintenance is considered, the following results follow. For the situation where the model is not used, a total of 57 maintenance hours are required, see [Table 7.3](#). 22 hours of this is used for maintenance during scheduled DT hours. Entering this in [Eq. 6.4](#) then gives:

$$MDI_{without\ model} = \frac{22:00\ hours}{57:00\ hours} * 100 = 38.6\%$$

In the situation when the model is used, there is a total of 57 hours and 50 minutes of DT for maintenance, see [Table 7.3](#). 51 hours of this is maintenance during scheduled DT hours. Entering this in [Eq. 6.4](#) then gives:

$$MDI_{with\ model} = \frac{51:00\ hours}{57:50\ hours} * 100 = 88.2\%$$

This already gives a much bigger difference for the MDI outcomes. This does make it clearer that for the situation where the model is used, most of the maintenance hours can be performed during scheduled DTs. As more scheduled DT hours are used to perform maintenance, there are fewer hours of unscheduled DTs.

It should still be noted that an assumption has been made about the time it takes to perform maintenance on a valve or a pump. If the repair time is much longer than estimated beforehand, the number of hours of unplanned DT will increase further and so will the total number of hours of DT. As a result, the MDI value may then be lower than it is currently. This then applies to both situations because the same hours were charged for carrying out maintenance within this research for both situations.

Looking at the cost KPIs, it should be noted that the total maintenance costs of the situation where the model is used are 53% lower than the total maintenance costs in the situation without the model. This creates a slightly distorted

picture when determining Eq. 6.16, Eq. 6.17 and Eq. 6.18. Nevertheless, it still clearly shows that in the situation where the model is used most of the costs go to predictive maintenance and that in situation without the model most of the costs go to corrective maintenance.

Finally, it should be noted that the model cannot prevent unplanned DTs. The model will only help reduce the number of unplanned DTs. The results also show that there are 2 components that still cause unplanned DT.

In short, it can be concluded that a model has been created in this research based on literature, theories and assumptions. This is a step towards providing insight into what a predictive maintenance strategy can do for reliability, with a focus on a soft drink production line. However, there are still snags in the model and research. Further research and data collection is therefore needed to validate the operation of the model in this study and then to be able to use it in the soft drink industry.

10. Recommendations

As mentioned earlier in the discussion, one of the biggest uncertainties in this research is that little data was available to use multiple parameters for the valves. In a follow-up research, it is advisable to look at the difference in behaviour of a butterfly valve and a double seat mixproof valve.

Now, looptijd has been used to determine whether a valve is still functioning or not. However, [Chapter 2.4.1](#) mentions other techniques that can help with valve condition monitoring. Some of these are valve torque, acoustic monitoring or valve flow coefficient. In addition, the pressure build-up on the valve can also be monitored. This can be used to determine at which processes and pressures the valves suffer.

For pumps, the choice has now been made to work with the NPSH margin and the BEP. However, there are also techniques available for pumps such as vibration monitoring, lubricant sampling or measuring energy consumption, see [Chapter 2.4.2](#) for more possible techniques. This adds other techniques to map the whole spectrum of possible failures of centrifugal pumps. It is important that in a follow-up study the various techniques are examined in more detail so that a choice can be made as to which combination of techniques is the best combination to represent the condition of the valves and pumps.

In addition, no data were available on pump condition monitoring. To improve the model and results, ways in which these data do become available should be explored. It is advisable to actively work on this in the future to get more out of the results of the model.

For the case study, no data was available on how maintenance was carried out in the past. It is therefore not known whether only certain components in a valve or pump were replaced or whether a whole new valve or pump was installed. In order to form a complete picture in the future of the influence of the various processes that cross the line, it is necessary to keep track of which components have been serviced to which. In this way, more insight can be gathered about the influence of the processes, the lifetime of certain components and thus a model can be refined. By combining this data with both production planning and data from sensors that can be placed that can monitor condition, process-specific information can be gathered. For example, which processes were at which temperatures and how is this reflected in the condition monitoring data up to the time of failure of the component. This would make it possible to determine component degradation based on multiple factors. A follow-up study could therefore reveal which factors and values should be considered essential for predicting maintenance.

In the situation that more data, as described earlier about maintenance data and other condition monitoring parameters, and knowledge becomes available then more complex models such as an ANN can also be explored. Knowledge here refers to more insight about the influences of processes on component parts. For example, what the degradation looks like of a valve when multiple production and cleaning processes at different temperatures have gone through it. Or what happens to the O-ring the moment the line has been idle for a long time.

Another option, if there is more data and knowledge available, that then also comes within reach is the creation of a digital twin of the production line and the behaviour of all components belonging to the processes passing over the line. This will make it possible to make more accurate predictions about the expected life of the components on the line based on which processes are expected to occur in the future. This makes it possible to align both production and maintenance planning.

This also makes it possible to see what other factors, such as production quantity/loss or maintenance costs, should be tracked and taken into account when considering when to carry out maintenance.

For follow-up research, there are still many possibilities that can be explored in terms of the degradation processes of the different components, what data should be collected and how this data should be interpreted. Research can also be done into what other ways of modelling then come within reach.

In this research, it was chosen to work with the failure probability, this was done because there was no reliable data available on the lifetime of the components. In a follow-up research, however, it would be better to work with the reliability/survival function. By linking it to the degradation of service life, a more accurate time factor can be given to the model.

For this research, on the advice of the company, the model did not distinguish between the importance of the components. If it turns out, however, that certain components are more important, for example because the delivery time is longer or because one component affects the other components, then weighting factors can be used. By adding weighting factors, this can also play a role in decision-making.

Within this research, the Python PGMPY package was used. It was found that a limited number of evidence values can be used. In the future, it is advisable to use a professional software package specifically for calculating and analysing a BN. This will prevent certain nodes from not being able to be used.

In addition, many assumptions about cost and duration of repairs have been made in this research. To achieve accurate results, it is important to track what a repair really costs and how long it takes. Even if a component has to be completely replaced, this can be included. This research only looked at the cost of mechanics. It did not look at what the cost of a downtime is in terms of delays and overtime of other staff. It is recommended that a follow-up research does include these costs further and use them in decision-making.

Also, this research used the assumption that all parts are always in stock and can be used immediately when maintenance is required. However, in today's world, it is not entirely plausible that all parts are always in stock and sometimes it will be necessary to wait for a part. In a follow-up research, it is important to look at the costs and production loss hours this entails.

Finally, it should be stated that the results in this research are largely based on assumptions. Much data and research is still needed to demonstrate whether the model actually achieves the results currently obtained from the synthetic datasets. A follow-up research can use the way of thinking behind the model and see if it can be applied to other production lines or maybe even in other industries.

Literature sources

- [1] Zaat, J.. (2019, October 23). 'Dodelijke frisdrank' en het gebrek aan een suikertaks | NTVG. NtvG. <https://www.ntvg.nl/artikelen/dodelijke-frisdrank-en-het-gebrek-aan-een-suikertaks>.
- [2] Magar, S. (2021, June 20). *Microbial spoilage of Beverages and their preservation*. Microbenotes. <https://microbenotes.com/beverages-spoilage-preservation/>
- [3] Ryan, R.. (2014). Safety of Food and Beverages: Soft Drinks and Fruit Juices. *Encyclopedia of Food Safety*, 3, 360-363. <https://doi.org/10.1016/B978-0-12-378612-8.00296-1>
- [4] Emmins, C. (1991). *Soft Drinks: Their Origins and History*. Dublin, Ireland:: Shire.
- [5] Schweppes Holdings Limited | Royal Warrant Holders Association. <https://members.royalwarrant.org/company/schweppes-holdings-limited?directory=1>
- [6] A Time For Innovation - Can Manufacturers Institute | Washington, DC. Cancentral. <https://www.cancentral.com/can-stats/history-of-the-can/time-innovation>
- [7] Steen, D., & Ashurst, P. R. (Eds.). (2008). *Carbonated soft drinks: formulation and manufacture*. John Wiley & Sons.
- [8] Medicinal Soft Drinks and Coca-Cola Fiends: The Toxic History of Soda Pop | Collectors Weekly.
- [9] Koss, J. P. (2005). "THE" PRODUCTION LINE. *Beverage World*, 124(1746), 40. <https://www.proquest.com/trade-journals/production-line/docview/213290219/se-2>
- [10] Lumix. (2020, May 14). *Importance of Predictive Maintenance in Food Processing Industry | lumixfps*. Lumixfps. <https://www.lumixfps.com.au/blog/importance-of-predictive-maintenance-in-food-processing-industry/>
- [11] Mccleery, A. (2022, July 31). *Non-alcoholic beverage producers experience surge in unplanned downtime as production slows - Food & Beverage Industry News*. Foodmag. <https://www.foodmag.com.au/non-alcoholic-beverage-producers-experience-surge-in-unplanned-downtime-as-production-slows/>
- [12] Jiménez, J. J. M., & Vingerhoeds, R. (2019, October). A System Engineering Approach to Predictive Maintenance Systems: from needs and desires to logical architecture. In *2019 International Symposium on Systems Engineering (ISSE)* (pp. 1-8). IEEE.
- [13] Verhagen, W. J., & De Boer, L. W. (2018). Predictive maintenance for aircraft components using proportional hazard models. *Journal of Industrial Information Integration*, 12, 23-30.
- [14] Liu, S., Chen, H., Shang, B., & Papanikolaou, A. (2022). Supporting Predictive Maintenance of a Ship by Analysis of Onboard Measurements. *Journal of Marine Science and Engineering*, 10(2), 215.
- [15] Tsarouhas, P. H. (2013). Evaluation of overall equipment effectiveness in the beverage industry: a case study. *International Journal of Production Research*, 51(2), 515-523.
- [16] Liberopoulos, G., & Tsarouhas, P. (2005). Reliability analysis of an automated pizza production line. *Journal of Food Engineering*, 69(1), 79-96.
- [17] Ashurst, P. R. (2016). *Chemistry and technology of soft drinks and fruit juices*. John Wiley & Sons, Incorporated.
- [18] Hoe wordt frisdrank gemaakt? - FWS. <https://www.fws.nl/onze-dranken/frisdrank/productieproces/>
- [19] Hyfoma. (2016, November 29). Frisdranken en siropen - Safe Food Factory. Safefoodfactory. <https://www.safefoodfactory.com/nl/knowledge/23-frisdranken-en-siropen/>
- [20] Moerman, F., Rizoulières P., & Majoor, F. A. (2014). Hygiene in food processing. In *10 - cleaning in place (cip) in food processing* (pp. 305–383). essay, Elsevier Ltd. <https://doi.org/10.1533/9780857098634.3.305>
- [21] Bürkert Fluid Control Systems UK. (2020, May 13). *Bürkert Spotlight: CIP (Clean in Place) application training* [Video]. YouTube. <https://www.youtube.com/watch?v=X3BhkCly7fl>
- [22] Ehedg. (2001). Hygienic design and safe use of double-seat mixproof valves. *Trends in Food Science & Technology*, 12(5-6), 203–206. [https://doi.org/10.1016/S0924-2244\(00\)00063-7](https://doi.org/10.1016/S0924-2244(00)00063-7)
- [23] Moerman, F., Rizoulières, P., & Majoor, F. A. (2014). Cleaning in place (CIP) in food processing. In *Hygiene in food processing* (pp. 305-383). Woodhead Publishing.
- [24] APV D4 / DA4 Series - Double seat mixproof valves. (n.d.). <https://www.spxflow.com/apv/products/d4-da4-series-double-seat-mixproof-valves/>
- [25] *Vlinderkleppen werking*. (n.d.). <https://eriks.nl/nl/producten/stromingstechniek/afsluiters/vlinderkleppen/vlinderkleppen-werking/>
- [26] *Butterfly Valves | Actuated | Manual | Pneumatic | SV/SVS Series | Manufacturer | SPX FLOW | APV*. (n.d.). <https://www.spxflow.com/apv/products/sv-svs-series-butterfly-valves/>
- [27] SPX FLOW, Inc. (2016, February 16). *SPX FLOW - APV - DELTA SV Butterfly Valve Actuation Demonstration* [Video]. YouTube. <https://www.youtube.com/watch?v=-iaFVFiiAbA>
- [28] Instrumentation Tools. (2021). PLC Valve Control Ladder Logic Programming. *Inst Tools*. <https://instrumentationtools.com/plc-controls-onoff-valve/>
- [29] Piontek, Z. Investigation into Condition-Based Monitoring Techniques for Valves, Pumps, and Storage Tanks Undergraduate Research–Spring 2020.
- [30] Dixon valve (2014). Mixproof valves. <https://europe.dixonvalve.com/sites/default/files/documents/Mix-Proof-Valves.pdf>
- [31] Expert knowledge, Soft drink Factory

- [32] Eaton, A., Ahmed, W. H., and Hassan, M. (November 18, 2021). "Evaluating the Performance Degradation of Centrifugal Pumps Using the Principal Component Analysis." *ASME. J. Pressure Vessel Technol.* April 2022; 144(2): 021405. <https://doi.org/10.1115/1.4052522>
- [33] Centrifugal Pumps Selection Guide: Types, Features, Applications | GlobalSpec.
- [34] Alfa Laval, Centrifugal pumps LKH.
- [35] Beebe, R. S.. (2004). *Predictive maintenance of pumps using condition monitoring*. Elsevier.
- [36] Zennaro, I., Battini, D., Sgarbossa, F., Persona, A. and De Marchi, R. (2018), "Micro downtime: Data collection, analysis and impact on OEE in bottling lines the San Benedetto case study", *International Journal of Quality & Reliability Management*, Vol. 35 No. 4, pp. 965-995. <https://doi.org/10.1108/IJQRM-11-2016-0202>
- [37] Brand, P. (2020, April 2). Downtime in Manufacturing: What's the True Cost? - Oden Technologies. Oden. <https://oden.io/blog/downtime-in-manufacturing-the-true-cost/>
- [38] Mendoza, K. (2022, September 1). iTWire - Unplanned downtime cripples drink production by 12.5%: report. Itwire. <https://itwire.com/it-industry-news/manufacturing-technology/unplanned-downtime-cripples-drink-production-by-12-5-report.html>
- [39] Peleg, L. (2022, January 5). Downtime Cost in Process Manufacturing | Precognize. Precog. <https://www.precog.co/blog/downtime-cost-process-manufacturing/>
- [40] (2022, July 6). Tackling food and beverage production challenges with predictive maintenance. Foodprocessing. <https://www.foodprocessing.com.au/content/processing/article/tackling-food-and-beverage-production-challenges-with-predictive-maintenance-1368320479>
- [41] Rowley, J. (2019, October 17). Predicting the end of unplanned downtime | Control Engineering. Controleng. <https://www.controleng.com/articles/predicting-the-end-of-unplanned-downtime/>
- [42] Sharif MA, Grosvenor RI. Process plant condition monitoring and fault diagnosis. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*. 1998;212(1):13-30. doi:10.1243/0954408981529268
- [43] CEN, E. (2001). EN 13306: maintenance terminology. *European committee for standardization*.
- [44] Shukla K, Nefti-Meziani S, Davis S. A heuristic approach on predictive maintenance techniques: Limitations and scope. *Advances in Mechanical Engineering*. 2022;14(6). doi:10.1177/16878132221101009
- [45] Poór, P., Ženíšek, D., & Basl, J. (2019, July). Historical overview of maintenance management strategies: Development from breakdown maintenance to predictive maintenance in accordance with four industrial revolutions. In *Proceedings of the international conference on industrial engineering and operations management, Pilsen, Czech Republic* (pp. 23-26).
- [46] Merkt, O. (2019, September). On the use of predictive models for improving the quality of industrial maintenance: An analytical literature review of maintenance strategies. In *2019 Federated Conference on Computer Science and Information Systems (FedCSIS)* (pp. 693-704). IEEE.
- [47] (2019, April 25). *Comprehensive Guide To Corrective Maintenance* | Limble CMMS. Limblecmms. <https://limblecmms.com/blog/corrective-maintenance/#:~:text=Immediate%20corrective%20maintenance%20refers%20to,are%20postponed%2Fscheduled%20for%20later.>
- [48] Yasin, M. N., Zin, R. M., Hamid, M. Y., Zakaria, M. A., & Deraman, R. (2019, April). Deferred maintenance of buildings: A review paper. In *IOP Conference Series: Materials Science and Engineering* (Vol. 513, No. 1, p. 012028). IOP Publishing.
- [49] Kothamasu, R., Huang, S. H., & VerDuin, W. H. (2006). System health monitoring and prognostics—a review of current paradigms and practices. *The International Journal of Advanced Manufacturing Technology*, 28(9), 1012-1024.
- [50] Marques, H., & Giacotto, A. (2019, October). Prescriptive maintenance: Building alternative plans for smart operations. In *FT2019. Proceedings of the 10th Aerospace Technology Congress, October 8-9, 2019, Stockholm, Sweden* (No. 162, pp. 231-236). Linköping University Electronic Press.
- [51] Abbasi, T., Lim, K. H., & San Yam, K. (2019, April). Predictive maintenance of oil and gas equipment using recurrent neural network. In *Iop conference series: Materials science and engineering* (Vol. 495, No. 1, p. 012067). IOP Publishing.
- [52] Vaz, L. (2021, November 5). Predictive Maintenance in Oil & Gas Industry: The Complete Guide. Birlasoft. <https://www.birlasoft.com/articles/predictive-maintenance-in-oil-gas-Industry>
- [53] Yiwei, W. A. N. G., Christian, G. O. G. U., Binaud, N., Christian, B. E. S., & Haftka, R. T. (2017). A cost driven predictive maintenance policy for structural airframe maintenance. *Chinese Journal of Aeronautics*, 30(3), 1242-1257.
- [54] Wang, J., Liang, Y., Zheng, Y., Gao, R. X., & Zhang, F. (2020). An integrated fault diagnosis and prognosis approach for predictive maintenance of wind turbine bearing with limited samples. *Renewable energy*, 145, 642-650.
- [55] Hsu, J. Y., Wang, Y. F., Lin, K. C., Chen, M. Y., & Hsu, J. H. Y. (2020). Wind turbine fault diagnosis and predictive maintenance through statistical process control and machine learning. *Ieee Access*, 8, 23427-23439.
- [56] Abid, K., Sayed Mouchaweh, M., & Cornez, L. (2018, September). Fault prognostics for the predictive maintenance of wind turbines: State of the art. In *Joint european conference on machine learning and knowledge discovery in databases* (pp. 113-125). Springer, Cham.
- [57] Kimera, D., & Nangolo, F. N. (2020). Predictive maintenance for ballast pumps on ship repair yards via machine learning. *Transportation Engineering*, 2, 100020.
- [58] Vafaei, N., Ribeiro, R. A., & Camarinha-Matos, L. M. (2019). Fuzzy early warning systems for condition based maintenance. *Computers & Industrial Engineering*, 128, 736-746.

- [59] Li, H., Parikh, D., He, Q., Qian, B., Li, Z., Fang, D., & Hampapur, A. (2014). Improving rail network velocity: A machine learning approach to predictive maintenance. *Transportation Research Part C: Emerging Technologies*, 45, 17-26.
- [60] Jimenez, J. J. M., Schwartz, S., Vingerhoeds, R., Grabot, B., & Salaün, M. (2020). Towards multi-model approaches to predictive maintenance: A systematic literature survey on diagnostics and prognostics. *Journal of Manufacturing Systems*, 56, 539-557.
- [61] Jardine, A. K., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing*, 20(7), 1483-1510.
- [62] Vingerhoeds, R. A., Janssens, P., Netten, B. D., & Fernández-Montesinos, M. A. (1995). Enhancing off-line and on-line condition monitoring and fault diagnosis. *Control Engineering Practice*, 3(11), 1515-1528.
- [63] Tsang, A.H.C. (1995), "Condition-based maintenance: tools and decision making", *Journal of Quality in Maintenance Engineering*, Vol. 1 No. 3, pp. 3 -17. <https://doi.org/10.1108/13552519510096350>
- [64] Edwards, D.J., Holt, G.D. and Harris, F.C. (1998), "Predictive maintenance techniques and their relevance to construction plant ", *Journal of Quality in Maintenance Engineering*, Vol. 4 No. 1, pp. 25-37. <https://doi.org/10.1108/13552519810369057>
- [65] Purarjomandlangrudi, A., Ghapanchi, A. H., & Esmalifalak, M. (2014). A data mining approach for fault diagnosis: An application of anomaly detection algorithm. *Measurement*, 55, 343-352.
- [66] 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.
- [67] Hao, L., Jinsong, Y., Ping, Z., & Kingshan, L. (2009, August). A review on fault prognostics in integrated health management. In *2009 9th International Conference on Electronic Measurement & Instruments* (pp. 4-267). IEEE.
- [68] Majstorović, V. D., & Milačić, V. R. (1990). Expert systems for maintenance in the CIM concept. *Computers in Industry*, 15(1-2), 83-93.
- [69] Freyermuth, B. (1992). Knowledge based incipient fault diagnosis of industrial robots. In *Fault Detection, Supervision and Safety for Technical Processes 1991* (pp. 369-375). Pergamon.
- [70] Boullart, L. (1992). A gentle introduction to artificial intelligence. In *Application of Artificial Intelligence in Process Control*, Pergamon Press, Oxford (pp. 5-40).
- [71] Aamodt, A., & Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI communications*, 7(1), 39-59.
- [72] 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.
- [73] Buckley, J. J., & Eslami, E. (2002). *An introduction to fuzzy logic and fuzzy sets* (Vol. 13). Springer Science & Business Media.
- [74] Menon, S., Jin, X., Chow, T. W., & Pecht, M. (2015). Evaluating covariance in prognostic and system health management applications. *Mechanical Systems and Signal Processing*, 58, 206-217.
- [75] Schmidt, B., & Wang, L. (2015). Predictive maintenance: literature review and future trends. In *The International Conference on Flexible Automation and Intelligent Manufacturing (FAIM), 23-26 June 2015, University of Wolverhampton, UK* (Vol. 1, pp. 232-239). The Choir Press.
- [76] Ghahramani, S. (2018). *Fundamentals of probability: with stochastic processes*. Chapman and Hall/CRC.
- [77] Chen, C., Wang, C., Lu, N., Jiang, B., & Xing, Y. (2021). A data-driven predictive maintenance strategy based on accurate failure prognostics. *Eksploracja i Niezawodność*, 23(2).
- [78] Javed, K., Gouriveau, R., & Zerhouni, N. (2017). State of the art and taxonomy of prognostics approaches, trends of prognostics applications and open issues towards maturity at different technology readiness levels. *Mechanical Systems and Signal Processing*, 94, 214-236.
- [79] Russell, S. J. (2010). *Artificial intelligence a modern approach*. Pearson Education, Inc..
- [80] What is Unsupervised Learning? | IBM. <https://www.ibm.com/topics/unsupervised-learning>
- [81] Bhatt, S. (2018, March 19). Reinforcement Learning 101. Towardsdatascience. <https://towardsdatascience.com/reinforcement-learning-101-e24b50e1d292>
- [82] Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science*, 2(6), 420.
- [83] Vachtsevanos, G. J. (2006). *Intelligent fault diagnosis and prognosis for engineering systems* (Vol. 456). Hoboken: Wiley.
- [84] Nasution, F. P., Sævik, S., & Gjøsteen, J. K. (2012). Fatigue analysis of copper conductor for offshore wind turbines by experimental and FE method. *Energy Procedia*, 24, 271-280.
- [85] Andymath.com. (2023, March 27). *Logistic Function* | <https://andymath.com/logistic-function/>
- [86] Npshr (n.d.). *Cavitatie pompen, oplossen met npsha en npshr*. <https://www.wdapompen.nl/Cavitatie.html>
- [87] *Useful information on pump cavitation*. (n.d.). <https://www.michael-smith-engineers.co.uk/resources/useful-info/pump-cavitation>
- [88] *Pump suction pipe design considerations*. (n.d.). <https://www.michael-smith-engineers.co.uk/resources/useful-info/suction-pipeline-design>
- [89] *What is Net Positive Suction Head (NPSH)?* | SimScale. (2023, May 11). SimScale. <https://www.simscale.com/docs/simwiki/cfd-computational-fluid-dynamics/what-is-net-positive-suction-head-npsh/>
- [90] Alfa Laval. (2022). *LKH Centrifugal Pump* (Version ESE00698-EN31) [Dataset].
- [91] Inventflow. (2020, August 21). *Wat is cavitatie?* - Inventflow. <https://inventflow.nl/nl/faq/cavitatie>

- [92] D, L. (2021, September 10). Symptoms of Bad Pump Seals: Why is My Mechanical Seal Leaking? *Mechanical Seals*. <https://www.gallagherseals.com/blog/symptoms-of-bad-pump-seals-why-is-my-mechanical-seal-leaking>
- [93] White, Frank M.. (2016). *Fluid Mechanics* (8th ed). Secacus, United State: McGraw Hill Education.
- [94] Gülich, Ing, J. F.. (2008). Centrifugal Pumps.
- [95] ANSI/HI 9.6.1-2012 *Guideline for NPSH Margin*.
- [96] Ferman, R. (2017). NPSH margin — How much? *Empowering Pumps and Equipment*. <https://empoweringpumps.com/npsH-net-positive-suction-head/>
- [97] Bloch, Heinz P. Budris, Allan R.. (2014). *Pump User's Handbook - Life Extension (4th Edition)*. River Publishers. Retrieved from: <https://app.knovel.com/hotlink/toc/id:kpPUHLEE12/pump-users-handbook-life/pump-users-handbook-life>
- [98] *Cavitation in centrifugal pumps | Nuclear-power.com*. (2021, October 27). Nuclear Power. <https://www.nuclear-power.com/nuclear-engineering/fluid-dynamics/centrifugal-pumps/cavitation/>
- [99] Cooper, P. A., Tchobanoglous, G., Garbus, R. O., Hart, R. J., Reh, C. W., Sloan, L. G., & Smith, E. C. (2008). Performance of centrifugal pumps. In *Elsevier eBooks* (p. 10.1-10.44). <https://doi.org/10.1016/b978-185617513-5.50017-2>
- [100] Ahonen, T. (2011). *Monitoring of centrifugal pump operation by a frequency converter*. Lappeenranta University of Technology.
- [101] Z. Lai, Q. Li, A. Zhao, W. Zhou, H. Xu and D. Wu, "Improving Reliability of Pumps in Parallel Pump Systems Using Particle Swam Optimization Approach," in *IEEE Access*, vol. 8, pp. 58427-58434, 2020, doi: 10.1109/ACCESS.2020.2980396.
- [102] P. Barringer, "A Life Cycle Cost Summary," in *Proceedings of the International Conference of Maintenance Societies (ICOMS@ -2003)*, Perth, Australia, 20–23 May 2003.
- [103] McKinney, R. (2021, December 16). Conserve Your Energy - The Benefits of Pump System Optimization. *Reliabilityweb*. https://reliabilityweb.com/articles/entry/conserve_your_energy_pump_benefits
- [104] *BEP @ Rated Flow for Centrifugal Pumps - API 610*. (2020, March 3). API 610. <https://pumpsdesign.com/bep-rated-flow/>
- [105] Bachus, L. (2013). Unwritten laws of pump operation. *World Pumps*, 2013(5), 38–39. [https://doi.org/10.1016/s0262-1762\(13\)70186-x](https://doi.org/10.1016/s0262-1762(13)70186-x)
- [106] Allpumps. (2021). How To Find The Best Efficiency Point For Your Pump. *All Pumps*. <https://allpumps.com.au/find-best-efficiency-point-pump/>
- [107] Ebara. (2021). *Centrifugal Pump [Dataset]*.
- [108] Ghahramani, S. (2018). *Fundamentals of probability: with stochastic processes*. Chapman and Hall/CRC.
- [109] Chen, C., Wang, C., Lu, N., Jiang, B., & Xing, Y. (2021). A data-driven predictive maintenance strategy based on accurate failure prognostics. *Eksplatacja i Niezawodność*, 23(2).
- [110] Jardine, A. K., & Tsang, A. H. (2021). *Maintenance, replacement, and reliability: theory and applications*. CRC press.
- [111] Heng, A., Tan, A. C., Mathew, J., Montgomery, N., Banjevic, D., & Jardine, A. K. (2009). Intelligent condition-based prediction of machinery reliability. *Mechanical Systems and Signal Processing*, 23(5), 1600-1614.
- [112] van Noortwijk, J. M. (2009). A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety*, 94(1), 2-21.
- [113] Conditional Reliability for a System. Weibull. <https://www.weibull.com/hotwire/issue193/>
- [114] Maguluri, G., & Zhang, C. H. (1994). Estimation in the mean residual life regression model. *Journal of the Royal Statistical Society: Series B (Methodological)*, 56(3), 477-489.
- [115] Clarotti, C. A., & Spizzichino, F. (1989). The Bayes predictive approach in reliability theory. *IEEE transactions on reliability*, 38(3), 379-382.
- [116] Dey, S., & Stori, J. A. (2005). A Bayesian network approach to root cause diagnosis of process variations. *International Journal of Machine Tools and Manufacture*, 45(1), 75-91.
- [117] Langseth, H., & Portinale, L. (2007). Bayesian networks in reliability. *Reliability Engineering & System Safety*, 92(1), 92-108.
- [118] Kelly, D. L., & Smith, C. L. (2009). Bayesian inference in probabilistic risk assessment—the current state of the art. *Reliability Engineering & System Safety*, 94(2), 628-643.
- [119] Ghahramani, Z. (2006). Learning dynamic Bayesian networks. *Adaptive Processing of Sequences and Data Structures: International Summer School on Neural Networks "ER Caianiello" Vietri sul Mare, Salerno, Italy September 6–13, 1997 Tutorial Lectures*, 168-197.
- [120] O'Connor, P., & Kleyner, A. (2012). *Practical reliability engineering*. John Wiley & Sons.
- [121] Ito, K., & Xiong, K. (2000). Gaussian filters for nonlinear filtering problems. *IEEE transactions on automatic control*, 45(5), 910-927.
- [122] Haug, A. J. (2005). A tutorial on Bayesian estimation and tracking techniques applicable to nonlinear and non-Gaussian processes.
- [123] Tsay, R. S. (2000). Time series and forecasting: Brief history and future research. *Journal of the American Statistical Association*, 95(450), 638-643.
- [124] Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2), 187-202.

- [125] Eleye-Datubo, A. G., Wall, A., Saajedi, A., & Wang, J. (2006). Enabling a powerful marine and offshore decision-support solution through Bayesian network technique. *Risk Analysis*, 26(3), 695-721.
- [126] Xing, E., Su, C., Hernandez, A., Chueng, M. & Li, X. (2019, January 16). Lecture 2: Bayesian Networks. Sailinglab. <https://sailinglab.github.io/pgm-spring-2019/notes/lecture-02/>
- [127] Sadigh, D. *CS 221: Artificial Intelligence: Principles and techniques*. Lecture 13 (2018) <https://web.stanford.edu/class/archive/cs/cs221/cs221.1186>
- [128] Jensen, F. V. (2001). *Bayesian Networks and Decision Graphs*, Springer-Verlag.
- [129] Jensen, F. V., & Nielsen, T. D. (2007). *Bayesian networks and decision graphs* (Vol. 2). New York: Springer.
- [130] Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: principles and techniques*. MIT press.
- [131] Pang, Ir. Y. (2021, March 3). *Operations and Maintenance, Lecture 3.4: BN & KBS*. 3mE, TU Delft.
- [132] Sorge, V. (n.d.). *Intro to AI: Lecture 8*. School of Computer Science, University of Birmingham. <https://www.cs.bham.ac.uk/~vxs/teaching/ai/slides/lecture8-slides.pdf>
- [133] Zhang, N.L. (2008, fall). COMP538: Introduction to Bayesian Networks Lecture 4: Inference in Bayesian Networks: The VE Algorithm. Department of Computer Science and Engineering, Hong Kong University of Science and Technology. [COMP538: Introduction to Bayesian Networks - Lecture 4: Inference in Bayesian Networks: The VE Algorithm \(hkust.edu.hk\)](https://www.comp.hkust.edu.hk/~nliu/COMP538/Introduction%20to%20Bayesian%20Networks%20-%20Lecture%204%20-%20Inference%20in%20Bayesian%20Networks%20-%20The%20VE%20Algorithm)
- [134] *Bayesian Networks*. (2017, February). <http://www.cs.ru.nl/~peterl/BN/>
- [135] Verbert, K., De Schutter, B., & Babuška, R. (2017). Timely condition-based maintenance planning for multi-component systems. *Reliability Engineering & System Safety*, 159, 310-321.
- [136] Sadigh, D. *CS 221: Artificial Intelligence: Principles and techniques*. Lecture 15 (2018) <https://web.stanford.edu/class/archive/cs/cs221/cs221.1186>
web.stanford.edu/class/archive/cs/cs221/cs221.1186/lectures/bayes3.pdf
- [137] Ankan, A., & Panda, A. (2015). pgmpy: Probabilistic graphical models using python. In *Proceedings of the 14th python in science conference (scipy 2015)* (Vol. 10).
- [138] Cai, B., Kong, X., Liu, Y., Lin, J., Yuan, X., Xu, H., & Ji, R. (2018). Application of Bayesian networks in reliability evaluation. *IEEE Transactions on Industrial Informatics*, 15(4), 2146-2157.
- [139] Schietekat, S., De Waal, A., & Gopaul, K. G. (2016). Validation & verification of a Bayesian network model for aircraft vulnerability. INCOSE.
- [140] Vojković, L., Kuzmanić Skelin, A., Mohovic, D., & Zec, D. (2021). The development of a Bayesian network framework with model validation for maritime accident risk factor assessment. *Applied Sciences*, 11(22), 10866.
- [141] Larrosa, J. & Rollon, E. (2015). *Bayesian Networks- Exact Inference by Variable Elimination*. Universitat Politècnica de Catalunya.
- [142] Institute of Electrical and Electronics Engineers. (1990). *IEEE Standard Computer Dictionary: Compilation of IEEE Standard Computer Glossaries: 610*. IEEE.
- [143] Conti, M., & Orcioni, S. (2020). Modeling of failure probability for reliability and component reuse of electric and electronic equipment. *Energies*, 13(11), 2843.
- [144] Menčík, J. (2016). Reliability of systems. In *InTech eBooks*. <https://doi.org/10.5772/62358>
- [145] Barringer, P. E. (1996). An overview of reliability engineering principles.
- [146] Barringer, P. E. (1997). Availability, Reliability, Maintainability, and Capability.
- [147] Barringer, P. E. (2004, May 25) Process and Equipment Reliability. Maintenance & Reliability Technology Summit, Chicago, IL.
- [148] What is a Key Performance Indicator (KPI)? - KPI.org. KPI. <https://www.kpi.org/kpi-basics/>
- [149] Prasad, N. D., & Radhakrishna, C. (2019, September). New Key Performance Indicators (KPI) For Substation Maintenance Performance. In *2019 International Conference on Computing, Power and Communication Technologies (GUCON)* (pp. 31-34). IEEE.
- [150] Peng, Y., Dong, M., Zuo, M.J., (2010, January 6). Current Status of Machine Prognostics in Condition-Based Maintenance: A review. *International Journal of Advanced Manufacturing Technology*, 50(4), pp.297-313.
- [151] Anh, D. T., Dąbrowski, K., & Skrzypek, K. (2018). The predictive maintenance concept in the maintenance department of the "Industry 4.0" production enterprise. *Foundations of Management*, 10(1), 283-292.
- [152] Pagano, D. (2023). A predictive maintenance model using Long Short-Term Memory Neural Networks and Bayesian inference. *Decision Analytics Journal*, 6, 100174. <https://doi.org/10.1016/j.dajour.2023.100174>
- [153] Manuel Inácio da Silva, C., Manuel Pereira Cabrita, C., & Carlos de Oliveira Matias, J. (2008). Proactive reliability maintenance: a case study concerning maintenance service costs. *Journal of Quality in Maintenance Engineering*, 14(4), 343-355.
- [154] CEN–European Committee for Standardization. (2007). EN 15341–European Standard–Maintenance Key Performance Indicators.
- [155] Torell, W., & Avelar, V. (2004). Mean time between failure: Explanation and standards. *white paper*, 78, 6-7.

Appendix A: Scientific paper

Reliability improvement of a soft drink production line using a Bayesian network

P.M.H. Freling
Delft University of Technology, The Netherlands

This paper contributes to the research on how a Bayesian network (BN) can contribute to improving reliability with specific application to soft drink production lines. The soft drink production industry is mostly using a corrective maintenance strategy. However, this maintenance strategy often causes the industry to face unplanned downtimes of the production lines. Unplanned downtimes are caused by equipment failure in the line. As a result, unplanned downtimes reduce the reliability of the production line. This paper discusses the research underlying how applying a BN can help improve the reliability of production lines of the soft drink industry.

Bayesian Network (BN) – Reliability – Soft drink production lines – Maintenance – Downtime (DT) – Mean Time Between Failure (MTBF) – Predictive maintenance (PdM) – Key Performance Indicator (KPI) – Maintenance Downtime Index (MDI)

Introduction

One of the solutions to reducing unplanned downtimes can be found in a different maintenance strategy. Within the maintenance strategies, according to EN13306 [1], a distinction can be made between reactive and proactive maintenance. Reactive maintenance is when maintenance is performed only at the moment when, in this case, components in the line have failed. Proactive maintenance, on the other hand, is a strategy to perform timely maintenance just before a component fails. Within both categories, different forms of maintenance strategies can be found, see Figure 1 [2].

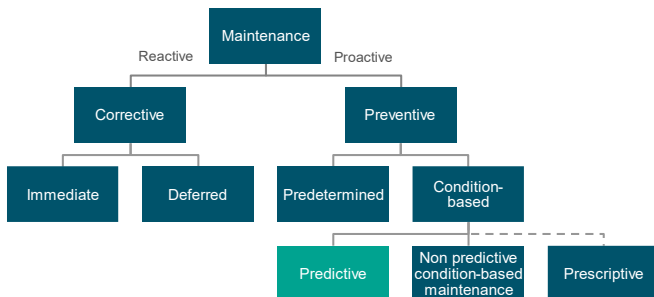


Figure 1: Different maintenance strategies [2]

Currently, the corrective maintenance strategy, which falls under reactive maintenance strategies, is mainly used within the soft drink industry, see Figure 1.

This means that maintenance is carried out only when a component has failed. As a result, when a component fails, it causes a downtime. Two types of downtimes can be distinguished [3]; planned and unplanned. Planned downtimes include downtimes that are scheduled in advance for cleaning, line rebuilds or maintenance. Unplanned downtime include downtimes caused by the failure of components in the production line. The moment unplanned downtime occurs, the line falls silent. As a result, nothing can be produced and sometimes everything in the line has to be discarded.

Next, it can then be determined which component is causing the unplanned downtime. After replacement or repair, the line has to be cleaned before new production can be started. Line reliability is strongly influenced by the number and duration of unplanned downtimes. The definition of reliability in this research is defined as: "The ability of a system or component to perform its required functions under stated conditions for a specified period of time." [4]. If there are many unplanned downtimes of the production line then this is an indication that reliability is poor [5,6].

In recent years, the beverage industry has been slowly looking at how to improve line reliability. One development in this is that there is growing interest in ways to monitor the condition of equipment in the line and how the resulting data can be used [6]. Some companies are already using sensors to collect data on the condition of equipment in the line. However, they often stop at the point of data collection and, in some cases, a simple analysis after a line has stopped. This research will address, through a literature and case study, how condition monitoring data can be used in a BN [7]. The aim is to provide insight into how reliability can be improved when a predictive maintenance strategy is applied. It is expected that with the PdM strategy, the number of unplanned downtimes can be reduced and thus the reliability of the production line can be improved. This research focuses on valves and pumps within a soft drink production line. The main research question is as follows:

How can a predictive maintenance strategy contribute to improving the reliability of a soft drink production line?

For this research, a case study was also done for an existing production line of a Dutch soft drink manufacturer. The line analysed for this research starts at the point the syrup is mixed. After this, the steps of heating, cooling, resting and carbonisation still have to take place, as can be seen in Figure 2.

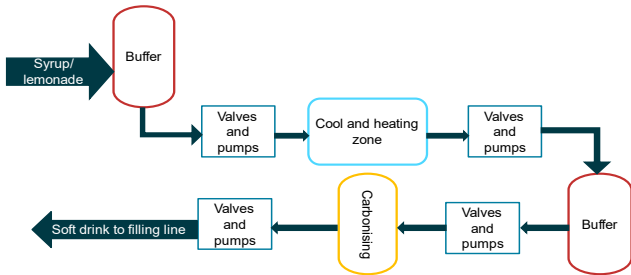


Figure 2: Components production line

To transport the soft drink through the line, valves and pumps are used. The valves and pumps are the components in the production line that are the focus of this research. It will be examined how condition monitoring techniques and data for the valves and pumps can be used to create a PdM strategy to improve the reliability of the production line [8].

The model

The model for PdM in this research is composed of several parts. There is a general part, a BN-based part and an MTBF part. To clarify when to use which part, a flow chart has been created, the flow chart can be found in Appendix A1.

The BN is made up of several layers. In the BN, all components from the production line are represented. The complete BN can be seen in Figure 3.

The production line (the PL node) is divided into five line segments, in the BN of Figure 3 the LS nodes. In each line segment there are a number of valves and one or two pumps. The valves are denoted by the $V_{x,x}$ nodes and the pumps by the $P_{x,x}$ nodes. Here, the first x represents the line segment to which the component belongs and the second x indicates the component number within that specific line segment. (As an example, $V_{1,2}$ is the second valve from line segment 1).

Valves

The condition of the valves is determined by the run time, in Dutch: looptijd, ($LT_{V_{x,x}}$ nodes), which is the time it takes a valve to open or close. It is known that once the looptijd increases this is an indication that the valve is about to fail. In this research, the relationship between looptijd and the probability of failure is represented by a logistic function [9].

Pumps

For pumps, there are two parameters that determine the probability of failure of a pump. The first is the parameter ($B_{P_{x,x}}$) is at what percentage of the BEP the pump operates at [10]. It is known that the further from the BEP the pump operates the earlier the pump fails. The other parameter ($G_{P_{x,x}}$) indicates whether cavitation occurs or not based on the NPSH margin [11] [12]. Both parameters are represented in the BN as parent nodes of the $P_{x,x}$ nodes.

Inference

By entering evidence values for the LT , B and G nodes, forward inference can be performed to determine the probability of failure and hence DT for the line segments and the entire production line.

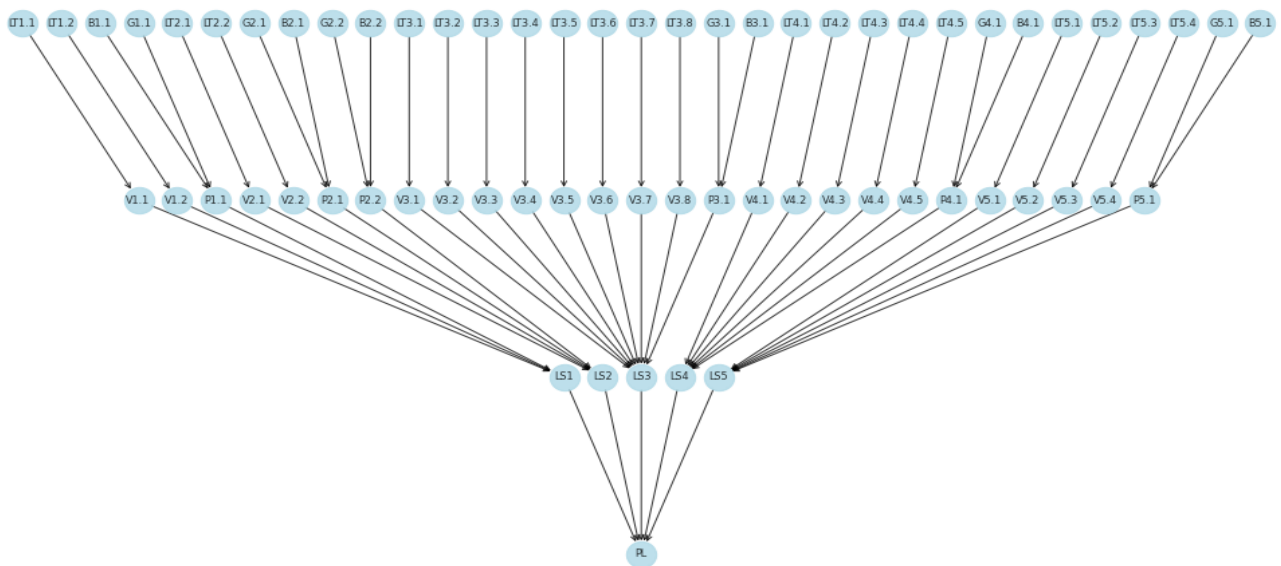


Figure 3: Bayesian network of the production line, in the top layer are the nodes of the parameters for the valves and pumps. The second layer contains the nodes of the valves and pumps. The third layer shows the line segments and the last layer the node of the production line.

The flowchart, see [Appendix A1](#), distinguishes between the BN section (sections framed in light blue), the MTBF section (sections framed in purple) and the general section (sections framed in pink).

Starting at the beginning (pink section at the top), it is first necessary to determine how long and which processes are sent across the line based on the schedule. Then, based on the historical process data, a synthetic dataset corresponding to the data for the future process can be generated.

The production planning dataset showing how long and which process goes over the line is used to determine the MTBF (the purple part). The synthetic process dataset is used in the BN (the blue part).

MTBF part (purple)

The time until planned downtime should be entered in the code for MTBF. The MTBF is determined from the historical data is unique for each component. The code uses a time counter to determine how long each component has been in service. Once the time of future production is added, the code checks for components that are soon or within 2% of the end of their useful life. If not, then from this section there is no indication that any components will fail until the next scheduled downtime. If there are components within the 2% margin, then it is necessary to determine which components they are and the remaining life of those components.

As output data from this section, a list of the critical components and their expected remaining life in hours is given.

BN section (blue)

The dataset created from the historical data contains values for LT, G and B nodes for every 15 minutes. The data is used as evidence to then make inferences with the BN to determine the output values for the PL and LS nodes.

As indicated earlier, values have been identified for the PL and LS nodes that indicate whether a potential failure is approaching. Therefore, after performing the inferences, it is necessary to check whether the values of the PL node are within the safe range. For the PL and LS nodes, there are indicator values that indicate that a component failure may occur in the near future. See [Table 1](#) for the values of the PL and LS nodes at warnings and possible failures.

If all PL and LS values are within the safe range, there is no indication from the BN that a potential failure will occur during the process. It is then assumed that production can be carried out without any problems.

If there are alarming LS values, it must be determined whether it is a warning value or a value that indicates that more is already going on. In the case that it is a warning value, then it must be taken into account that a potential failure may occur in the next two hours. The component(s) responsible for this must then be determined. Then the time until the next scheduled downtime has to be considered. As output data, it is given which components have a critical value and how much time is left until the next stop.

If the value of the LS nodes already exceeds the warning value(s) then it is only necessary to determine which component(s) are causing this. This is collected and used as output data.

General section (pink part at the bottom)

A list is made of the critical components based on the output data, along with how much time is left from MTBF and from the BN. It then needs to be determined whether there is enough time to complete production without unplanned downtime or not. In case there is enough time left then production can be completed before maintenance is carried out. It can then be calculated how much time is needed for maintenance and then what the cost will be.

If there is not enough time to fully run the complete production then maintenance/inspections must be carried out before any production can be started at all. Then it has to be redetermined what the schedule is then and whether more components need maintenance. Afterwards, the time for maintenance and the costs must be calculated.

Once maintenance has been performed on a component then the usage time must be reset in the MTBF code.

With this model, it is possible, based on the synthetic datasets, for both planning and process data, to see prior to a production if there are components in the line that cannot survive it. In this way, it is possible to predict whether maintenance is needed or not. In this research, a synthetic production schedule for a year was used.

Table 1: Warning and Failure values BN

	PL(1)	LS1(1)	LS2(1)	LS3(1)	LS4(1)	LS5(1)
Warning	0.6025	0.5073	0.5270	0.5799	0.5111	0.5421
Potential failure	0.6031	0.5167	0.5327	0.5828	0.5115	0.5485

To determine whether the model contributes to improving reliability, two situations are used, both of which are tested against some KPIs.

In the first situation, the model is not used and the current combination of corrective and a small piece of preventive maintenance is used. Here, the synthetic datasets are analysed and the KPIs are calculated.

The second situation does use the model. This then works with PdM. For this situation, the flow chart is used and the KPIs are calculated.

The first KPI focuses on the ratio of the number of hours of DT used for planned maintenance to the total number of hours of DT [13], see Eq. 1.

$$MDI = \frac{DT \text{ hours for scheduled maintenance}}{\frac{total \text{ DT hours}}{DT \text{ hours for scheduled maintenance}}} = \frac{DT \text{ hours for scheduled maintenance}}{Scheduled + unscheduled \text{ DT hours}} \quad (Eq. 1)$$

To determine whether the reliability of the line is improving, the MDI must increase. Indeed, this means that more DT hours are used for planned maintenance. In addition, the total number of DT hours is composed of both planned DT hours and unplanned DT hours [13]. A higher MDI is caused either by more hours of planned maintenance during DT hours (higher numerator) or by the total number of DT hours decreasing (lower denominator). This is a first indication that line reliability is improving.

The second, third and fourth KPI are related to costs. These are the costs focused on corrective, preventive and predictive maintenance relative to total maintenance costs [14]. It is assumed within this research that predictive maintenance costs are equal to condition-based maintenance costs. This can be determined using Eq. 2, Eq. 3 and Eq. 4.

$$E15 = \frac{Corrective \text{ maintenance cost}}{Total \text{ maintenance cost}} * 100 \quad (Eq. 2)$$

$$E16 = \frac{Preventive \text{ maintenance cost}}{Total \text{ maintenance cost}} * 100 \quad (Eq. 3)$$

$$E17 = \frac{Condition \text{ based maintenance cost}}{Total \text{ maintenance cost}} * 100 \quad (Eq. 4)$$

It is assumed that line reliability increases when total maintenance costs decrease. In addition, an indicator of improved reliability is when a lower percentage goes to corrective maintenance costs and a higher percentage goes to predictive maintenance costs. This indicates that fewer corrective maintenance actions are needed, which in turn is an indication that the reliability of the line is increasing.

By comparing the outcomes of the KPIs for both situations, the impact of the model, with the PdM strategy, on the reliability of the production line can be determined.

Results

This research looked at a period of one year. The synthetic datasets are based on this. The results are therefore also based on one year of production.

In the situation where the model is not used, the findings are that 57 hours of maintenance were required. 22 hours of this was used for maintenance during scheduled DT hours. 35 hours is during unscheduled DT hours. Looking at the costs, the total maintenance cost is €49,120.-.

In the situation where the model is used, the findings are that 57 hours and 50 minutes of maintenance was performed. 51 hours of this is maintenance performed during the scheduled DT hours. 6:50 hours of maintenance took place during unscheduled DT hours. In this research, unscheduled DT hours are determined by the hours needed for corrective maintenance and the hours left to complete repairs when the scheduled DT hours are just too short. Of the 6:50 hours of unscheduled DT, 2 hours are for corrective maintenance, where the line stopped and had to be flushed empty, and the remaining hours are the hours needed on top of the scheduled DT hours to complete the maintenance. Suppose there was 50 minutes of scheduled DT, but maintenance was required and the maintenance took 1 hour, that's 10 minutes of unscheduled DT. The total maintenance cost comes out to €23,120.-.

Table 2: Hours downtime

	Without model	Time used for maintenance	With model	Time used for maintenance
Scheduled for change and PdM	318:25:00	8:00:00	318:25:00	37:00:00
Scheduled for preventive maintenance	200:00:00	14:00:00	200:00:00	14:00:00
Unscheduled	35:00:00	35:00:00	6:50:00	6:50:00
Total downtime	553:25:00		525:15:00	
Maintenance during planned DT		22:00:00		51:00:00

KPIs

The overall results serve as the basis for calculating the KPIs.

MDI

To determine the MDI [13], Eq. 1 has to be filled in. This involves determining what the total number of DT hours was in both situations. Then it is necessary to determine how many hours of the DTs were used for planned maintenance, see Table 2 for the results. Filling in the numbers from Table 2 in Eq. 1 then gives the following results:

$$MDI_{without\ model} = \frac{22:00\ hours}{553:25\ hours} * 100 = 3.98\%$$

This can also be done when the model is used. Eq. 1 will then be like:

$$MDI_{with\ model} = \frac{51:00\ hours}{525:15\ hours} * 100 = 9.71\%$$

From this, it can be seen that in the situation where the model is used, the MDI is higher compared to the situation without the model. This is caused by both a lower total number of DT hours and by more hours of planned maintenance times DT hours. Here, the first indication that the model contributes to improving reliability is in its favour.

Cost KPIs

For cost KPIs, the percentage of corrective, preventive and predictive maintenance costs relative to total maintenance costs was considered [14]. It should be noted that the total maintenance costs of the two situations already differ enormously. In the situation where the model is used, total maintenance costs are 53% lower than when the model is not used, see Figure 4. The reason for this is that there were 28:10 more unplanned DT hours in the situation without the model. As a result, the line had to be emptied before maintenance could be carried out. This ticks up in the costs involved. In the situation where the model is used, there are only 2 times when the line has to be emptied for corrective maintenance.

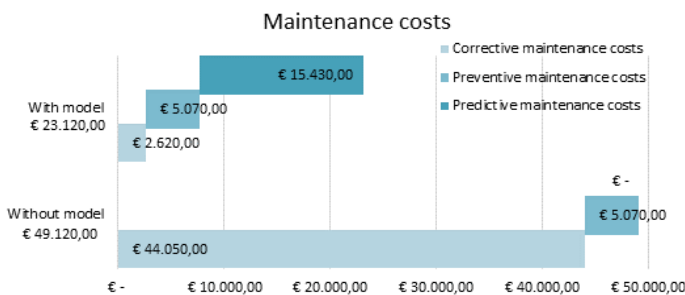


Figure 4: Maintenance costs

If Eq. 2, Eq. 3 and Eq. 4 are then entered, this gives the following results, see Table 3.

Table 3: Cost KPIs results

KPI	Without model	With model
E15	89.68%	11.33%
E16	10.32%	21.93%
E17	0.00%	66.74%

It can be inferred that when the model is used, the largest part of maintenance costs is spent on predictive maintenance and the smallest part on corrective maintenance. Because of the difference in total maintenance costs, the percentage of E16, Eq. 3, is higher in the situation with the model. When looking at the numbers, see Figure 4, it is the same in both situations.

In the situation where the model is used, besides the reduction in total maintenance costs, it can be seen from the second, third and fourth KPI that costs shift from corrective to predictive maintenance. With this, it can be determined that fewer corrective maintenance actions are needed, thereby improving the reliability of the line. This is the second indication that reliability improves when the model is used.

Discussion

As can be seen in the results, using the model ensures that there is an increase in the number of hours that planned maintenance is performed during pre-scheduled DT hours. In addition, there is a reduction in the total number of DT hours. This indicates an increase in line reliability.

However, the model works on many assumptions. An average time for maintenance is assumed for valves and pumps. In the situation that more time may be needed than planned, this will result in a lower MDI. This is because more unscheduled DT hours will then be needed to carry out the maintenance and the number of DT hours used for planned maintenance will remain the same. Keeping the same number in the numerator but a higher number for the denominator results in a lower MDI. To calculate the MDI, the total number of hours of planned DT and number of hours of unscheduled DT are used. Because the line has a large number of hours of scheduled downtime due to the lack of production, the value of the MDI calculation ends up to be low in both situations. Approaching the calculation for MDI slightly differently and looking at the ratio between the number of hours of DT used for scheduled maintenance compared to the total number of hours of DT for maintenance gives the following results.

For the situation where the model is not used, a total of 57 maintenance hours are required, see Table 2. 22 hours of this is used for maintenance during scheduled DT hours. Entering this in Eq. 1 then gives:

$$MDI_{without\ model} = \frac{22:00\ hours}{57:00\ hours} * 100 = 38.6\%$$

In the situation when the model is used, there is a total of 57 hours and 50 minutes of DT for maintenance, see [Table 2](#). 51 hours of this is maintenance during scheduled DT hours. Entering this in [Eq. 1](#) then gives:

$$MDI_{with\ model} = \frac{51:00\ hours}{57:50\ hours} * 100 = 88.2\%$$

Here it can be clearly seen that despite requiring a little more time for maintenance, the MDI still works out in favour of the situation with the model.

When looking at costs, it had been noted earlier that total maintenance costs are 53% lower when the model is used. To calculate the total maintenance costs, several assumptions were made for both situations, including that all parts are always in stock and that in the time span of the synthetic dataset only repairs are carried out and no complete valves or pumps are replaced. If this is not the case, the total maintenance costs will be many times higher. This is due to the fact that there is then a need to take into account that there will be schedule outages, requiring employees to work overtime to make up for lost time. In addition, the cost of replacing a component is many times higher than carrying out a repair.

This model is based on historical data for the valves and assumptions for the dates of the pumps. For the valves in the BN, only looptijd was considered and no other parameters. This was done because no other data was available. For the pumps, no data was available so it was assumed that in the BN, the condition of the pumps can be determined with the probability of cavitation and at what percentage of BEP the pump is operating. However, there are many other possible ways in which the condition of the valves and pumps can be determined. It is therefore impossible to say whether other parameters need to be included to achieve a better model. In addition, the model works with the MTBF of each component. For the valves, this is determined based on historical failure data. For the pumps, based on the manufacturers' data, there is a need to perform maintenance on the pump at least every year for seals and bearings. Although, this may give a slightly distorted view of how long a pump can operate without maintenance. The moment sufficient data is available for the pumps then it can be determined whether the MTBF assumption is correct. If less frequent maintenance is required on the pumps then this will not only give a reduction in the total number of hours of maintenance but also a reduction in costs.

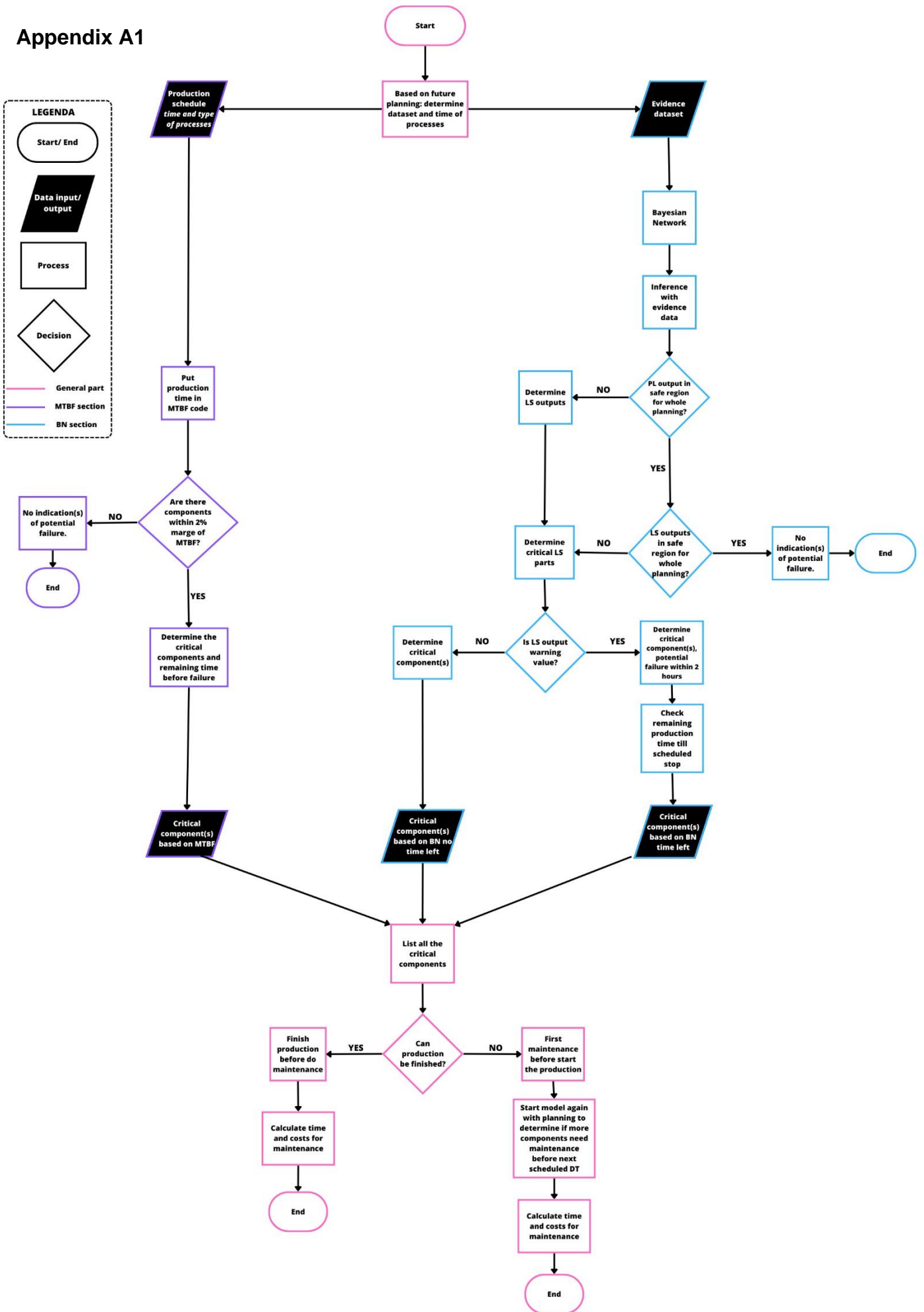
Finally, the model relies on assumptions to indicate what a predictive maintenance strategy can deliver in terms of reduction in unplanned DTs and reduction in maintenance costs. It is assumed that these are the indications that the reliability of the line improves. However, the results also show that the model does not detect all failures in time and

unplanned DTs still occur. Albeit to a lesser extent than when the model is not used, but completely eliminating unplanned DTs is not possible with this model.

References

- [1] CEN, E. (2001). EN 13306: maintenance terminology. *European committee for standardization*.
- [2] Shukla K, Nefti-Meziani S, Davis S. A heuristic approach on predictive maintenance techniques: Limitations and scope. *Advances in Mechanical Engineering*. 2022;14(6). doi:10.1177/16878132221101009
- [3] Zennaro, I., Battini, D., Sgarbossa, F., Persona, A. and De Marchi, R. (2018), "Micro downtime: Data collection, analysis and impact on OEE in bottling lines the San Benedetto case study", *International Journal of Quality & Reliability Management*, Vol. 35 No. 4, pp. 965-995. <https://doi.org/10.1108/IJQRM-11-2016-0202>
- [4] Institute of Electrical and Electronics Engineers. (1990). *IEEE Standard Computer Dictionary: Compilation of IEEE Standard Computer Glossaries: 610*. IEEE.
- [5] Lumix. (2020, May 14). *Importance of Predictive Maintenance in Food Processing Industry | lumixfps*. Lumixfps. <https://www.lumixfps.com.au/blog/importance-of-predictive-maintenance-in-food-processing-industry/>
- [6] McCleery, A. (2022, July 31). *Non-alcoholic beverage producers experience surge in unplanned downtime as production slows - Food & Beverage Industry News*. Foodmag. <https://www.foodmag.com.au/non-alcoholic-beverage-producers-experience-surge-in-unplanned-downtime-as-production-slows/>
- [7] 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.
- [8] Piontek, Z. Investigation into Condition-Based Monitoring Techniques for Valves, Pumps, and Storage Tanks Undergraduate Research–Spring 2020.
- [9] Andymath.com. (2023, March 27). *Logistic Function |*. <https://andymath.com/logistic-function/>
- [10] Cooper, P. A., Tchobanoglous, G., Garbus, R. O., Hart, R. J., Reh, C. W., Sloan, L. G., & Smith, E. C. (2008). Performance of centrifugal pumps. In *Elsevier eBooks* (p. 10.1-10.44). <https://doi.org/10.1016/b978-185617513-5.50017-2>
- [11] ANSI/HI 9.6.1-2012 *Guideline for NPSH Margin*.
- [12] Ferman, R. (2017). NPSH margin — How much? *Empowering Pumps and Equipment*. <https://empoweringpumps.com/npsh-net-positive-suction-head/>
- [13] Prasad, N. D., & Radhakrishna, C. (2019, September). New Key Performance Indicators (KPI) For Substation Maintenance Performance. In *2019 International Conference on Computing, Power and Communication Technologies (GUCON)* (pp. 31-34). IEEE.
- [14] CEN–European Committee for Standardization. (2007). EN 15341–European Standard–Maintenance Key Performance Indicators.

Appendix A1



Appendix B: CIP

There are many different ways in which CIP can be carried out. Looking at the different studies, it can be established that a general CIP process includes at least the following steps [9] [12].

The first step within CIP is to clean the system to flush away product residues. This is also called the product flush. Water is pumped into the system for cleaning.

The second step is the pre-rinse. Clean drinking water at 25 degrees Celsius is used for this. Another option for this step is to use water with alkaline, at 45 degrees Celsius. This step takes between 3 and 10 minutes and is meant to remove 95% of product residues.

The third step is cleaning with a chemical liquid. For this, water is heated and a caustic is added to create an alkaline wash containing between 1-3% caustic. This is then heated between 55-90° C, depending on the chemicals used. The system is then cleaned with this liquid for between 10 and 30 minutes. Once the liquid comes out of the system, it is cleaned to be used again later.

The fourth step is to clean the system again with room-temperature water to flush out the chemicals and other dirt from the system.

The fifth step is cleaning with an acidic solution. This can remove the residual alkaline liquid that has not been flushed away with the water from step four. On average, the acid solution contains between 0.5 and 2% acid. The temperature of the acid solution is between 50 and 70° C and it is kept in the system for between 3 and 20 minutes.

The sixth step is again cleaning the system with water. Only when no more residues of the aforementioned fluids are detected in the water at the outlet then this step is complete.

The seventh step is to disinfect the system. This can be done either with water and disinfectants at room temperature, or water and disinfectants between 70-95° C. The choice depends on the type of disinfection and which micro-organisms need to be removed with this step. This step takes between 10 and 60 minutes.

The eighth and second-to-last step is cleaning with fresh water to remove the residues of the disinfectant liquid. This takes between 5 and 10 minutes.

The final step is drying the system so that it is ready for the next production process.

All steps are summarised in [Figure 51](#).

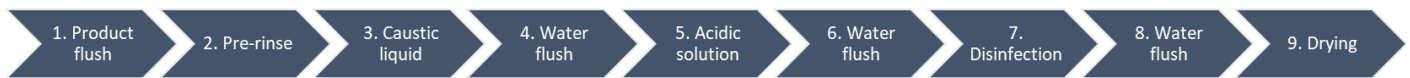


Figure 51: Steps of CIP

Appendix C: Vapour pressure and density

Vapour pressure

To determine whether cavitation is occurring, the NPSHa must be considered. This depends on the inlet velocity, density and vapour pressure of the fluid to be pumped. Vapour pressure and density depend on temperature.

Raoult's law:

$$P_A = X_A P_A^0$$

Looking at an X amount of sugar (sucrose) dissolved in water, the vapour pressure of the mixture can be determined. This is needed to determine the vapour pressure from inside the pump. Note that the mole mass depends on the temperature.

An example is given below.

The vapour pressure is to be determined from a mixture of sugar ($C_{12}H_{22}O_{11}$) and water (H_2O), where there are 158.0 g of sugar (sucrose) dissolved in 641.6 g of water at 25°C. The molar mass of sucrose is 342.3 g/mol. The molar mass of water is 18.01528 g/mol. The vapour pressure of water at 25°C is 23.76 mmHg.

Then Raoult's law must be used:

$$P_{Solution} = X_{H_2O} P_{H_2O}^0$$

First, determine the number of mol of both sucrose and water. It is given that 158.0 g of sucrose is dissolved.

$$\text{moles } C_{12}H_{22}O_{11} = 158 * \left(\frac{1}{342,3}\right) = 0,462 \text{ mol}$$

$$\text{moles } H_2O = 641,6 * \left(\frac{1}{18}\right) = 35,6 \text{ mol}$$

Next, the mole fraction of H₂O needs to be determined:

$$X_{H_2O} = \left(\frac{\text{mol } H_2O}{\text{mol } H_2O + \text{mol } C_{12}H_{22}O_{11}}\right) = \left(\frac{35,6}{35,6 + 0,462}\right) = 0,987$$

Then, finally, the vapour pressure can be determined for the solution of the sugar in the water.

$$P_{Solution} = (0,987)(23,76) = 23,5 \text{ mmHg}$$

Adding the sugar reduces the vapour pressure. The more sugar is dissolved in the water the lower the vapour pressure will be.

Here are the tables shown that can be used to determine the correct vapour pressure for each liquid and temperature.

Density

The density of the liquid can be determined by dividing the total mass by the total volume of the liquid. Note here that the density of water changes as the temperature changes.

$$\rho = \frac{m}{V}$$

m: mass in kg

V: volume in L or in m³

Here are the tables shown that can be used to determine the correct vapour pressure and density for some liquid compositions and temperatures.

1000L		30 degrees			
sugar [%]	Vapour pressure [mm Hg]	Vapour pressure [Pa]	density [kg/L]	density [kg/m3]	
0,25	30,94065239	4125,081044	1,1435	1143,5	
0,20	31,15110544	4153,139143	1,1144	1114,4	
0,17	31,27875711	4170,157966	1,095	1095	
0,14	31,36444103	4181,581549	1,081143	1081,143	
0,13	31,4259318	4189,779645	1,07075	1070,75	
0,11	31,47220834	4195,949343	1,062667	1062,667	
0,10	31,50829554	4200,760574	1,0562	1056,2	
0,09	31,53722494	4204,61751	1,050909	1050,909	

1000L		40 degrees			
sugar [%]	Vapour pressure [mm Hg]	Vapour pressure [Pa]	density [kg/L]	density [kg/m3]	
0,09	54,8572	7313,692278	1,045455	1045,455	
0,10	54,80658	7306,943155	1,0508	1050,8	
0,11	54,74343	7298,524246	1,057333	1057,333	
0,13	54,66246	7287,7284	1,0655	1065,5	
0,14	54,55486	7273,383521	1,076	1076	
0,17	54,40494	7253,395303	1,09	1090	
0,20	54,18159	7223,618119	1,1096	1109,6	
0,25	53,81339	7174,529091	1,139	1139	

1000L		50 degrees			
sugar [%]	Vapour pressure [mm Hg]	Vapour pressure [Pa]	density [kg/L]	density [kg/m3]	
0,09	91,7851	12237,01	1,041818	1041,818	
0,10	91,70006	12225,67	1,0472	1047,2	
0,11	91,59398	12211,53	1,053778	1053,778	
0,13	91,45796	12193,39	1,062	1062	
0,14	91,27721	12169,29	1,072571	1072,571	
0,17	91,02537	12135,72	1,086667	1086,667	
0,20	90,65021	12085,7	1,1064	1106,4	
0,25	90,03176	12003,25	1,136	1136	

1000L		60 degrees			
sugar [%]	Vapour pressure [mm Hg]	Vapour pressure [Pa]	density [kg/L]	density [kg/m3]	
0,09	148,159	19752,90341	1,037273	1037,273	
0,10	148,021	19734,50996	1,0427	1042,7	
0,11	147,8489	19711,56627	1,049333	1049,333	
0,13	147,6282	19682,14551	1,057625	1057,625	
0,14	147,335	19643,05417	1,068286	1068,286	
0,17	146,9265	19588,58645	1,0825	1082,5	
0,20	146,3179	19507,44889	1,1024	1102,4	
0,25	145,3147	19373,70308	1,13225	1132,25	

1000L		70 degrees			
sugar [%]	Vapour pressure [mm Hg]	Vapour pressure [Pa]	density [kg/L]	density [kg/m3]	
0,09	231,7488	30897,29	1,032727	1032,727	
0,10	231,5319	30868,38	1,0382	1038,2	
0,11	231,2613	30832,31	1,044889	1044,889	
0,13	230,9144	30786,06	1,05325	1053,25	
0,14	230,4535	30724,6	1,064	1064	
0,17	229,8112	30638,98	1,078333	1078,333	
0,20	228,8546	30511,43	1,0984	1098,4	
0,25	227,2777	30301,2	1,1285	1128,5	

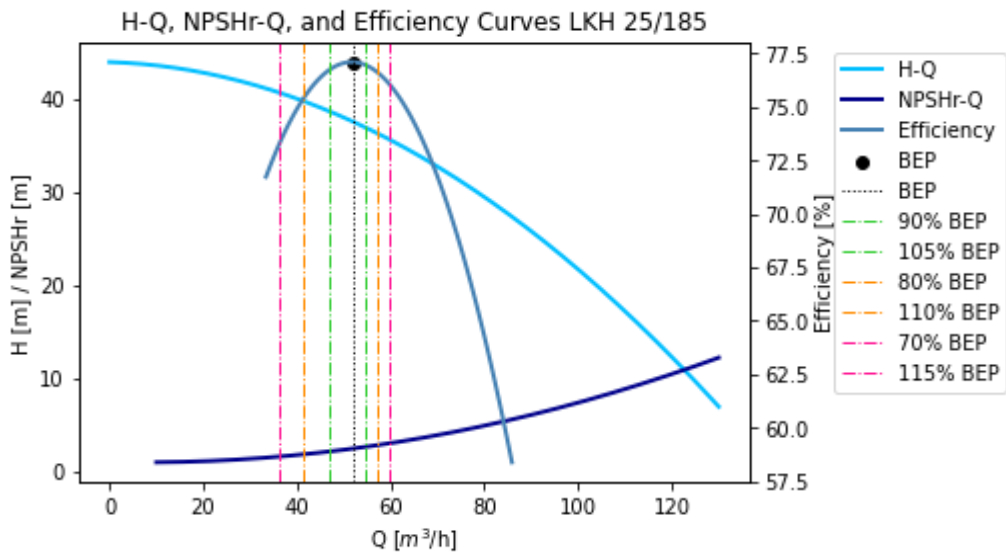
1000L		80 degrees			
sugar [%]	Vapour pressure [mm Hg]	Vapour pressure [Pa]	density [kg/L]	density [kg/m3]	
0	355,1542	47350	0,972	972	
0,09	352,1416	46948,35	1,027273	1027,273	
0,10	351,81	46904,15	1,0328	1032,8	
0,11	351,3964	46849	1,039556	1039,556	
0,13	350,8661	46778,3	1,048	1048	
0,14	350,1614	46684,35	1,058857	1058,857	
0,17	349,1797	46553,46	1,073333	1073,333	
0,20	347,7173	46358,5	1,0936	1093,6	
0,25	345,3071	46037,16	1,124	1124	

1000L		90 degrees			
sugar [%]	Vapour pressure [mm Hg]	Vapour pressure [Pa]	density [kg/L]	density [kg/m3]	
0,09	521,3755	69511,01423	1,020909	1020,909	
0,10	520,881	69445,09129	1,0265	1026,5	
0,11	520,2643	69362,86324	1,033333	1033,333	
0,13	519,4734	69257,42716	1,041875	1041,875	
0,14	518,4227	69117,34333	1,052857	1052,857	
0,17	516,9588	68922,1753	1,0675	1067,5	
0,20	514,7784	68631,48116	1,088	1088	
0,25	511,1851	68152,40208	1,11875	1118,75	

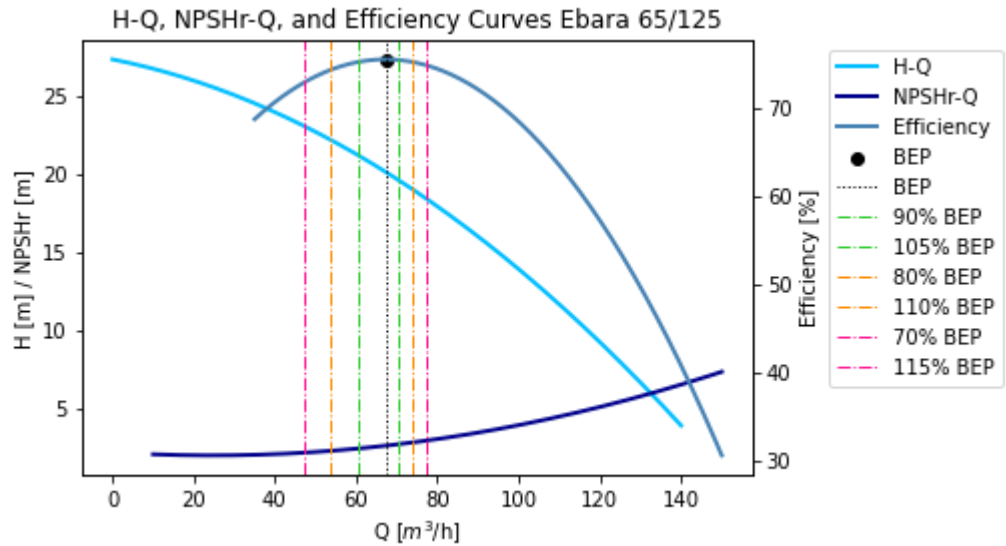
1000L		100 degrees			
sugar [%]	Vapour pressure [mm Hg]	Vapour pressure [Pa]	density [kg/L]	density [kg/m3]	
0	759,8125021	101300,002	0,958	958	
0,09	753,273994	100428,2726	1,014545455	1014,545	
0,10	752,5544329	100332,339	1,0202	1020,2	
0,11	751,6569125	100212,6795	1,027111111	1027,111	
0,13	750,5060985	100059,2503	1,03575	1035,75	
0,14	748,9771504	99855,40727	1,046857143	1046,857	
0,17	746,8470589	99571,41843	1,061666667	1061,667	
0,20	743,6745427	99148,45105	1,0824	1082,4	
0,25	738,44649	98451,43468	1,1135	1113,5	

Appendix D: Pump curves

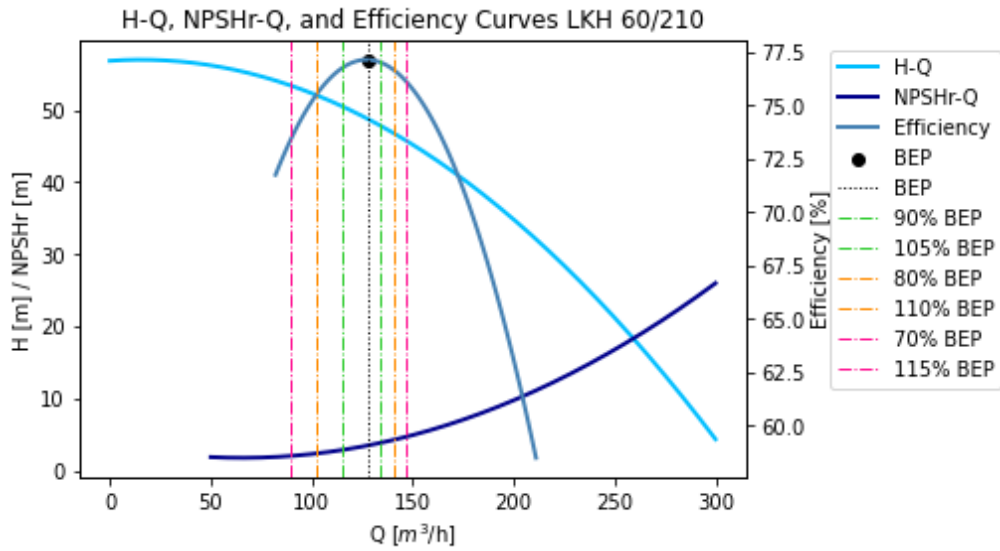
Pump curve of P_{1.1} and P_{2.2}



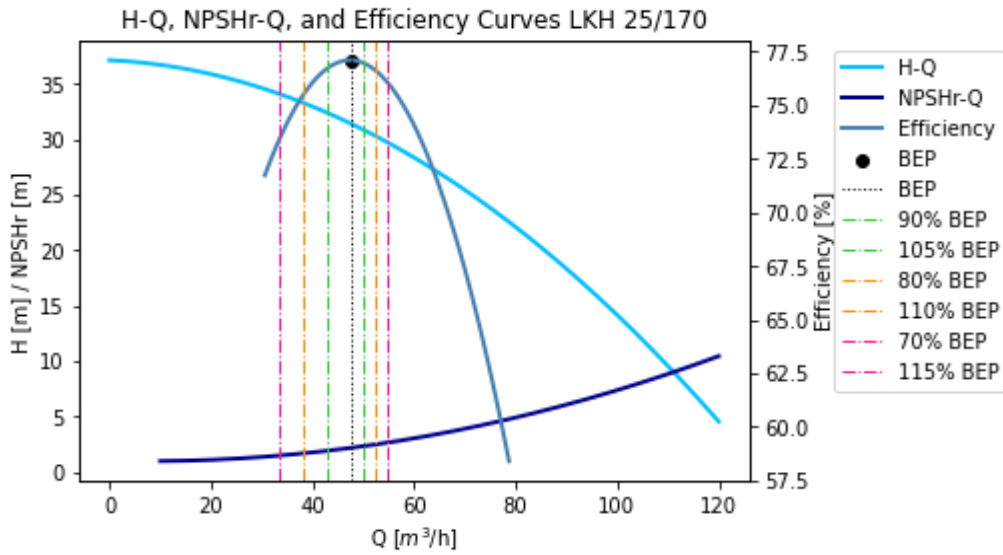
Pump curve of P_{2.1}



Pump curve of P_{3.1}



Pump curve of P_{4.1} and P_{5.1}



Appendix E: Method selection

<i>Stochastic</i>	Advantages	Disadvantages	When to consider	When to avoid
Aggregate reliability functions	Simple and well understood by reliability engineering community	Failures must be statistically independent and identically distributed	Sample size is statistically significant and representative of individual sample; and	Only a small number of failures can be attributed to individual failure modes; or
	Numerous software options available	In most cases will require a statistically significant sample size pertaining to each failure mode for reliable RUL predictions	Small set of dominant failure modes; and	Significant number of possible failure modes that cannot be easily differentiated, or historically have not been; or
	Theoretically can be performed at all equipment hierarchy levels, especially when a small number of failure modes dominate	Warnings prior to actual failure are not readily available	PDF is not exponential; and	Hazard rate is constant; or
	Confidence limits are available for RUL predictions		Reliability growth is not occurring (unless specific growth model used); and	Past operating conditions are not representative of current environment or usage; or
			Condition monitoring data is not available; and	The specific asset is critical to plant safety or operations (e.g. unsparred) and warning is required prior to failure
			RUL prediction is predominantly used for overall maintenance management rather than tracking of a specific asset (e.g. when redundancy is available) so gradual escalation of warning levels are not required	
RUL PDF	Simple and easy adaptation of basic reliability approaches	Available accuracy and precision is dependent on forecasting interval	Sample size is statistically significant and representative of individual sample; and	Only a small number of failures can be attributed to individual failure modes; or
	Only requires that time at which failure has not occurred is monitored	In most cases will require a statistically significant sample size pertaining to each	Small set of dominant failure modes; and	Significant number of possible failure modes that cannot be easily differentiated, or historically have not been; or

(i.e. no condition monitoring data)	failure mode for reliable RUL predictions.		
Theoretically can be performed at all equipment hierarchy levels, especially when a small number of failure modes dominate	Assumes that hazard is a function of operating time rather than external risk factors	PDF is not exponential; and	Hazard rate is constant; or
Confidence limits are available for RUL predictions		Reliability growth is not occurring; and	Past operating conditions are not representative of current environment or usage; or
Accuracy and precision increases as RUL decreases resulting in the ability to set useful warning limits		Condition monitoring data is not available; and	Failure is hidden and no failure finding is being undertaken; or
		Operating age can be tracked to confirm absence of failure; and	High level of accuracy and precision is required a long time into the future
		Only final estimates need to be particularly accurate and precise	

--	--	--	--

Static Bayesian Networks	Can readily manage incomplete data sets	Cannot model previously unanticipated faults and/or root causes	Incomplete, multivariate data available; and	Root causes of failure unknown; or
	Allow/force user to learn about causal relationship;	Computational difficulty of exploring a previously unknown network	Root causes of failure known; and	Expert plant and modelling knowledge unavailable; or
	Captures and integrates expert knowledge	A Bayesian network is only as useful as the prior knowledge is reliable	Process and plant configuration is relatively static or network is confirmed up to date; and	Training data unavailable
	Algorithms available to avoid the over fitting of data	Results may be sensitive to selection of prior distribution	Modelling experts are available	
	Computer software available for modelling	Modelling experts required in addition to domain experts		
	Confidence limits are intrinsically provided.			

Markov, Semi-Markov models	Well established approach and able to model numerous system designs and failure scenarios	Reasonably large volume of data required for training;	Simple to develop and implement;	Repairable system; or
	Can readily manage incomplete data sets	Assumes a single monotonic, non temporal failure degradation pattern (i.e. different stages of failure cannot be accounted for)	Incomplete, multivariate data available; and	Temporal measurement data as model inputs; or
	Computationally efficient once developed	Cannot model previously unanticipated faults and/or root causes	Root causes of failure known; and	Sufficient data related to failure mode is not available for training; or
	Provide confidence limits as part of their RUL prediction	More complex semi-Markov models are required if failures or failure progression times are not exponentially distributed	Process and plant configuration is relatively static or network is confirmed up to date; and	Failure being modelled has more than one discrete stage (e.g. Crack initiation, growth, final failure etc)
		Not appropriate for repairable systems that are only partially restored	Relatively accurate and precise RUL estimate required	
Hidden Markov, semi-Markov models	Can model different stages of degradation so failure trend does not need to be monotonic	Large volume of data required for training, proportional to number of hidden states	Repairable systems; and	Sufficient data related to failure mode is not available for training; or
	Can model spatial and temporal data	Cannot model previously unanticipated faults and/or root causes	Root causes of failure known; and	Suitable hardware for computation is not available
	Specific knowledge of failure mechanism progression is not required	More complex Hidden semi-Markov models are required if failures or failure progression times are not exponentially distributed	Failure being modelled has more than one discrete stage (e.g. Crack initiation, growth, final failure, etc.)	
	Can readily manage incomplete data sets	Computationally intensive, particularly	Temporal data to be used as model inputs	

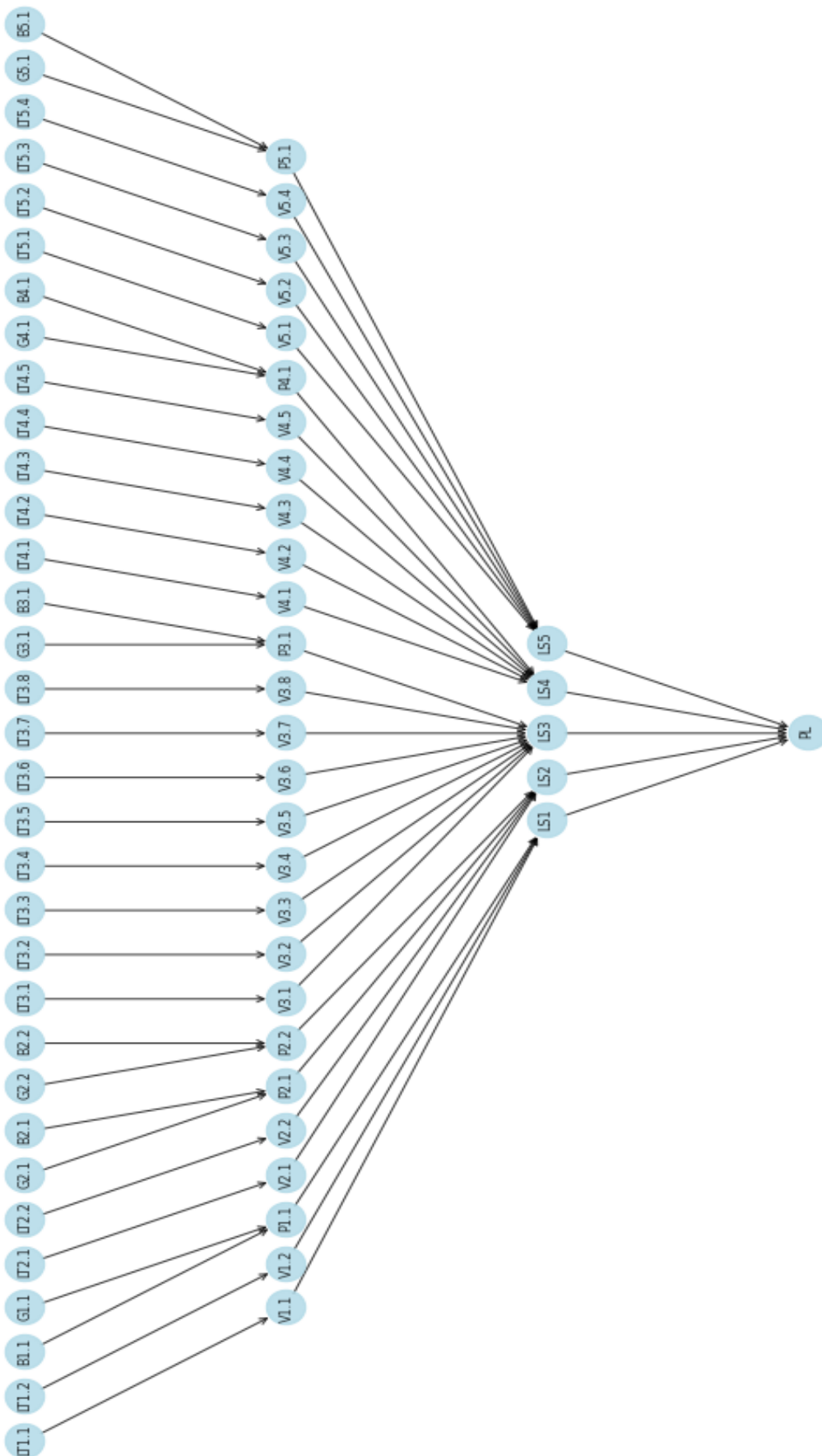
		for a large number of hidden states		
	Provide confidence limits as part of their RUL prediction		Relatively accurate and precise RUL required	
Bayesian techniques with Kalman Filters	Can be used to model multivariate, dynamic processes	Process and measurement noise must be Gaussian	Multivariate posterior distribution; and	Multiplicative noise; or
	Basic KF is computationally efficient, particularly for systems with a large number of states	Some variants diverge easily	Additive noise; and	Single variable posterior distribution; or
	Can accommodate incomplete and noisy measurements	Variants for non-linear systems are more computationally intensive than basic Kalman filters	Condition monitoring data is available; and	Covariate data is not available for the failures of interest.
	Variants available for non-linear processes	Measurement data required	Relatively accurate and precise RUL estimate required	
	Other advantages depend on underlying Bayesian technique	Other disadvantages depend on underlying Bayesian technique		
Bayesian techniques with Particle Filters	Can be used to model multivariate, dynamic processes	A large number of samples (or resampling) are required to avoid degeneracy problem	Multivariate or non-standard posterior distribution	Typical deterministic posterior distribution; or
	Noise does not need to be either linear or Gaussian	Can be more computationally intensive than basic Kalman filters	Non-linear, non-Gaussian noise; and	Linear, Gaussian noise; or
	More accurate than Kalman filter variants for non-linear systems	Measurement data required	Relatively accurate and precise RUL estimate required	Multiplicative noise; or
	Other advantages depend on underlying Bayesian technique	Other disadvantages depend on underlying Bayesian technique		Single variable posterior distribution; or

				Covariate data is not available for the failures of interest
<i>Statistical</i>				
Trend extrapolation	Simplest technique to apply and explain	Few failures have a well defined monotonic, single-parameter trend	Single defined failure mode associated with a single monitored (or calculated) parameter that can be described with a monotonic trend; and	Incipient failure cannot be related to a simple measurable input; or
	Easy to set alarms	Interpretability is affected by process/measurement noise and variations in operating conditions	Operating conditions are stable or do not affect monitored parameter; and	Varying operating conditions that affect the measured parameter but are not related to failure; or
	Advanced software tools not required	Availability of confidence limits dependant on amount of data at the different states of failure development	Measurements are repeatable, reliable and not highly sensitive to measurement processes (e.g. online sensors)	Trend is not monotonic; or
				Data highly dependent on measurement process; or
				Data is subject to high levels of process or measurement noise; or
				Reliable confidence limits are required on the extrapolated RUL estimate
ARMA Models & variants	Advanced ARMA related techniques available for non-stationary data	Basic ARMA models assume stationarity of the process and noise	Hazard rate is a linear relationship of covariates and noise; and	Hazard rate is not a linear relationship of covariates and noise; or
	Historical failure data is not required	Does not integrate prior or expert knowledge	Short-term predictions required; and	When historical or expert data is available in addition to measurement data; or
	Usually computationally efficient and therefore can be performed in real-time	Sensitive to noise and initial conditions	Hazard rate is independent of age (i.e. exponential distribution); and	Long term predictions are required; or
	An understanding of detailed failure	Significant data required for model	Measurement data is available for modelling and application but	Sufficiently large volume of data is not available for model construction and validation

	mechanisms not required	development and validation	historical failure data is not	
	Provide accurate and reliable short term predictions of RUL	Long term predictions of RUL are less reliable		
PHM	COTS software available	All relevant covariates must be included in the model	Times to failure are independent and identically distributed;	Failures have not occurred previously or have no associated covariate data
	Accounts for age dependant and independent hazards	Mixing different types of covariates in one model may be problematic	Covariates have a multiplicative effect on the baseline hazard rate; and	Hazard rate is not multiplicative; or
	Models are simple to develop	Strict (albeit implied) assumptions regarding nature of underlying process	A number of covariates are available and required to describe change in risk; and	Failures cannot be segregated into individual (or dominating) failure modes; or
	Confidence limits can be calculated	Historical data required pertaining to individual failure modes	Process represented by the covariates is stationary (unless using Dynamic PHM); and	Covariates related to the failure modes being modelled cannot be measured; or
		Multi-collinearity, monotonicity and large covariate values that can cause a failure of the model parameter estimation process	Associated covariate data is available for the failure modes being modelled; and	Process represented by the covariates is non-stationary (unless using Dynamic PHM); or
		No clear guidelines on the selection of parametric estimation technique	Only the final RUL estimate and confidence limit is required (not an estimate of a precursor to failure)	If a precursor to failure is to be predicted rather than final failure itself
		Parameter selection often manual and time consuming		
		Traditional PHM equation assumes covariates describe a stationary process. Dynamic PHM is more involved		
		Can only be used to develop models for failures that have been experienced		

		previously and for which associate covariate data is available		
		Too easy to develop a model that may be statistically adequate but does not represent any actual failure phenomenon (i.e. is physically meaningless		

Appendix F: Overall Bayesian Network



Appendix G: Conditional probability tables

Node $LT_{V1.1}$	Probability
LT1.1(1)	0,04
LT1.1(2)	0,04
LT1.1(3)	0,04
LT1.1(4)	0,04
LT1.1(5)	0,04
LT1.1(6)	0,04
LT1.1(7)	0,04
LT1.1(8)	0,04
LT1.1(9)	0,04
LT1.1(10)	0,04
LT1.1(11)	0,04
LT1.1(12)	0,04
LT1.1(13)	0,04
LT1.1(14)	0,04
LT1.1(15)	0,04
LT1.1(16)	0,04
LT1.1(17)	0,04
LT1.1(18)	0,04
LT1.1(19)	0,04
LT1.1(20)	0,04
LT1.1(21)	0,04
LT1.1(22)	0,04
LT1.1(23)	0,04
LT1.1(24)	0,04
LT1.1(25)	0,04

Node $LT_{V1.2}$	Probability
LT1.2(1)	0,03333
LT1.2(2)	0,03333
LT1.2(3)	0,03333
LT1.2(4)	0,03333
LT1.2(5)	0,03333
LT1.2(6)	0,03333
LT1.2(7)	0,03333
LT1.2(8)	0,03333
LT1.2(9)	0,03333
LT1.2(10)	0,03333
LT1.2(11)	0,03333
LT1.2(12)	0,03333
LT1.2(13)	0,03333
LT1.2(14)	0,03333
LT1.2(15)	0,03333
LT1.2(16)	0,03333
LT1.2(17)	0,03333
LT1.2(18)	0,03333
LT1.2(19)	0,03333
LT1.2(20)	0,03333
LT1.2(21)	0,03333
LT1.2(22)	0,03333
LT1.2(23)	0,03333
LT1.2(24)	0,03333
LT1.2(25)	0,03333
LT1.2(26)	0,03333
LT1.2(27)	0,03333
LT1.2(28)	0,03333
LT1.2(29)	0,03333
LT1.2(30)	0,03333

Node $G_{P1.1}$	Probability
G1.1(1)	0,5
G1.1(2)	0,5

Node $B_{P1.1}$	Probability
B1.1(1)	0,05
B1.1(2)	0,10
B1.1(3)	0,35
B1.1(4)	0,30
B1.1(5)	0,20

$G1.1$	$B1.1$	$P1.1(0)$	$P1.1(1)$
G1.1(1)	B1.1(1)	0,95	0,05
G1.1(1)	B1.1(2)	0,92	0,08
G1.1(1)	B1.1(3)	0,53	0,47
G1.1(1)	B1.1(4)	0,10	0,90
G1.1(1)	B1.1(5)	0,05	0,95
G1.1(2)	B1.1(1)	0,30	0,70
G1.1(2)	B1.1(2)	0,20	0,80
G1.1(2)	B1.1(3)	0,15	0,85
G1.1(2)	B1.1(4)	0,10	0,90
G1.1(2)	B1.1(5)	0,01	0,99

$LT1.1$	$V1.1(0)$	$V1.1(1)$
LT1.1(1)	0,9964	0,0036
LT1.1(2)	0,9945	0,0055
LT1.1(3)	0,9916	0,0084
LT1.1(4)	0,9871	0,0129
LT1.1(5)	0,9804	0,0196
LT1.1(6)	0,9704	0,0296
LT1.1(7)	0,9554	0,0446
LT1.1(8)	0,9333	0,0667
LT1.1(9)	0,9014	0,0986
LT1.1(10)	0,8566	0,1434
LT1.1(11)	0,796	0,204
LT1.1(12)	0,7183	0,2817
LT1.1(13)	0,6249	0,3751
LT1.1(14)	0,5213	0,4787
LT1.1(15)	0,4157	0,5843
LT1.1(16)	0,3174	0,6826
LT1.1(17)	0,233	0,767
LT1.1(18)	0,1656	0,8344
LT1.1(19)	0,1148	0,8852
LT1.1(20)	0,0782	0,9218
LT1.1(21)	0,0525	0,9475
LT1.1(22)	0,0349	0,9651
LT1.1(23)	0,0231	0,9769
LT1.1(24)	0,0152	0,9848
LT1.1(25)	0,0100	0,9900

LT1.2	V1.2(0)	V1.2(1)
LT1.2(1)	0,9956	0,0044
LT1.2(2)	0,9938	0,0062
LT1.2(3)	0,9913	0,0087
LT1.2(4)	0,9877	0,0123
LT1.2(5)	0,9827	0,0173
LT1.2(6)	0,9758	0,0242
LT1.2(7)	0,9661	0,0339
LT1.2(8)	0,9528	0,0472
LT1.2(9)	0,9346	0,0654
LT1.2(10)	0,9101	0,0899
LT1.2(11)	0,8775	0,1225
LT1.2(12)	0,8353	0,1647
LT1.2(13)	0,7822	0,2178
LT1.2(14)	0,7177	0,2823
LT1.2(15)	0,6428	0,3572
LT1.2(16)	0,5602	0,4398
LT1.2(17)	0,4741	0,5259
LT1.2(18)	0,3896	0,6104
LT1.2(19)	0,3112	0,6888
LT1.2(20)	0,2423	0,7577
LT1.2(21)	0,1846	0,8154
LT1.2(22)	0,1381	0,8619
LT1.2(23)	0,1019	0,8981
LT1.2(24)	0,0743	0,9257
LT1.2(25)	0,0538	0,9462
LT1.2(26)	0,0387	0,9613
LT1.2(27)	0,0277	0,9723
LT1.2(28)	0,0198	0,9802
LT1.2(29)	0,0141	0,9859
LT1.2(30)	0,0101	0,9900

V1.1	V1.2	P1.1	LS1(0)	LS1(1)
V1.1(0)	V1.2(0)	P1.1(0)	0,99	0,01
V1.1(0)	V1.2(0)	P1.1(1)	0,40	0,60
V1.1(0)	V1.2(1)	P1.1(0)	0,20	0,80
V1.1(0)	V1.2(1)	P1.1(1)	0,30	0,70
V1.1(1)	V1.2(0)	P1.1(0)	0,50	0,50
V1.1(1)	V1.2(0)	P1.1(1)	0,30	0,70
V1.1(1)	V1.2(1)	P1.1(0)	0,20	0,80
V1.1(1)	V1.2(1)	P1.1(1)	0,01	0,99

Node LTV2.1	Probability
LT2.1(1)	0,0666667
LT2.1(2)	0,0666667
LT2.1(3)	0,0666667
LT2.1(4)	0,0666667
LT2.1(5)	0,0666667
LT2.1(6)	0,0666667
LT2.1(7)	0,0666667
LT2.1(8)	0,0666667
LT2.1(9)	0,0666667
LT2.1(10)	0,0666667
LT2.1(11)	0,0666667
LT2.1(12)	0,0666667
LT2.1(13)	0,0666667
LT2.1(14)	0,0666667
LT2.1(15)	0,0666667

Node LTV2.2	Probability
LT2.2(1)	0,0666667
LT2.2(2)	0,0666667
LT2.2(3)	0,0666667
LT2.2(4)	0,0666667
LT2.2(5)	0,0666667
LT2.2(6)	0,0666667
LT2.2(7)	0,0666667
LT2.2(8)	0,0666667
LT2.2(9)	0,0666667
LT2.2(10)	0,0666667
LT2.2(11)	0,0666667
LT2.2(12)	0,0666667
LT2.2(13)	0,0666667
LT2.2(14)	0,0666667
LT2.2(15)	0,0666667

Node Gp2.1	Probability
G2.1(1)	0,5
G2.1(2)	0,5

Node Gp2.2	Probability
G2.2(1)	0,5
G2.2(2)	0,5

Node Bp1.2	Probability
B2.1(1)	0,02
B2.1(2)	0,10
B2.1(3)	0,35
B2.1(4)	0,33
B2.1(5)	0,20

Node Bp2.2	Probability
B2.2(1)	0,05
B2.2(2)	0,09
B2.2(3)	0,33
B2.2(4)	0,32
B2.2(5)	0,21

G2.1	B2.1	P2.1(0)	P2.1(1)
G2.1(1)	B2.1(1)	0,95	0,05
G2.1(1)	B2.1(2)	0,92	0,08
G2.1(1)	B2.1(3)	0,53	0,47
G2.1(1)	B2.1(4)	0,10	0,90
G2.1(1)	B2.1(5)	0,05	0,95
G2.1(2)	B2.1(1)	0,25	0,75
G2.1(2)	B2.1(2)	0,20	0,80
G2.1(2)	B2.1(3)	0,15	0,85
G2.1(2)	B2.1(4)	0,10	0,90
G2.1(2)	B2.1(5)	0,01	0,99

G2.2	B2.2	P2.2(0)	P2.2(1)
G2.2(1)	B2.2(1)	0,95	0,05
G2.2(1)	B2.2(2)	0,92	0,08
G2.2(1)	B2.2(3)	0,53	0,47
G2.2(1)	B2.2(4)	0,10	0,90
G2.2(1)	B2.2(5)	0,05	0,95
G2.2(2)	B2.2(1)	0,40	0,60
G2.2(2)	B2.2(2)	0,35	0,65
G2.2(2)	B2.2(3)	0,15	0,85
G2.2(2)	B2.2(4)	0,10	0,90
G2.2(2)	B2.2(5)	0,01	0,99

LT2.1	V2.1(0)	V2.1(1)	LT2.2	V2.2(0)	V2.2(1)
LT2.1(1)	0,9747	0,0253	LT2.2(1)	0,9958	0,0042
LT2.1(2)	0,9554	0,0446	LT2.2(2)	0,9913	0,0087
LT2.1(3)	0,9223	0,0777	LT2.2(3)	0,9824	0,0176
LT2.1(4)	0,8682	0,1318	LT2.2(4)	0,9645	0,0355
LT2.1(5)	0,7852	0,2148	LT2.2(5)	0,9299	0,0701
LT2.1(6)	0,6697	0,3303	LT2.2(6)	0,8661	0,1339
LT2.1(7)	0,5294	0,4706	LT2.2(7)	0,7593	0,2407
LT2.1(8)	0,3843	0,6157	LT2.2(8)	0,6061	0,3939
LT2.1(9)	0,2572	0,7428	LT2.2(9)	0,4287	0,5713
LT2.1(10)	0,1612	0,8388	LT2.2(10)	0,2679	0,7321
LT2.1(11)	0,0963	0,9037	LT2.2(11)	0,1515	0,8485
LT2.1(12)	0,0558	0,9442	LT2.2(12)	0,0801	0,9199
LT2.1(13)	0,0318	0,9682	LT2.2(13)	0,0407	0,9593
LT2.1(14)	0,0179	0,9821	LT2.2(14)	0,0203	0,9797
LT2.1(15)	0,0100	0,9900	LT2.2(15)	0,0100	0,9900

V2.1	V2.2	P2.1	P2.2	LS1(0)	LS1(1)
V2.1(0)	V2.2(0)	P2.1(0)	P2.2(0)	0,99	0,01
V2.1(0)	V2.2(0)	P2.1(0)	P2.2(1)	0,45	0,55
V2.1(0)	V2.2(0)	P2.1(1)	P2.2(0)	0,49	0,51
V2.1(0)	V2.2(0)	P2.1(1)	P2.2(1)	0,19	0,81
V2.1(0)	V2.2(1)	P2.1(0)	P2.2(0)	0,48	0,52
V2.1(0)	V2.2(1)	P2.1(0)	P2.2(1)	0,35	0,65
V2.1(0)	V2.2(1)	P2.1(1)	P2.2(0)	0,31	0,69
V2.1(0)	V2.2(1)	P2.1(1)	P2.2(1)	0,23	0,77
V2.1(1)	V2.2(0)	P2.1(0)	P2.2(0)	0,44	0,56
V2.1(1)	V2.2(0)	P2.1(0)	P2.2(1)	0,22	0,78
V2.1(1)	V2.2(0)	P2.1(1)	P2.2(0)	0,28	0,72
V2.1(1)	V2.2(0)	P2.1(1)	P2.2(1)	0,20	0,80
V2.1(1)	V2.2(1)	P2.1(0)	P2.2(0)	0,25	0,75
V2.1(1)	V2.2(1)	P2.1(0)	P2.2(1)	0,15	0,85
V2.1(1)	V2.2(1)	P2.1(1)	P2.2(0)	0,05	0,95
V2.1(1)	V2.2(1)	P2.1(1)	P2.2(1)	0,01	0,99

Node LTV3.1	Probability
LT3.1(1)	0,05
LT3.1(2)	0,05
LT3.1(3)	0,05
LT3.1(4)	0,05
LT3.1(5)	0,05
LT3.1(6)	0,05
LT3.1(7)	0,05
LT3.1(8)	0,05
LT3.1(9)	0,05
LT3.1(10)	0,05
LT3.1(11)	0,05
LT3.1(12)	0,05
LT3.1(13)	0,05
LT3.1(14)	0,05
LT3.1(15)	0,05
LT3.1(16)	0,05
LT3.1(17)	0,05
LT3.1(18)	0,05
LT3.1(19)	0,05
LT3.1(20)	0,05

Node LTV3.2	Probability
LT3.2(1)	0,05
LT3.2(2)	0,05
LT3.2(3)	0,05
LT3.2(4)	0,05
LT3.2(5)	0,05
LT3.2(6)	0,05
LT3.2(7)	0,05
LT3.2(8)	0,05
LT3.2(9)	0,05
LT3.2(10)	0,05
LT3.2(11)	0,05
LT3.2(12)	0,05
LT3.2(13)	0,05
LT3.2(14)	0,05
LT3.2(15)	0,05
LT3.2(16)	0,05
LT3.2(17)	0,05
LT3.2(18)	0,05
LT3.2(19)	0,05
LT3.2(20)	0,05

Node LTV3.3	Probability
LT3.3(1)	0,0667
LT3.3(2)	0,0667
LT3.3(3)	0,0667
LT3.3(4)	0,0667
LT3.3(5)	0,0667
LT3.3(6)	0,0667
LT3.3(7)	0,0667
LT3.3(8)	0,0667
LT3.3(9)	0,0667
LT3.3(10)	0,0667
LT3.3(11)	0,0667
LT3.3(12)	0,0667
LT3.3(13)	0,0667
LT3.3(14)	0,0667
LT3.3(15)	0,0667

Node LTV3.6	Probability
LT3.6(1)	0,0333
LT3.6(2)	0,0333
LT3.6(3)	0,0333
LT3.6(4)	0,0333
LT3.6(5)	0,0333
LT3.6(6)	0,0333
LT3.6(7)	0,0333
LT3.6(8)	0,0333
LT3.6(9)	0,0333
LT3.6(10)	0,0333
LT3.6(11)	0,0333
LT3.6(12)	0,0333
LT3.6(13)	0,0333
LT3.6(14)	0,0333
LT3.6(15)	0,0333
LT3.6(16)	0,0333
LT3.6(17)	0,0333
LT3.6(18)	0,0333
LT3.6(19)	0,0333
LT3.6(20)	0,0333
LT3.6(21)	0,0333
LT3.6(22)	0,0333
LT3.6(23)	0,0333
LT3.6(24)	0,0333
LT3.6(25)	0,0333
LT3.6(26)	0,0333
LT3.6(27)	0,0333
LT3.6(28)	0,0333
LT3.6(29)	0,0333
LT3.6(30)	0,0333

Node LTV3.4	Probability
LT3.4(1)	0,0667
LT3.4(2)	0,0667
LT3.4(3)	0,0667
LT3.4(4)	0,0667
LT3.4(5)	0,0667
LT3.4(6)	0,0667
LT3.4(7)	0,0667
LT3.4(8)	0,0667
LT3.4(9)	0,0667
LT3.4(10)	0,0667
LT3.4(11)	0,0667
LT3.4(12)	0,0667
LT3.4(13)	0,0667
LT3.4(14)	0,0667
LT3.4(15)	0,0667

Node LTV3.5	Probability
LT3.5(1)	0,0667
LT3.5(2)	0,0667
LT3.5(3)	0,0667
LT3.5(4)	0,0667
LT3.5(5)	0,0667
LT3.5(6)	0,0667
LT3.5(7)	0,0667
LT3.5(8)	0,0667
LT3.5(9)	0,0667
LT3.5(10)	0,0667
LT3.5(11)	0,0667
LT3.5(12)	0,0667
LT3.5(13)	0,0667
LT3.5(14)	0,0667
LT3.5(15)	0,0667

Node LTV3.7	Probability
LT3.7(1)	0,0333
LT3.7(2)	0,0333
LT3.7(3)	0,0333
LT3.7(4)	0,0333
LT3.7(5)	0,0333
LT3.7(6)	0,0333
LT3.7(7)	0,0333
LT3.7(8)	0,0333
LT3.7(9)	0,0333
LT3.7(10)	0,0333
LT3.7(11)	0,0333
LT3.7(12)	0,0333
LT3.7(13)	0,0333
LT3.7(14)	0,0333
LT3.7(15)	0,0333
LT3.7(16)	0,0333
LT3.7(17)	0,0333
LT3.7(18)	0,0333
LT3.7(19)	0,0333
LT3.7(20)	0,0333
LT3.7(21)	0,0333
LT3.7(22)	0,0333
LT3.7(23)	0,0333
LT3.7(24)	0,0333
LT3.7(25)	0,0333
LT3.7(26)	0,0333
LT3.7(27)	0,0333
LT3.7(28)	0,0333
LT3.7(29)	0,0333
LT3.7(30)	0,0333

Node LTV3.8	Probability
LT3.8(1)	0,0667
LT3.8(2)	0,0667
LT3.8(3)	0,0667
LT3.8(4)	0,0667
LT3.8(5)	0,0667
LT3.8(6)	0,0667
LT3.8(7)	0,0667
LT3.8(8)	0,0667
LT3.8(9)	0,0667
LT3.8(10)	0,0667
LT3.8(11)	0,0667
LT3.8(12)	0,0667
LT3.8(13)	0,0667
LT3.8(14)	0,0667
LT3.8(15)	0,0667

Node Gp3.1	Probability
G3.1(1)	0,5
G3.1(2)	0,5

Node Bp3.1	Probability
B3.1(1)	0,05
B3.1(2)	0,08
B3.1(3)	0,35
B3.1(4)	0,32
B3.1(5)	0,2

G3.1	B3.1	P3.1(0)	P3.1(1)
G3.1(1)	B3.1(1)	0,95	0,05
G3.1(1)	B3.1(2)	0,92	0,08
G3.1(1)	B3.1(3)	0,53	0,47
G3.1(1)	B3.1(4)	0,10	0,90
G3.1(1)	B3.1(5)	0,05	0,95
G3.1(2)	B3.1(1)	0,28	0,72
G3.1(2)	B3.1(2)	0,19	0,81
G3.1(2)	B3.1(3)	0,15	0,85
G3.1(2)	B3.1(4)	0,09	0,91
G3.1(2)	B3.1(5)	0,01	0,99

LT3.1	V3.1(0)	V3.1(1)
LT3.1(1)	0,997	0,003
LT3.1(2)	0,995	0,005
LT3.1(3)	0,991	0,009
LT3.1(4)	0,985	0,015
LT3.1(5)	0,974	0,026
LT3.1(6)	0,955	0,045
LT3.1(7)	0,925	0,075
LT3.1(8)	0,878	0,123
LT3.1(9)	0,806	0,194
LT3.1(10)	0,706	0,294
LT3.1(11)	0,581	0,419
LT3.1(12)	0,446	0,555
LT3.1(13)	0,317	0,683
LT3.1(14)	0,212	0,788
LT3.1(15)	0,135	0,865
LT3.1(16)	0,083	0,917
LT3.1(17)	0,050	0,951
LT3.1(18)	0,029	0,971
LT3.1(19)	0,017	0,983
LT3.1(20)	0,010	0,990

LT3.2	V3.2(0)	V3.2(1)
LT3.2(1)	0,9926	0,0074
LT3.2(2)	0,9878	0,0122
LT3.2(3)	0,9801	0,0199
LT3.2(4)	0,9676	0,0324
LT3.2(5)	0,9477	0,0523
LT3.2(6)	0,9166	0,0834
LT3.2(7)	0,8697	0,1303
LT3.2(8)	0,8019	0,1981
LT3.2(9)	0,7108	0,2892
LT3.2(10)	0,5986	0,4014
LT3.2(11)	0,475	0,525
LT3.2(12)	0,3545	0,6455
LT3.2(13)	0,25	0,75
LT3.2(14)	0,1682	0,8318
LT3.2(15)	0,1093	0,8907
LT3.2(16)	0,0693	0,9307
LT3.2(17)	0,0432	0,9568
LT3.2(18)	0,0267	0,9733
LT3.2(19)	0,0164	0,9836
LT3.2(20)	0,01	0,99

LT3.3	V3.3(0)	V3.3(1)
LT3.3(1)	0,9912	0,0088
LT3.3(2)	0,9831	0,0169
LT3.3(3)	0,9676	0,0324
LT3.3(4)	0,9388	0,0612
LT3.3(5)	0,8874	0,1126
LT3.3(6)	0,8019	0,1981
LT3.3(7)	0,6754	0,3246
LT3.3(8)	0,5166	0,4834
LT3.3(9)	0,3545	0,6455
LT3.3(10)	0,2201	0,7799
LT3.3(11)	0,1266	0,8734
LT3.3(12)	0,0693	0,9307
LT3.3(13)	0,0369	0,9631
LT3.3(14)	0,0193	0,9807
LT3.3(15)	0,01	0,99

LT3.4	V3.4(0)	V3.4(1)
LT3.4(1)	0,9855	0,0145
LT3.4(2)	0,9731	0,0269
LT3.4(3)	0,9507	0,0493
LT3.4(4)	0,9113	0,0887
LT3.4(5)	0,8455	0,1545
LT3.4(6)	0,7446	0,2554
LT3.4(7)	0,6084	0,3916
LT3.4(8)	0,4529	0,5471
LT3.4(9)	0,3061	0,6939
LT3.4(10)	0,1903	0,8097
LT3.4(11)	0,1113	0,8887
LT3.4(12)	0,0626	0,9374
LT3.4(13)	0,0344	0,9656
LT3.4(14)	0,0186	0,9814
LT3.4(15)	0,01	0,99

LT3.5	V3.5(0)	V3.5(1)
LT3.5(1)	0,9934	0,0066
LT3.5(2)	0,9869	0,0131
LT3.5(3)	0,9743	0,0257
LT3.5(4)	0,9502	0,0498
LT3.5(5)	0,9058	0,0942
LT3.5(6)	0,8288	0,1712
LT3.5(7)	0,7092	0,2908
LT3.5(8)	0,5513	0,4487
LT3.5(9)	0,3822	0,6178
LT3.5(10)	0,2376	0,7624
LT3.5(11)	0,1357	0,8643
LT3.5(12)	0,0733	0,9267
LT3.5(13)	0,0383	0,9617
LT3.5(14)	0,0197	0,9803
LT3.5(15)	0,01	0,99

LT3.6	V3.6(0)	V3.6(1)
LT3.6(1)	0,9937	0,0063
LT3.6(2)	0,9912	0,0088
LT3.6(3)	0,9878	0,0122
LT3.6(4)	0,9831	0,0169
LT3.6(5)	0,9766	0,0234
LT3.6(6)	0,9676	0,0324
LT3.6(7)	0,9554	0,0446
LT3.6(8)	0,9388	0,0612
LT3.6(9)	0,9166	0,0834
LT3.6(10)	0,8874	0,1126
LT3.6(11)	0,8496	0,1504
LT3.6(12)	0,8019	0,1981
LT3.6(13)	0,7437	0,2563
LT3.6(14)	0,6754	0,3246
LT3.6(15)	0,5986	0,4014
LT3.6(16)	0,5166	0,4834
LT3.6(17)	0,4338	0,5662
LT3.6(18)	0,3545	0,6455
LT3.6(19)	0,2824	0,7176
LT3.6(20)	0,2201	0,7799
LT3.6(21)	0,1682	0,8318
LT3.6(22)	0,1266	0,8734
LT3.6(23)	0,0941	0,9059
LT3.6(24)	0,0693	0,9307
LT3.6(25)	0,0507	0,9493
LT3.6(26)	0,0369	0,9631
LT3.6(27)	0,0267	0,9733
LT3.6(28)	0,0193	0,9807
LT3.6(29)	0,0149	0,9861
LT3.6(30)	0,01	0,99

LT3.7	V3.7(0)	V3.7(1)
LT3.7(1)	0,989	0,011
LT3.7(2)	0,985	0,015
LT3.7(3)	0,979	0,021
LT3.7(4)	0,972	0,028
LT3.7(5)	0,962	0,038
LT3.7(6)	0,948	0,052
LT3.7(7)	0,931	0,070
LT3.7(8)	0,907	0,093
LT3.7(9)	0,878	0,123
LT3.7(10)	0,840	0,160
LT3.7(11)	0,793	0,207
LT3.7(12)	0,737	0,263
LT3.7(13)	0,672	0,328
LT3.7(14)	0,600	0,400
LT3.7(15)	0,523	0,477
LT3.7(16)	0,446	0,555
LT3.7(17)	0,370	0,630
LT3.7(18)	0,301	0,699
LT3.7(19)	0,239	0,761
LT3.7(20)	0,187	0,813
LT3.7(21)	0,144	0,856
LT3.7(22)	0,110	0,890
LT3.7(23)	0,083	0,917
LT3.7(24)	0,062	0,938
LT3.7(25)	0,046	0,954
LT3.7(26)	0,034	0,966
LT3.7(27)	0,025	0,975
LT3.7(28)	0,019	0,982
LT3.7(29)	0,014	0,986
LT3.7(30)	0,010	0,990

LT3.8	V3.8(0)	V3.8(1)
LT3.8(1)	0,972	0,028
LT3.8(2)	0,951	0,049
LT3.8(3)	0,9157	0,0843
LT3.8(4)	0,8585	0,1415
LT3.8(5)	0,7723	0,2277
LT3.8(6)	0,6547	0,3453
LT3.8(7)	0,5145	0,4855
LT3.8(8)	0,372	0,628
LT3.8(9)	0,2488	0,7512
LT3.8(10)	0,1562	0,8438
LT3.8(11)	0,0938	0,9062
LT3.8(12)	0,0547	0,9453
LT3.8(13)	0,0313	0,9687
LT3.8(14)	0,0178	0,9822
LT3.8(15)	0,01	0,99

V3.1	V3.2	V3.3	V3.4	V3.5	V3.6	V3.7	V3.8	P3.1	LS3(0)	LS3(1)
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(0)	V3.7(0)	V3.8(0)	P3.1(0)	0,99	0,01
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(0)	V3.7(0)	V3.8(0)	P3.1(1)	0,40	0,60
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(0)	V3.7(0)	V3.8(1)	P3.1(0)	0,42	0,58
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(0)	V3.7(0)	V3.8(1)	P3.1(1)	0,38	0,62
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(0)	V3.7(1)	V3.8(0)	P3.1(0)	0,49	0,51
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(0)	V3.7(1)	V3.8(0)	P3.1(1)	0,38	0,62
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(0)	V3.7(1)	V3.8(1)	P3.1(0)	0,38	0,62
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(0)	V3.7(1)	V3.8(1)	P3.1(1)	0,28	0,72
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(1)	V3.7(0)	V3.8(0)	P3.1(0)	0,42	0,58
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(1)	V3.7(0)	V3.8(0)	P3.1(1)	0,38	0,62
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(1)	V3.7(0)	V3.8(1)	P3.1(0)	0,39	0,61
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(1)	V3.7(0)	V3.8(1)	P3.1(1)	0,28	0,72
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(1)	V3.7(1)	V3.8(0)	P3.1(0)	0,32	0,68
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(1)	V3.7(1)	V3.8(0)	P3.1(1)	0,28	0,72
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(1)	V3.7(1)	V3.8(1)	P3.1(0)	0,29	0,71
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(0)	V3.6(1)	V3.7(1)	V3.8(1)	P3.1(1)	0,34	0,66
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(1)	V3.6(0)	V3.7(0)	V3.8(0)	P3.1(0)	0,42	0,58
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(1)	V3.6(0)	V3.7(0)	V3.8(0)	P3.1(1)	0,41	0,59
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(1)	V3.6(0)	V3.7(0)	V3.8(1)	P3.1(0)	0,42	0,58
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(1)	V3.6(0)	V3.7(0)	V3.8(1)	P3.1(1)	0,28	0,72
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(1)	V3.6(0)	V3.7(1)	V3.8(0)	P3.1(0)	0,43	0,57
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(1)	V3.6(0)	V3.7(1)	V3.8(0)	P3.1(1)	0,28	0,72
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(1)	V3.6(0)	V3.7(1)	V3.8(1)	P3.1(0)	0,45	0,55
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(1)	V3.6(0)	V3.7(1)	V3.8(1)	P3.1(1)	0,41	0,59
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(1)	V3.6(1)	V3.7(0)	V3.8(0)	P3.1(0)	0,47	0,53
V3.1(0)	V3.2(0)	V3.3(0)	V3.4(0)	V3.5(1)	V3.6(1)	V3.7(0)	V3.8(0)	P3.1(1)	0,28	0,72

V3.1(1)	V3.2(1)	V3.3(1)	V3.4(0)	V3.5(1)	V3.6(1)	V3.7(0)	V3.8(0)	P3.1(0)	0,28	0,72
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(0)	V3.5(1)	V3.6(1)	V3.7(0)	V3.8(0)	P3.1(1)	0,21	0,79
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(0)	V3.5(1)	V3.6(1)	V3.7(0)	V3.8(1)	P3.1(0)	0,20	0,80
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(0)	V3.5(1)	V3.6(1)	V3.7(0)	V3.8(1)	P3.1(1)	0,18	0,82
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(0)	V3.5(1)	V3.6(1)	V3.7(1)	V3.8(0)	P3.1(0)	0,23	0,77
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(0)	V3.5(1)	V3.6(1)	V3.7(1)	V3.8(0)	P3.1(1)	0,11	0,89
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(0)	V3.5(1)	V3.6(1)	V3.7(1)	V3.8(1)	P3.1(0)	0,12	0,88
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(0)	V3.5(1)	V3.6(1)	V3.7(1)	V3.8(1)	P3.1(1)	0,06	0,94
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(0)	V3.7(0)	V3.8(0)	P3.1(0)	0,31	0,69
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(0)	V3.7(0)	V3.8(0)	P3.1(1)	0,32	0,68
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(0)	V3.7(0)	V3.8(1)	P3.1(0)	0,31	0,69
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(0)	V3.7(0)	V3.8(1)	P3.1(1)	0,30	0,70
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(0)	V3.7(1)	V3.8(0)	P3.1(0)	0,28	0,72
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(0)	V3.7(1)	V3.8(0)	P3.1(1)	0,17	0,83
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(0)	V3.7(1)	V3.8(1)	P3.1(0)	0,18	0,82
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(0)	V3.7(1)	V3.8(1)	P3.1(1)	0,08	0,92
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(1)	V3.7(0)	V3.8(0)	P3.1(0)	0,22	0,78
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(1)	V3.7(0)	V3.8(0)	P3.1(1)	0,12	0,88
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(1)	V3.7(0)	V3.8(1)	P3.1(0)	0,16	0,84
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(1)	V3.7(0)	V3.8(1)	P3.1(1)	0,11	0,89
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(1)	V3.7(1)	V3.8(0)	P3.1(0)	0,18	0,82
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(1)	V3.7(1)	V3.8(0)	P3.1(1)	0,09	0,91
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(1)	V3.7(1)	V3.8(1)	P3.1(0)	0,10	0,90
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(0)	V3.6(1)	V3.7(1)	V3.8(1)	P3.1(1)	0,13	0,87
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(1)	V3.6(0)	V3.7(0)	V3.8(0)	P3.1(0)	0,35	0,65
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(1)	V3.6(0)	V3.7(0)	V3.8(0)	P3.1(1)	0,12	0,88
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(1)	V3.6(0)	V3.7(0)	V3.8(1)	P3.1(0)	0,13	0,87
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(1)	V3.6(0)	V3.7(0)	V3.8(1)	P3.1(1)	0,14	0,86
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(1)	V3.6(0)	V3.7(1)	V3.8(0)	P3.1(0)	0,15	0,85
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(1)	V3.6(0)	V3.7(1)	V3.8(0)	P3.1(1)	0,12	0,88
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(1)	V3.6(0)	V3.7(1)	V3.8(1)	P3.1(0)	0,13	0,87
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(1)	V3.6(1)	V3.7(0)	V3.8(0)	P3.1(1)	0,20	0,80
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(1)	V3.6(1)	V3.7(0)	V3.8(1)	P3.1(0)	0,21	0,79
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(1)	V3.6(1)	V3.7(0)	V3.8(1)	P3.1(1)	0,11	0,89
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(1)	V3.6(1)	V3.7(1)	V3.8(0)	P3.1(0)	0,22	0,78
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(1)	V3.6(1)	V3.7(1)	V3.8(0)	P3.1(1)	0,08	0,92
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(1)	V3.6(1)	V3.7(1)	V3.8(1)	P3.1(0)	0,10	0,90
V3.1(1)	V3.2(1)	V3.3(1)	V3.4(1)	V3.5(1)	V3.6(1)	V3.7(1)	V3.8(1)	P3.1(1)	0,01	0,99

Node	
LTv4.1	Probability
LT4.1(1)	0,0667
LT4.1(2)	0,0667
LT4.1(3)	0,0667
LT4.1(4)	0,0667
LT4.1(5)	0,0667
LT4.1(6)	0,0667
LT4.1(7)	0,0667
LT4.1(8)	0,0667
LT4.1(9)	0,0667
LT4.1(10)	0,0667
LT4.1(11)	0,0667
LT4.1(12)	0,0667
LT4.1(13)	0,0667
LT4.1(14)	0,0667
LT4.1(15)	0,0667

Node	
LTv4.2	Probability
LT4.2(1)	0,0667
LT4.2(2)	0,0667
LT4.2(3)	0,0667
LT4.2(4)	0,0667
LT4.2(5)	0,0667
LT4.2(6)	0,0667
LT4.2(7)	0,0667
LT4.2(8)	0,0667
LT4.2(9)	0,0667
LT4.2(10)	0,0667
LT4.2(11)	0,0667
LT4.2(12)	0,0667
LT4.2(13)	0,0667
LT4.2(14)	0,0667
LT4.2(15)	0,0667

Node	
LTv4.3	Probability
LT4.3(1)	0,0667
LT4.3(2)	0,0667
LT4.3(3)	0,0667
LT4.3(4)	0,0667
LT4.3(5)	0,0667
LT4.3(6)	0,0667
LT4.3(7)	0,0667
LT4.3(8)	0,0667
LT4.3(9)	0,0667
LT4.3(10)	0,0667
LT4.3(11)	0,0667
LT4.3(12)	0,0667
LT4.3(13)	0,0667
LT4.3(14)	0,0667
LT4.3(15)	0,0667

Node	
LTv4.4	Probability
LT4.4(1)	0,0667
LT4.4(2)	0,0667
LT4.4(3)	0,0667
LT4.4(4)	0,0667
LT4.4(5)	0,0667
LT4.4(6)	0,0667
LT4.4(7)	0,0667
LT4.4(8)	0,0667
LT4.4(9)	0,0667
LT4.4(10)	0,0667
LT4.4(11)	0,0667
LT4.4(12)	0,0667
LT4.4(13)	0,0667
LT4.4(14)	0,0667
LT4.4(15)	0,0667

Node	
LTv4.5	Probability
LT4.5(1)	0,0667
LT4.5(2)	0,0667
LT4.5(3)	0,0667
LT4.5(4)	0,0667
LT4.5(5)	0,0667
LT4.5(6)	0,0667
LT4.5(7)	0,0667
LT4.5(8)	0,0667
LT4.5(9)	0,0667
LT4.5(10)	0,0667
LT4.5(11)	0,0667
LT4.5(12)	0,0667
LT4.5(13)	0,0667
LT4.5(14)	0,0667
LT4.5(15)	0,0667

Node Gp4.1	
Probability	
G4.1(1)	0,5
G4.1(2)	0,5

Node Bp4.1	
Probability	
B4.1(1)	0,02
B4.1(2)	0,09
B4.1(3)	0,33
B4.1(4)	0,33
B4.1(5)	0,23

LT4.1	V4.1(0)	V4.1(1)
LT4.1(1)	0,990	0,010
LT4.1(2)	0,981	0,019
LT4.1(3)	0,964	0,036
LT4.1(4)	0,933	0,068
LT4.1(5)	0,878	0,123
LT4.1(6)	0,788	0,212
LT4.1(7)	0,659	0,342
LT4.1(8)	0,500	0,500
LT4.1(9)	0,342	0,659
LT4.1(10)	0,212	0,788
LT4.1(11)	0,123	0,878
LT4.1(12)	0,068	0,933
LT4.1(13)	0,036	0,964
LT4.1(14)	0,019	0,981
LT4.1(15)	0,010	0,990

LT4.2	V4.2(0)	V4.2(1)
LT4.2(1)	0,987	0,013
LT4.2(2)	0,976	0,024
LT4.2(3)	0,955	0,045
LT4.2(4)	0,919	0,081
LT4.2(5)	0,857	0,143
LT4.2(6)	0,759	0,241
LT4.2(7)	0,625	0,375
LT4.2(8)	0,468	0,532
LT4.2(9)	0,317	0,683
LT4.2(10)	0,197	0,803
LT4.2(11)	0,115	0,885
LT4.2(12)	0,064	0,936
LT4.2(13)	0,035	0,965
LT4.2(14)	0,019	0,981
LT4.2(15)	0,010	0,990

LT4.3	V4.3(0)	V4.3(1)
LT4.3(1)	0,992	0,008
LT4.3(2)	0,985	0,015
LT4.3(3)	0,971	0,029
LT4.3(4)	0,945	0,055
LT4.3(5)	0,897	0,103
LT4.3(6)	0,816	0,184
LT4.3(7)	0,692	0,308
LT4.3(8)	0,534	0,466
LT4.3(9)	0,368	0,632
LT4.3(10)	0,229	0,771
LT4.3(11)	0,131	0,869
LT4.3(12)	0,071	0,929
LT4.3(13)	0,038	0,962
LT4.3(14)	0,020	0,981
LT4.3(15)	0,010	0,990

LT4.4	V4.4(0)	V4.4(1)
LT4.4(1)	0,977	0,023
LT4.4(2)	0,959	0,041
LT4.4(3)	0,929	0,071
LT4.4(4)	0,878	0,123
LT4.4(5)	0,798	0,202
LT4.4(6)	0,685	0,315
LT4.4(7)	0,545	0,455
LT4.4(8)	0,397	0,603
LT4.4(9)	0,266	0,734
LT4.4(10)	0,166	0,834
LT4.4(11)	0,099	0,901
LT4.4(12)	0,057	0,943
LT4.4(13)	0,032	0,968
LT4.4(14)	0,018	0,982
LT4.4(15)	0,010	0,990

LT4.5	V4.5(0)	V4.5(1)
LT4.5(1)	0,996	0,004
LT4.5(2)	0,993	0,008
LT4.5(3)	0,985	0,015
LT4.5(4)	0,969	0,031
LT4.5(5)	0,937	0,063
LT4.5(6)	0,878	0,123
LT4.5(7)	0,776	0,224
LT4.5(8)	0,625	0,375
LT4.5(9)	0,446	0,555
LT4.5(10)	0,279	0,721
LT4.5(11)	0,157	0,843
LT4.5(12)	0,083	0,917
LT4.5(13)	0,042	0,958
LT4.5(14)	0,021	0,980
LT4.5(15)	0,010	0,990

G4.1	B4.1	P4.1(0)	P4.1(1)
G4.1(1)	B4.1(1)	0,95	0,05
G4.1(1)	B4.1(2)	0,92	0,08
G4.1(1)	B4.1(3)	0,53	0,47
G4.1(1)	B4.1(4)	0,10	0,90
G4.1(1)	B4.1(5)	0,05	0,95
G4.1(2)	B4.1(1)	0,21	0,79
G4.1(2)	B4.1(2)	0,20	0,80
G4.1(2)	B4.1(3)	0,12	0,88
G4.1(2)	B4.1(4)	0,07	0,93
G4.1(2)	B4.1(5)	0,01	0,99

V4.1	V4.2	V4.3	V4.4	V4.5	P4.1	LS4(0)	LS4(1)
V4.1(0)	V4.2(0)	V4.3(0)	V4.4(0)	V4.5(0)	P4.1(0)	0,999	0,001
V4.1(0)	V4.2(0)	V4.3(0)	V4.4(0)	V4.5(0)	P4.1(1)	0,490	0,510
V4.1(0)	V4.2(0)	V4.3(0)	V4.4(0)	V4.5(1)	P4.1(0)	0,470	0,530
V4.1(0)	V4.2(0)	V4.3(0)	V4.4(0)	V4.5(1)	P4.1(1)	0,430	0,570
V4.1(0)	V4.2(0)	V4.3(0)	V4.4(1)	V4.5(0)	P4.1(0)	0,480	0,520
V4.1(0)	V4.2(0)	V4.3(0)	V4.4(1)	V4.5(0)	P4.1(1)	0,460	0,540
V4.1(0)	V4.2(0)	V4.3(0)	V4.4(1)	V4.5(1)	P4.1(0)	0,420	0,580
V4.1(0)	V4.2(0)	V4.3(0)	V4.4(1)	V4.5(1)	P4.1(1)	0,320	0,680
V4.1(0)	V4.2(0)	V4.3(1)	V4.4(0)	V4.5(0)	P4.1(0)	0,490	0,510
V4.1(0)	V4.2(0)	V4.3(1)	V4.4(0)	V4.5(0)	P4.1(1)	0,410	0,590
V4.1(0)	V4.2(0)	V4.3(1)	V4.4(0)	V4.5(1)	P4.1(0)	0,420	0,580
V4.1(0)	V4.2(0)	V4.3(1)	V4.4(0)	V4.5(1)	P4.1(1)	0,330	0,670
V4.1(0)	V4.2(0)	V4.3(1)	V4.4(1)	V4.5(0)	P4.1(0)	0,410	0,590
V4.1(0)	V4.2(0)	V4.3(1)	V4.4(1)	V4.5(0)	P4.1(1)	0,340	0,660
V4.1(0)	V4.2(0)	V4.3(1)	V4.4(1)	V4.5(1)	P4.1(0)	0,350	0,650
V4.1(0)	V4.2(0)	V4.3(1)	V4.4(1)	V4.5(1)	P4.1(1)	0,180	0,820
V4.1(0)	V4.2(1)	V4.3(0)	V4.4(0)	V4.5(0)	P4.1(0)	0,490	0,510
V4.1(0)	V4.2(1)	V4.3(0)	V4.4(0)	V4.5(0)	P4.1(1)	0,430	0,570
V4.1(0)	V4.2(1)	V4.3(0)	V4.4(0)	V4.5(1)	P4.1(0)	0,420	0,580
V4.1(0)	V4.2(1)	V4.3(0)	V4.4(0)	V4.5(1)	P4.1(1)	0,330	0,670
V4.1(0)	V4.2(1)	V4.3(0)	V4.4(1)	V4.5(0)	P4.1(0)	0,410	0,590
V4.1(0)	V4.2(1)	V4.3(0)	V4.4(1)	V4.5(0)	P4.1(1)	0,320	0,680
V4.1(0)	V4.2(1)	V4.3(0)	V4.4(1)	V4.5(1)	P4.1(0)	0,310	0,690
V4.1(0)	V4.2(1)	V4.3(0)	V4.4(1)	V4.5(1)	P4.1(1)	0,220	0,780
V4.1(0)	V4.2(1)	V4.3(1)	V4.4(0)	V4.5(0)	P4.1(0)	0,450	0,550
V4.1(0)	V4.2(1)	V4.3(1)	V4.4(0)	V4.5(0)	P4.1(1)	0,300	0,700
V4.1(0)	V4.2(1)	V4.3(1)	V4.4(0)	V4.5(1)	P4.1(0)	0,320	0,680
V4.1(0)	V4.2(1)	V4.3(1)	V4.4(0)	V4.5(1)	P4.1(1)	0,210	0,790
V4.1(0)	V4.2(1)	V4.3(1)	V4.4(1)	V4.5(0)	P4.1(0)	0,290	0,710
V4.1(0)	V4.2(1)	V4.3(1)	V4.4(1)	V4.5(0)	P4.1(1)	0,130	0,870
V4.1(0)	V4.2(1)	V4.3(1)	V4.4(1)	V4.5(1)	P4.1(0)	0,120	0,880
V4.1(0)	V4.2(1)	V4.3(1)	V4.4(1)	V4.5(1)	P4.1(1)	0,070	0,930
V4.1(1)	V4.2(0)	V4.3(0)	V4.4(0)	V4.5(0)	P4.1(0)	0,490	0,510
V4.1(1)	V4.2(0)	V4.3(0)	V4.4(0)	V4.5(0)	P4.1(1)	0,390	0,610
V4.1(1)	V4.2(0)	V4.3(0)	V4.4(0)	V4.5(1)	P4.1(0)	0,380	0,620
V4.1(1)	V4.2(0)	V4.3(0)	V4.4(0)	V4.5(1)	P4.1(1)	0,310	0,690
V4.1(1)	V4.2(0)	V4.3(0)	V4.4(1)	V4.5(0)	P4.1(0)	0,460	0,540
V4.1(1)	V4.2(0)	V4.3(0)	V4.4(1)	V4.5(0)	P4.1(1)	0,400	0,600
V4.1(1)	V4.2(0)	V4.3(0)	V4.4(1)	V4.5(1)	P4.1(0)	0,300	0,700
V4.1(1)	V4.2(0)	V4.3(0)	V4.4(1)	V4.5(1)	P4.1(1)	0,200	0,800
V4.1(1)	V4.2(0)	V4.3(1)	V4.4(0)	V4.5(0)	P4.1(0)	0,310	0,690
V4.1(1)	V4.2(0)	V4.3(1)	V4.4(0)	V4.5(0)	P4.1(1)	0,210	0,790
V4.1(1)	V4.2(0)	V4.3(1)	V4.4(0)	V4.5(1)	P4.1(0)	0,220	0,780
V4.1(1)	V4.2(0)	V4.3(1)	V4.4(0)	V4.5(1)	P4.1(1)	0,170	0,830
V4.1(1)	V4.2(0)	V4.3(1)	V4.4(1)	V4.5(0)	P4.1(0)	0,230	0,770
V4.1(1)	V4.2(0)	V4.3(1)	V4.4(1)	V4.5(0)	P4.1(1)	0,170	0,830
V4.1(1)	V4.2(0)	V4.3(1)	V4.4(1)	V4.5(1)	P4.1(0)	0,190	0,810
V4.1(1)	V4.2(0)	V4.3(1)	V4.4(1)	V4.5(1)	P4.1(1)	0,100	0,900
V4.1(1)	V4.2(1)	V4.3(0)	V4.4(0)	V4.5(0)	P4.1(0)	0,390	0,610
V4.1(1)	V4.2(1)	V4.3(0)	V4.4(0)	V4.5(0)	P4.1(1)	0,280	0,720

V4.1(1)	V4.2(1)	V4.3(0)	V4.4(0)	V4.5(1)	P4.1(0)	0,270	0,730
V4.1(1)	V4.2(1)	V4.3(0)	V4.4(0)	V4.5(1)	P4.1(1)	0,180	0,820
V4.1(1)	V4.2(1)	V4.3(0)	V4.4(1)	V4.5(0)	P4.1(0)	0,260	0,740
V4.1(1)	V4.2(1)	V4.3(0)	V4.4(1)	V4.5(0)	P4.1(1)	0,170	0,830
V4.1(1)	V4.2(1)	V4.3(0)	V4.4(1)	V4.5(1)	P4.1(0)	0,160	0,840
V4.1(1)	V4.2(1)	V4.3(0)	V4.4(1)	V4.5(1)	P4.1(1)	0,090	0,910
V4.1(1)	V4.2(1)	V4.3(1)	V4.4(0)	V4.5(0)	P4.1(0)	0,250	0,750
V4.1(1)	V4.2(1)	V4.3(1)	V4.4(0)	V4.5(0)	P4.1(1)	0,130	0,870
V4.1(1)	V4.2(1)	V4.3(1)	V4.4(0)	V4.5(1)	P4.1(0)	0,140	0,860
V4.1(1)	V4.2(1)	V4.3(1)	V4.4(0)	V4.5(1)	P4.1(1)	0,060	0,940
V4.1(1)	V4.2(1)	V4.3(1)	V4.4(1)	V4.5(0)	P4.1(0)	0,150	0,850
V4.1(1)	V4.2(1)	V4.3(1)	V4.4(1)	V4.5(0)	P4.1(1)	0,070	0,930
V4.1(1)	V4.2(1)	V4.3(1)	V4.4(1)	V4.5(1)	P4.1(0)	0,080	0,920
V4.1(1)	V4.2(1)	V4.3(1)	V4.4(1)	V4.5(1)	P4.1(1)	0,010	0,990

Node LTV5.2	Probability
LT5.2(1)	0,0667
LT5.2(2)	0,0667
LT5.2(3)	0,0667
LT5.2(4)	0,0667
LT5.2(5)	0,0667
LT5.2(6)	0,0667
LT5.2(7)	0,0667
LT5.2(8)	0,0667
LT5.2(9)	0,0667
LT5.2(10)	0,0667
LT5.2(11)	0,0667
LT5.2(12)	0,0667
LT5.2(13)	0,0667
LT5.2(14)	0,0667
LT5.2(15)	0,0667

Node LTV5.3	Probability
LT5.3(1)	0,0667
LT5.3(2)	0,0667
LT5.3(3)	0,0667
LT5.3(4)	0,0667
LT5.3(5)	0,0667
LT5.3(6)	0,0667
LT5.3(7)	0,0667
LT5.3(8)	0,0667
LT5.3(9)	0,0667
LT5.3(10)	0,0667
LT5.3(11)	0,0667
LT5.3(12)	0,0667
LT5.3(13)	0,0667
LT5.3(14)	0,0667
LT5.3(15)	0,0667

Node LTV5.4	Probability
LT5.4(1)	0,0667
LT5.4(2)	0,0667
LT5.4(3)	0,0667
LT5.4(4)	0,0667
LT5.4(5)	0,0667
LT5.4(6)	0,0667
LT5.4(7)	0,0667
LT5.4(8)	0,0667
LT5.4(9)	0,0667
LT5.4(10)	0,0667
LT5.4(11)	0,0667
LT5.4(12)	0,0667
LT5.4(13)	0,0667
LT5.4(14)	0,0667
LT5.4(15)	0,0667

Node Gp5.1	Probability
G5.1(1)	0,5
G5.1(2)	0,5

Node Bp5.1	Probability
B5.1(1)	0,01
B5.1(2)	0,05
B5.1(3)	0,36
B5.1(4)	0,32
B5.1(5)	0,26

G5.1	B5.1	P5.1(0)	P5.1(1)
G5.1(1)	B5.1(1)	0,95	0,05
G5.1(1)	B5.1(2)	0,92	0,08
G5.1(1)	B5.1(3)	0,53	0,47
G5.1(1)	B5.1(4)	0,10	0,90
G5.1(1)	B5.1(5)	0,05	0,95
G5.1(2)	B5.1(1)	0,27	0,73
G5.1(2)	B5.1(2)	0,20	0,80
G5.1(2)	B5.1(3)	0,11	0,89
G5.1(2)	B5.1(4)	0,08	0,92
G5.1(2)	B5.1(5)	0,01	0,99

Node LTV5.1	Probability	LT5.1	V5.1(0)	V5.1(1)	LT5.2	V5.2(0)	V5.2(1)
LT5.1(1)	0,02	LT5.1(1)	0,9929	0,0071	LT5.2(1)	0,982	0,018
LT5.1(2)	0,02	LT5.1(2)	0,9914	0,0086	LT5.2(2)	0,967	0,033
LT5.1(3)	0,02	LT5.1(3)	0,9896	0,0104	LT5.2(3)	0,940	0,060
LT5.1(4)	0,02	LT5.1(4)	0,9874	0,0126	LT5.2(4)	0,895	0,105
LT5.1(5)	0,02	LT5.1(5)	0,9847	0,0153	LT5.2(5)	0,822	0,178
LT5.1(6)	0,02	LT5.1(6)	0,9815	0,0185	LT5.2(6)	0,715	0,285
LT5.1(7)	0,02	LT5.1(7)	0,9776	0,0224	LT5.2(7)	0,576	0,424
LT5.1(8)	0,02	LT5.1(8)	0,9729	0,0271	LT5.2(8)	0,424	0,576
LT5.1(9)	0,02	LT5.1(9)	0,9673	0,0327	LT5.2(9)	0,285	0,715
LT5.1(10)	0,02	LT5.1(10)	0,9606	0,0394	LT5.2(10)	0,178	0,822
LT5.1(11)	0,02	LT5.1(11)	0,9525	0,0475	LT5.2(11)	0,105	0,895
LT5.1(12)	0,02	LT5.1(12)	0,9429	0,0571	LT5.2(12)	0,060	0,940
LT5.1(13)	0,02	LT5.1(13)	0,9314	0,0686	LT5.2(13)	0,033	0,967
LT5.1(14)	0,02	LT5.1(14)	0,9179	0,0821	LT5.2(14)	0,018	0,982
LT5.1(15)	0,02	LT5.1(15)	0,902	0,098	LT5.2(15)	0,010	0,990
LT5.1(16)	0,02	LT5.1(16)	0,8834	0,1166			
LT5.1(17)	0,02	LT5.1(17)	0,8618	0,1382	LT5.3	V5.3(0)	V5.3(1)
LT5.1(18)	0,02	LT5.1(18)	0,8369	0,1631	LT5.3(1)	0,992	0,008
LT5.1(19)	0,02	LT5.1(19)	0,8086	0,1914	LT5.3(2)	0,985	0,015
LT5.1(20)	0,02	LT5.1(20)	0,7766	0,2234	LT5.3(3)	0,971	0,029
LT5.1(21)	0,02	LT5.1(21)	0,741	0,259	LT5.3(4)	0,945	0,055
LT5.1(22)	0,02	LT5.1(22)	0,702	0,298	LT5.3(5)	0,897	0,103
LT5.1(23)	0,02	LT5.1(23)	0,6597	0,3403	LT5.3(6)	0,816	0,184
LT5.1(24)	0,02	LT5.1(24)	0,6147	0,3853	LT5.3(7)	0,692	0,308
LT5.1(25)	0,02	LT5.1(25)	0,5677	0,4323	LT5.3(8)	0,534	0,466
LT5.1(26)	0,02	LT5.1(26)	0,5195	0,4805	LT5.3(9)	0,368	0,632
LT5.1(27)	0,02	LT5.1(27)	0,4708	0,5292	LT5.3(10)	0,229	0,771
LT5.1(28)	0,02	LT5.1(28)	0,4227	0,5773	LT5.3(11)	0,131	0,869
LT5.1(29)	0,02	LT5.1(29)	0,3761	0,6239	LT5.3(12)	0,071	0,929
LT5.1(30)	0,02	LT5.1(30)	0,3316	0,6684	LT5.3(13)	0,038	0,962
LT5.1(31)	0,02	LT5.1(31)	0,2899	0,7101	LT5.3(14)	0,020	0,981
LT5.1(32)	0,02	LT5.1(32)	0,2515	0,7485	LT5.3(15)	0,010	0,990
LT5.1(33)	0,02	LT5.1(33)	0,2167	0,7833			
LT5.1(34)	0,02	LT5.1(34)	0,1855	0,8145	LT5.4	V5.4(0)	V5.4(1)
LT5.1(35)	0,02	LT5.1(35)	0,1578	0,8422	LT5.4(1)	0,972	0,028
LT5.1(36)	0,02	LT5.1(36)	0,1336	0,8664	LT5.4(2)	0,951	0,049
LT5.1(37)	0,02	LT5.1(37)	0,1127	0,8873	LT5.4(3)	0,916	0,084
LT5.1(38)	0,02	LT5.1(38)	0,0946	0,9054	LT5.4(4)	0,859	0,142
LT5.1(39)	0,02	LT5.1(39)	0,0792	0,9208	LT5.4(5)	0,772	0,228
LT5.1(40)	0,02	LT5.1(40)	0,0661	0,9339	LT5.4(6)	0,655	0,345
LT5.1(41)	0,02	LT5.1(41)	0,0551	0,9449	LT5.4(7)	0,515	0,486
LT5.1(42)	0,02	LT5.1(42)	0,0458	0,9542	LT5.4(8)	0,372	0,628
LT5.1(43)	0,02	LT5.1(43)	0,038	0,962	LT5.4(9)	0,249	0,751
LT5.1(44)	0,02	LT5.1(44)	0,0315	0,9685	LT5.4(10)	0,156	0,844
LT5.1(45)	0,02	LT5.1(45)	0,026	0,974	LT5.4(11)	0,094	0,906
LT5.1(46)	0,02	LT5.1(46)	0,0215	0,9785	LT5.4(12)	0,055	0,945
LT5.1(47)	0,02	LT5.1(47)	0,0178	0,9822	LT5.4(13)	0,031	0,969
LT5.1(48)	0,02	LT5.1(48)	0,0147	0,9853	LT5.4(14)	0,018	0,982
LT5.1(49)	0,02	LT5.1(49)	0,0121	0,9879	LT5.4(15)	0,010	0,990
LT5.1(50)	0,02	LT5.1(50)	0,01	0,99			

V5.1	V5.2	V5.3	V5.4	P5.1	LS5(0)	LS5(1)
V5.1(0)	V5.2(0)	V5.3(0)	V5.4(0)	P5.1(0)	0,99	0,01
V5.1(0)	V5.2(0)	V5.3(0)	V5.4(0)	P5.1(1)	0,40	0,60
V5.1(0)	V5.2(0)	V5.3(0)	V5.4(1)	P5.1(0)	0,46	0,54
V5.1(0)	V5.2(0)	V5.3(0)	V5.4(1)	P5.1(1)	0,31	0,69
V5.1(0)	V5.2(0)	V5.3(1)	V5.4(0)	P5.1(0)	0,47	0,53
V5.1(0)	V5.2(0)	V5.3(1)	V5.4(0)	P5.1(1)	0,32	0,68
V5.1(0)	V5.2(0)	V5.3(1)	V5.4(1)	P5.1(0)	0,35	0,65
V5.1(0)	V5.2(0)	V5.3(1)	V5.4(1)	P5.1(1)	0,34	0,66
V5.1(0)	V5.2(1)	V5.3(0)	V5.4(0)	P5.1(0)	0,48	0,52
V5.1(0)	V5.2(1)	V5.3(0)	V5.4(0)	P5.1(1)	0,31	0,69
V5.1(0)	V5.2(1)	V5.3(0)	V5.4(1)	P5.1(0)	0,34	0,66
V5.1(0)	V5.2(1)	V5.3(0)	V5.4(1)	P5.1(1)	0,24	0,76
V5.1(0)	V5.2(1)	V5.3(1)	V5.4(0)	P5.1(0)	0,40	0,60
V5.1(0)	V5.2(1)	V5.3(1)	V5.4(0)	P5.1(1)	0,28	0,72
V5.1(0)	V5.2(1)	V5.3(1)	V5.4(1)	P5.1(0)	0,27	0,73
V5.1(0)	V5.2(1)	V5.3(1)	V5.4(1)	P5.1(1)	0,04	0,96
V5.1(1)	V5.2(0)	V5.3(0)	V5.4(0)	P5.1(0)	0,45	0,55
V5.1(1)	V5.2(0)	V5.3(0)	V5.4(0)	P5.1(1)	0,37	0,63
V5.1(1)	V5.2(0)	V5.3(0)	V5.4(1)	P5.1(0)	0,38	0,62
V5.1(1)	V5.2(0)	V5.3(0)	V5.4(1)	P5.1(1)	0,26	0,74
V5.1(1)	V5.2(0)	V5.3(1)	V5.4(0)	P5.1(0)	0,31	0,69
V5.1(1)	V5.2(0)	V5.3(1)	V5.4(0)	P5.1(1)	0,17	0,83
V5.1(1)	V5.2(0)	V5.3(1)	V5.4(1)	P5.1(0)	0,08	0,92
V5.1(1)	V5.2(0)	V5.3(1)	V5.4(1)	P5.1(1)	0,07	0,93
V5.1(1)	V5.2(1)	V5.3(0)	V5.4(0)	P5.1(0)	0,33	0,67
V5.1(1)	V5.2(1)	V5.3(0)	V5.4(0)	P5.1(1)	0,16	0,84
V5.1(1)	V5.2(1)	V5.3(0)	V5.4(1)	P5.1(0)	0,14	0,86
V5.1(1)	V5.2(1)	V5.3(0)	V5.4(1)	P5.1(1)	0,09	0,91
V5.1(1)	V5.2(1)	V5.3(1)	V5.4(0)	P5.1(0)	0,15	0,85
V5.1(1)	V5.2(1)	V5.3(1)	V5.4(0)	P5.1(1)	0,08	0,92
V5.1(1)	V5.2(1)	V5.3(1)	V5.4(1)	P5.1(0)	0,10	0,90
V5.1(1)	V5.2(1)	V5.3(1)	V5.4(1)	P5.1(1)	0,01	0,99

LS1	LS2	LS3	LS4	LS5	PL(0)	PL(1)
LS1(0)	LS2(0)	LS3(0)	LS4(0)	LS5(0)	0,90	0,10
LS1(0)	LS2(0)	LS3(0)	LS4(0)	LS5(1)	0,30	0,70
LS1(0)	LS2(0)	LS3(0)	LS4(1)	LS5(0)	0,30	0,70
LS1(0)	LS2(0)	LS3(0)	LS4(1)	LS5(1)	0,20	0,80
LS1(0)	LS2(0)	LS3(1)	LS4(0)	LS5(0)	0,25	0,75
LS1(0)	LS2(0)	LS3(1)	LS4(0)	LS5(1)	0,19	0,81
LS1(0)	LS2(0)	LS3(1)	LS4(1)	LS5(0)	0,17	0,83
LS1(0)	LS2(0)	LS3(1)	LS4(1)	LS5(1)	0,11	0,89
LS1(0)	LS2(1)	LS3(0)	LS4(0)	LS5(0)	0,26	0,74
LS1(0)	LS2(1)	LS3(0)	LS4(0)	LS5(1)	0,20	0,80
LS1(0)	LS2(1)	LS3(0)	LS4(1)	LS5(0)	0,23	0,77
LS1(0)	LS2(1)	LS3(0)	LS4(1)	LS5(1)	0,16	0,84
LS1(0)	LS2(1)	LS3(1)	LS4(0)	LS5(0)	0,26	0,74
LS1(0)	LS2(1)	LS3(1)	LS4(0)	LS5(1)	0,20	0,80
LS1(0)	LS2(1)	LS3(1)	LS4(1)	LS5(0)	0,18	0,82
LS1(0)	LS2(1)	LS3(1)	LS4(1)	LS5(1)	0,05	0,95
LS1(1)	LS2(0)	LS3(0)	LS4(0)	LS5(0)	0,22	0,78
LS1(1)	LS2(0)	LS3(0)	LS4(0)	LS5(1)	0,11	0,89
LS1(1)	LS2(0)	LS3(0)	LS4(1)	LS5(0)	0,15	0,85
LS1(1)	LS2(0)	LS3(0)	LS4(1)	LS5(1)	0,14	0,86
LS1(1)	LS2(0)	LS3(1)	LS4(0)	LS5(0)	0,27	0,73
LS1(1)	LS2(0)	LS3(1)	LS4(0)	LS5(1)	0,17	0,83
LS1(1)	LS2(0)	LS3(1)	LS4(1)	LS5(0)	0,16	0,84
LS1(1)	LS2(0)	LS3(1)	LS4(1)	LS5(1)	0,06	0,94
LS1(1)	LS2(1)	LS3(0)	LS4(0)	LS5(0)	0,28	0,72
LS1(1)	LS2(1)	LS3(0)	LS4(0)	LS5(1)	0,19	0,81
LS1(1)	LS2(1)	LS3(0)	LS4(1)	LS5(0)	0,20	0,80
LS1(1)	LS2(1)	LS3(0)	LS4(1)	LS5(1)	0,05	0,95
LS1(1)	LS2(1)	LS3(1)	LS4(0)	LS5(0)	0,15	0,85
LS1(1)	LS2(1)	LS3(1)	LS4(0)	LS5(1)	0,06	0,94
LS1(1)	LS2(1)	LS3(1)	LS4(1)	LS5(0)	0,05	0,95
LS1(1)	LS2(1)	LS3(1)	LS4(1)	LS5(1)	0,01	0,99

Appendix H: Independence checks

Line segment 1

(LS1 ⊥ G2.1, LT1.2, P5.1, LT3.7, B2.2, LT3.1, LT5.3, LT5.4, V3.6, LT3.2, V5.2, V4.5, LT3.3, B3.1, B4.1, V3.2, G1.1, V3.8, V4.3, LT5.1, V2.1, LT3.6, LS4, G3.1, V3.5, LT4.4, LS3, V5.4, B2.1, LT4.5, V3.7, V2.2, V3.3, LS2, LT2.2, V4.4, P2.2, LT4.3, G4.1, LT2.1, P2.1, LT4.2, B5.1, G2.2, P4.1, P3.1, V4.1, LT3.8, LT3.4, G5.1, LT3.5, B1.1, LT5.2, V3.1, LT4.1, V3.4, V5.1, LT1.1, V4.2, LS5, V5.3 | **P1.1, V1.1, V1.2)**

Line segment 2

(LS2 ⊥ G2.1, LT1.2, P1.1, P5.1, LT3.7, B2.2, LT3.1, LT5.3, LT5.4, V3.6, LT3.2, V5.2, V4.5, LT3.3, B3.1, B4.1, V3.2, G1.1, V3.8, V4.3, LT5.1, LS4, LT3.6, G3.1, V3.5, LT4.4, LS3, V5.4, B2.1, LT4.5, V3.7, V3.3, LT2.2, V4.4, LT4.3, G4.1, LT2.1, LT4.2, B5.1, G2.2, P4.1, P3.1, V4.1, LT3.8, LT3.4, LS1, G5.1, LT3.5, B1.1, LT5.2, V3.1, LT4.1, V3.4, V5.1, LT1.1, V4.2, V1.1, LS5, V5.3, V1.2 | **V2.1, P2.1, V2.2, P2.2)**

Line segment 3

(LS3 ⊥ G2.1, LT1.2, P1.1, P5.1, LT3.7, B2.2, LT3.1, LT5.3, LT5.4, LT3.2, V5.2, V4.5, LT3.3, B3.1, B4.1, G1.1, V4.3, LT5.1, V2.1, LS4, LT3.6, G3.1, LT4.4, V5.4, B2.1, LT4.5, V2.2, LS2, LT2.2, V4.4, P2.2, LT4.3, G4.1, LT2.1, P2.1, LT4.2, B5.1, G2.2, P4.1, V4.1, LT3.8, LT3.4, LS1, G5.1, LT3.5, B1.1, LT5.2, LT4.1, V5.1, LT1.1, V4.2, V1.1, LS5, V5.3, V1.2 | **V3.1, V3.2, V3.3, V3.4, V3.5, V3.6, V3.7, V3.8, P3.1)**

Line segment 4

(LS4 ⊥ G2.1, LT1.2, P1.1, P5.1, LT3.7, B2.2, LT3.1, LT5.3, LT5.4, V3.6, LT3.2, V5.2, LT3.3, B3.1, B4.1, V3.2, G1.1, V3.8, LT5.1, V2.1, G3.1, LT3.6, V3.5, LT4.4, LS3, V5.4, B2.1, LT4.5, V3.7, V2.2, V3.3, LS2, LT2.2, P2.2, LT4.3, G4.1, LT2.1, P2.1, LT4.2, B5.1, G2.2, P3.1, LT3.8, LT3.4, LS1, G5.1, LT3.5, B1.1, LT5.2, V3.1, LT4.1, V3.4, V5.1, LT1.1, V1.1, LS5, V5.3, V1.2 | **V4.1, V4.2, V4.3, V4.4, V4.5, P4.1)**

Line segment 5

(LS5 ⊥ G2.1, LT1.2, P1.1, LT3.7, B2.2, LT3.1, LT5.3, LT5.4, V3.6, LT3.2, V4.5, LT3.3, B3.1, B4.1, V3.2, G1.1, V3.8, V4.3, LT5.1, V2.1, LT3.6, LS4, G3.1, V3.5, LT4.4, LS3, B2.1, LT4.5, V3.7, V2.2, V3.3, LS2, LT2.2, V4.4, P2.2, LT4.3, G4.1, LT2.1, P2.1, LT4.2, B5.1, G2.2, P4.1, P3.1, V4.1, LT3.8, LT3.4, LS1, G5.1, LT3.5, B1.1, LT5.2, V3.1, V3.4, LT1.1, V4.2, V1.1, LT4.1, V1.2 | **P5.1, V5.1, V5.2, V5.3, V5.4)**

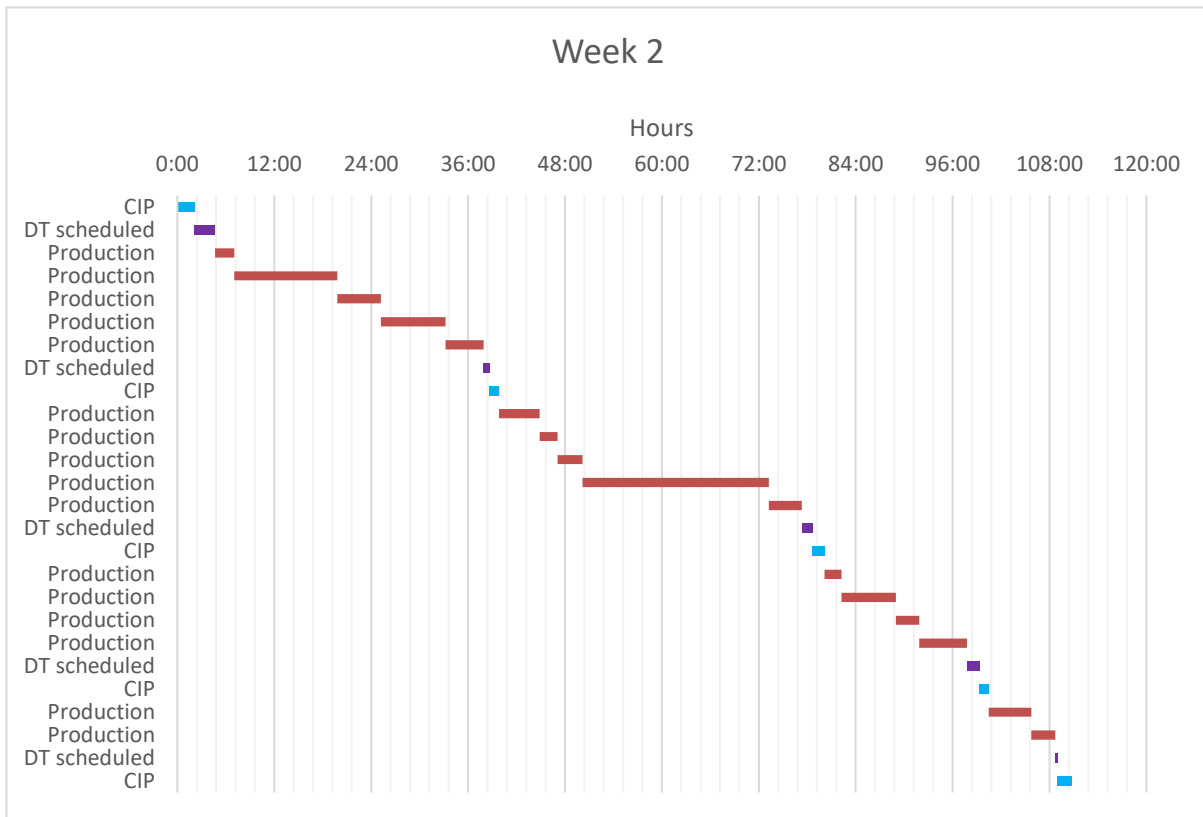
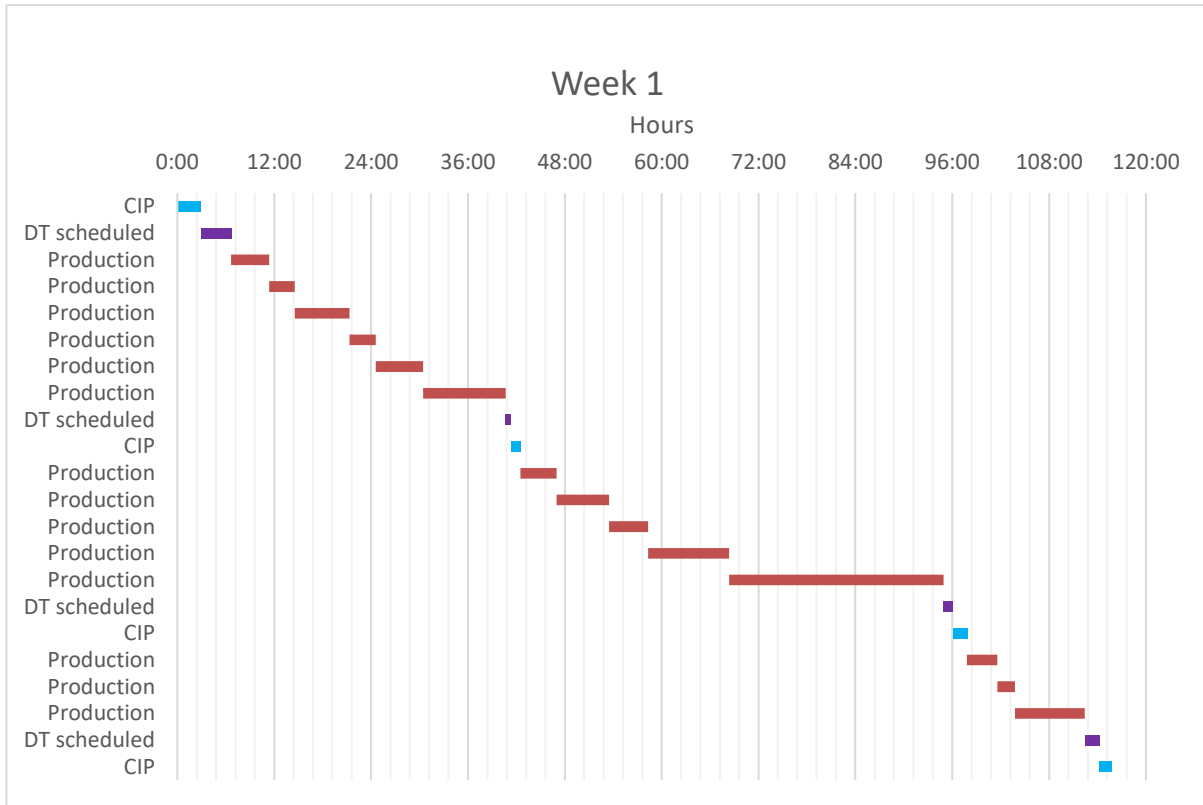
Appendix I: Mean Time Between Failure

Start values of MTBF and verification of number of failures

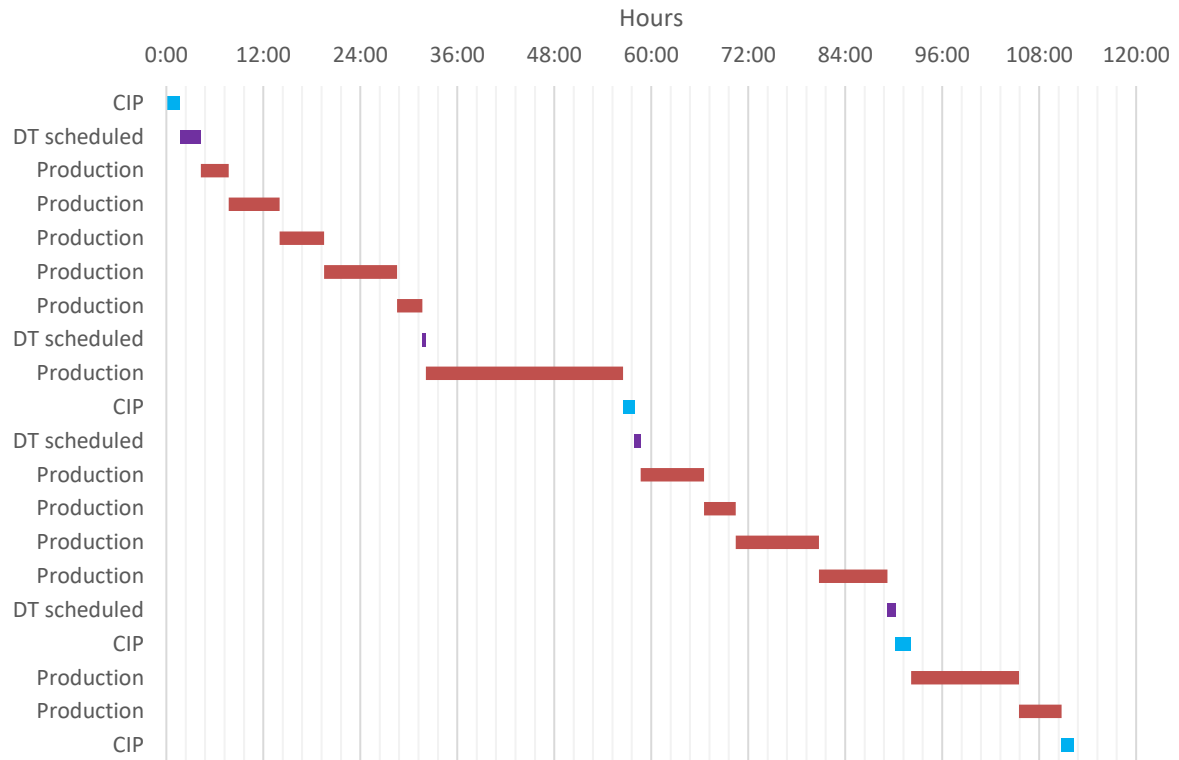
Component	MTBF [hr]	Start time MTBF code	Number of failures in synthetic dataset
V1.1	3323:41:07	1723:00:00	2
V1.2	586:31:58	380:00:00	8
P1.1	4985:31:40	2485:00:00	1
V2.1	4985:31:40	3285:00:00	1
V2.2	4985:31:40	1885:00:00	1
P2.1	4985:31:40	185:00:00	1
P2.2	4985:31:40	3085:00:00	1
V3.1	4985:31:40	2185:00:00	1
V3.2	4985:31:40	2585:00:00	1
V3.3	3323:41:07	3223:00:00	2
V3.4	4985:31:40	2485:00:00	1
V3.5	3323:41:07	2121:00:00	2
V3.6	1107:53:42	107:00:00	4
V3.7	4985:31:40	3485:00:00	1
V3.8	767:00:15	167:00:00	6
P3.1	4985:31:40	1185:00:00	1
V4.1	4985:31:40	3885:00:00	1
V4.2	4985:31:40	1185:00:00	1
V4.3	4985:31:40	2785:00:00	1
V4.4	4985:31:40	585:00:00	1
V4.5	4985:31:40	4585:00:00	1
P4.1	4985:31:40	1185:00:00	1
V5.1	4985:31:40	285:00:00	1
V5.2	4985:31:40	2285:00:00	1
V5.3	4985:31:40	4685:00:00	1
V5.4	3323:41:07	2623:00:00	2
P5.1	4985:31:40	4385:00:00	1

Appendix J: Synthetic production planning

Synthetic production planning example



Week 3



Week 4

