BERGISCHE
UNIVERSITÄT
WUPPERTAL

Fakultät 7- Maschinenbau und Sicherheitstechnik
Fachgebiet Verkehrssicherheit und Zuverlässigkeit
Jun.-Prof. Dr. Antoine Tordeux

Masterthesis

Instandhaltungsstrategien in der Industrie 4.0 - Vergleich von zustandsorientierten und herkömmlichen Instandhaltungsstrategien mit beispielhafter Datensatzbearbeitung

---

Zur Erreichung des akademischen Grades
Master of Science

Name:	Nico Lischka
Matrikel-Nr.:	1861210
Studiengang:	Master Qualitätsingenieurswesen
Betreuer:	Jun-. Prof. Dr. Antoine Tordeux
Datum der Ausgabe:	10.05.2021
Datum der Abgabe:	08.10.2021

Eidesstattliche Erklärung

Name, Vorname: Lischka, Nico

Erklärung

Hiermit, erkläre ich, dass ich die von mir eingereichte Abschlussarbeit (Master-Thesis) selbstständig verfasst und keine andere als die angegebene Quelle und Hilfsmittel benutzt sowie Stellen der Abschlussarbeit, die anderen Werken dem Wortlaut oder Sinn nach entnommen wurden, in jedem Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Ich bin damit einverstanden, dass die Arbeit durch Dritte eingesehen und unter Wahrung urheberrechtlicher Grundsätze zitiert werden darf.

Datum und Ort: Ahlen, den 8.10.2021

Unterschrift: N. Lischka

Zusammenfassung

In der Masterthesis wurde die zustandsorientierte Instandhaltungsstrategie mit den herkömmlichen Strategien verglichen. Für jede Strategie oder auch Strategiekombination gibt es einen geeigneten Einsatzbereich. Auf Grund dessen sollte eine zustandsorientierte Instandhaltungsstrategie angewandt werden, wenn die technische Ausstattung gegeben ist, ein Fehler zwingend zu vermeiden ist, die Prozesse beherrscht werden, die Daten analysiert werden können und die Planung der Fertigung und Instandhaltung möglich ist.

Bei der darauffolgenden Analyse mit R wurde ein Datensatze zu Triebwerken der NASA, welche in Kalifornien im Jahre 2008 ermittelt wurden, verwendet. Hierbei wurden unterschiedliche Algorithmen und Regressionen für die Prognose des RULs angewendet. Es ließ sich feststellen, dass die Algorithmen, die auf maschinellem Lernen basieren, wie z.B. die Support-Vector-Machine und das neuronale Netz im Vergleich zu den Regressionen besser abgeschnitten haben. Die Ergebnisse liefern Werte, die in Ordnung sind, aber noch nicht überragend. Aus diesem Grund müssten im Nachgang an diese Thesis einige Faktoren verändert werden, wie zum Beispiel die Datenerfassung, um bessere Prognoseergebnisse zu erhalten.

Abstract

In this thesis, the condition-based maintenance strategy was compared with the conventional strategies. There is suitable area of application for each strategy or strategy combination. Based on this, a condition-based maintenance strategy should be applied of the technical equipment is given, a defect must be avoided, the processes are controlled, the data can be analyzed and the planning of production and maintenance is possible.

In the subsequent analysis with R, a data set of NASA engines determined in California in 2008 was used. Different algorithms and regressions were used to predict the RUL. It was found that the algorithms based on machine learning, such as the support vector machine and the neural network, performed better compared to the regressions. The results provide values are ok, but still not outstanding. For this reason, some factors would have to be changed in the follow-up of this thesis, such as data collection, to get better prediction results.

Abkürzungsverzeichnis

AHC	Agglomerative hierarchische Clustermethode
AI	Artifizielle Intelligenz
CPS	Cyber-Physische-Systeme
DIN	Deutsche Industrie Norm
DoE	Design of Experiments
FIT	Failure in Time
FMEA	Fehlermöglichkeits- und -einflussanalyse
HFD	Human-Focussed-Design
HLEG	High Level Expert Group
HMM	Hidden-Markov-Modell
KMA	K-means-Algorithmus
KNN	K-nearest-neighbors
LR	Lineare Regression
MDBF	Mean Distance between Failure
MSE	Mean square error
MTBF	Mean Time between Failure
MTTF	Mean Time to Failure
$MTTF_d$	Mean Time to dangerous Failure
NLR	Nicht lineare Regression
NN	Neuronales Netz
PC	Principal Component
QFD	Quality-Function-Deployment
RF	Random Forest
RMSE	Root mean square error
RPZ	Risikoprioritätszahl
RUL	Remaining useful life
SVM	Support-Vector-Machine
UL	Useful life
VDI	Verein Deutscher Ingenieure

Tabellenverzeichnis

Tabelle 1: DMADV-Phasen, Ziele, Methoden 9
Tabelle 2: Messungen am Triebwerk 42
Tabelle 3: Daten zu den Hauptkomponenten Teil 1 46
Tabelle 4: Daten zu den Hauptkomponenten Teil 2 47
Tabelle 5: Kennzahlen der 10 Durchläufe; Algorithmen 1-4 64
Tabelle 6: Kennzahlen der 10 Durchläufe; Algorithmen 5-8 64

Abbildungsverzeichnis

Abbildung 1: Vier Stufen der Industriellen Revolution 4
Abbildung 2: Mensch-Maschine-Betriebsmittel Vernetzung 6
Abbildung 3: Ausfallverhalten Badewannenkurve 8
Abbildung 4: Normalverteilung, Anpassung der Toleranz / erlaubten Abweichung 12
Abbildung 5: Differenzierung von Big Data 14
Abbildung 6: Struktur von Lernmethoden 16
Abbildung 7: Unterteilung der Instandhaltung 20
Abbildung 8: Instandhaltungsstrategien 22
Abbildung 9: Aufbau aktive Redundanz 24
Abbildung 10: Aufbau passive Redundanz 24
Abbildung 11: Framework für prädiktive Instandhaltung, 10 Bestandteile 27
Abbildung 12: Direkte Abbaumodellierung 30
Abbildung 13:Exponentielle Abbaumodellierung 30
Abbildung 14: Lineare Abbaumodellierung 30
Abbildung 15: Weibullverteilung 30
Abbildung 16: Beispielbild SciDAVis 37
Abbildung 17: Beispielhafte Plots Matlap 38
Abbildung 18: Aufbau Optischer Sensor 40
Abbildung 19: Darstellung des Triebwerkquerschnittes 41
Abbildung 20: Standardabweichung der Messreihen 43
Abbildung 21: Darstellung der Korrelation der Sensoren 44
Abbildung 22: Korrelation Sensormessreihen und RUL 45
Abbildung 23: Korrelation der Sensoren 4 und 11 in zwei Phasen 46
Abbildung 24: Standardabweichung der „Principal Components“ 47
Abbildung 25: Varianz der „Principal Components“, kumuliert 48
Abbildung 26: Anteil der PC-Varianz an der Gesamtvarianz 48
Abbildung 27: Korrelation PCs und RUL 49
Abbildung 28: Korrelation PC1 und PC2 50
Abbildung 29: Korrelation PC9 und PC11 50
Abbildung 30: Funktion der Einheiten in Abhängigkeit von PC1 und t 51
Abbildung 31: Lineare Funktion der Einheiten in Abhängigkeit von PC1 und t 51
Abbildung 32: Funktionen der Einheiten nach HMM-Einteilung 53

Abbildung 33: Funktion der Einheiten nach k-Means-Einteilung 53
Abbildung 34: Funktion der Einheiten nach hierarchischer Clustermethode 53
Abbildung 35: Korrelation von PC1 und PC2 nach HMM Einteilung 54
Abbildung 36: Korrelation von PC1 und PC2 nach k-Means-Einteilung 54
Abbildung 37: Korrelation von PC1 und PC2 nach hierarchischer Clustermethode 54
Abbildung 38: Histogramm der Verteilung der realen RULs 57
Abbildung 39: Ergebnisse lineare Regression 58
Abbildung 40: Ergebnisse des quadratischen Modells 59
Abbildung 41: Ergebnis exponentielles Modell 60
Abbildung 42: Ergebnis Support-Vector-Machine 60
Abbildung 43: Ergebnis Random Forest 61
Abbildung 44: Ergebnis K-Nearest Neighbors 62
Abbildung 45: Ergebnis Neuronales Netzwerk 62
Abbildung 46: Ergebnis Mittelwert der Algorithmen 63
Abbildung 47: Entwicklung der RMSE über die zehn Durchläufe 65
Abbildung 48: Testung der RMSE über zehn Durchläufe 66

Inhaltsverzeichnis

Eidesstattliche Erklärung I
Zusammenfassung II
Abkürzungsverzeichnis III
Tabellenverzeichnis IV
Abbildungsverzeichnis V
Inhaltsverzeichnis VII

1. Einleitung 1
1.1 Ausgangssituation 1
1.2 Problemstellung 1
1.3 Zielsetzung 1
1.4 Aufbau der Arbeit 2
2. Stand der Technik 3
2.1 Industrie 4.0 3
2.2 Zuverlässigkeit in der Industrie 4.0 5
2.4 Condition Monitoring 13
2.5 Big Data 13
2.6 RUL 15
2.7 Artifizielle Intelligenz 15
3. Analyse von Instandhaltungsstrategien 19
3.1 Was ist Instandhaltung? 19
3.2 Untergliederung von Instandhaltung 19
3.3 Entwicklung von Instandhaltung 21
3.4 Was für Instandhaltungsstrategien gibt es? 22
3.4.5 Vergleich der Instandhaltungsstrategien: zustandsorientiert - herkömmlich 34
4. Praktische Datenauswertung eines Turbinendatensatzes 36
4.1 Was ist die Software R? 36
4.2 Alternative Softwareanwendungen 36
4.3 Datengewinnung 39
4.4 Vorbereitung des Datensatzes (deskriptiv) 42

4.5 Klassifizierung des Datensatzes 51
4.6 Vorhersage des RUL 55
4.7 Auswertung der Ergebnisse 63
5. Fazit 67

Literatur- und Quellenverzeichnis 68
Anlage 74

1. Einleitung

Die Basis für die Masterthesis mit dem Thema „Instandhaltungsstrategien in der Industrie 4.0 -Vergleich von zustandsorientierten und herkömmlichen Instandhaltungsstrategien mit beispielhafter Datensatzbearbeitung“ bildet die vierte industrielle Revolution. Hierdurch sind neue Anforderungen an die Unternehmen aufgekommenen. Dazu gehören unter anderem das Thema der Digitalisierung, Big Data oder auch Condition Monitoring, welche im Folgenden, Kapitel 2 näher betrachtet werden.

1.1 Ausgangssituation

Zum aktuellen Zeitpunkt sind viele Unternehmen bezogen auf die Anforderungen von Industrie 4.0 in der Findungsphase, bei der Fragen betrachtet werden, wie: Betrifft uns die Idee bzw. Vorgabe; wie könnten wir diese umsetzen, inwieweit wollen wir das umsetzen und stimmt das mit den Unternehmenszielen überein. Ganz besonders kann hierbei auch die Instandhaltungsstrategie des Unternehmens in Frage gestellt werden, ob die gelebte Strategie die Richtige ist und auch zu den neuen Anforderungen nach der Industrie 4.0 passt.

Sehr wichtig für die Instandhaltung ist, dass die Zuverlässigkeit der Anlagen beherrscht wird. Um die Zuverlässigkeit der Anlagen zu kennen und dementsprechend auch das Ausfallverhalten, müssen relevante Kennzahlen ermittelt werden. Zuverlässigkeit, deren Bedeutung und auch wie diese bestimmt wird, wird in Kapitel 2.2 erläutert.

1.2 Problemstellung

Das Problem bei der Produktionsarbeit der Zukunft stellt die steigende Verknüpfung, zwischen den Menschen, den Maschinen und den Produkten dar. Durch diese Verknüpfungen wird ein bestimmtes Maß an technischer Ausstattung vorausgesetzt. Dieser Druck der durch die vierte industrielle Revolution mit der Digitalisierung auf die Industrie ausgeübt wird, stellt viele Unternehmen vor Herausforderungen. Wenn es einem Unternehmen gelungen ist die technische Ausstattung zu beschaffen und in Betrieb zu nehmen, stehen diese Unternehmen vor dem nächsten Problem. Durch die neue Technik und die daraus entstandenen Verknüpfungen werden vielen Daten erfasst, doch die Auswertung dieser ist für viele Unternehmen eine Herausforderung. Außerdem erschwert die zunehmende Technik die Instandhaltung der produzierenden Anlagen. Am besten sollten die Anlagen keine Fehler und Ausfälle haben, da die Kosten für den Stillstand immer weiter zunehmen.

1.3 Zielsetzung

Die Zielsetzung dieser Masterthesis ist es die bestehenden Instandhaltungsstrategien etwas näher zu beleuchten und diese mit der immer weiter aufkommenden zustandsorientierten

Instandhaltungsstrategie zu vergleichen. Außerdem sollen Antworten auf die folgenden Fragen geliefert werden. Gibt es große Unterschiede zwischen den Strategien und muss jedes Unternehmen jetzt auf die moderne, datenbasierte Instandhaltungsstrategie wechseln? Was bedeutet die neue Instandhaltungsstrategie für die Zuverlässigkeit der Anlagen und welche Vorteile bietet die zustandsorientierte Instandhaltung dem Unternehmen, welches versucht die Anforderungen der Industrie 4.0 zu erfüllen? Nachdem diese ganzen Fragen behandelt wurden, wird ein Triebwerksdatensatz der NASA analysiert, um zu untersuchen, wie eine solche Prognose des Instandhaltungszeitraum auf Basis von Daten funktionieren könnte.

1.4 Aufbau der Arbeit

Die Thesis ist in fünf Bereiche unterteilt. Nach der Einleitung, in der die Ausgangslage, die Problemstellung und die Zielsetzung beschrieben wurden, wird der allgemeine Stand der Technik beschrieben. Hier werden die Grundlagen für die darauffolgenden Bereiche erklärt, wie zum Beispiel Industrie 4.0, Digitalisierung oder Big Data. Nachdem die Grundlagen geschaffen wurden, wird der Themenbereich der Instandhaltung bearbeitet. Hier ist vor allem auch die Analyse der Instandhaltungsstrategien zu erwähnen. Als vierter Teil wird ein veröffentlichter Datensatz der NASA analysiert. Hierbei sollen die Restnutzungsdauern der betrachteten Triebwerke vorhergesagt werden. Zu guter Letzt werden die Ergebnisse des dritten und vierten Teils kombiniert und in einem Fazit beschrieben.

2. Stand der Technik

2.1 Industrie 4.0

Der Begriff Industrie 4.0 ist aktuell in aller Munde. Viele Unternehmen in der Industrie versuchen sich an den Richtlinien, Strategien oder Konzepten zu orientieren, doch was genau ist unter Industrie 4.0 zu verstehen?

Erstmals konnte man den Begriff Industrie 4.0 auf der Hannover-Messe im Jahr 2011 vernehmen. Zwei Jahre später, ebenfalls auf der Hannover-Messe, wurde ein Bericht zu dem Thema „Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0“ von dem Arbeitskreis Industrie 4.0 geleitet von Siegfried Dais und Henning Kagermann veröffentlicht. Inhaltlich beschäftigt sich der Bericht mit der vierten industriellen Revolution, welche sich vor allem als Zukunftsprojekt beschreiben lässt. Dieses Projekt lässt sich in zwei Bereiche unterteilen, zum einen in den Bereich des Bedarfs und zum anderen in den Bereich der Technologie. [1]

Die vierte industrielle Revolution soll Unternehmen viele Vorteile bringen, ebenso wie die drei industriellen Revolutionen vorher zu deren Zeit. Es begann mit der ersten industriellen Revolution in der zweiten Hälfte des 18 Jhd., hierbei sticht besonders die Mechanisierung der Produktion durch Dampfmaschinen heraus.[2] Neben der Mechanisierung gab es viele gesellschaftliche Veränderungen, die auch die sozialen Verhältnisse, Lebensbedingungen und auch Arbeitsbedingungen prägten.[3] Ab den 1870ern kann von der zweiten industriellen Revolution gesprochen werden. Chemie und elektrische Energie erhielten immer mehr Einfluss auf die Industrie und erzeugten einen Entwicklungsschub[4]. Die angloamerikanische Version der zweiten industriellen Revolution beschreibt die in den 1920ern eingeführte Massenproduktion, welche von F. Taylor und H. Ford geprägt wurde. Die dritte industrielle Revolution, oder auch digitale Revolution, begann gegen Ende des 20 Jhd. Durch die zunehmende digitale Vernetzung und den digitalen Wandel der Lebensstile wurde die Industrie maßgeblich verändert. Es wurden neue Steuerungen, Überwachungen, Technologien oder Geschäftsmodelle eingeführt, wodurch sich die gesellschaftlichen Verhältnisse erneut deutlich änderten.[5]

In der Abbildung 1 ist deutlich zu erkennen, wie die Komplexität mit der Zeit und den industriellen Revolutionen immer weiter zunimmt. Die Anforderungen und Erwartungen an die Produkte steigen und die Konkurrenz nimmt zu, es wird immer schwerer für Unternehmen sich auf dem Markt zu etablieren.

---

[1] Vgl. Kagermann H. (April 2021). Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. Industriellen Revolution. VDI-Nachrichten.
[2] Vgl. Deppe K. (April 2019). Arbeit 4.0 – Geschichte der Arbeit. Planet-wissen.
[3] Vgl. Seite „Industrielle Revolution“ (September 2021). In: Wikipedia – Die freie Enzyklopädie.
[4] Vgl. Hahn H.-W. (November 2011). Die industrielle Revolution in Deutschland. De Gruyter Oldenbourg. Seite 42.
[5] Vgl. Deppe K. (April 2019). Arbeit 4.0 – Geschichte der Arbeit. Planet-wissen.

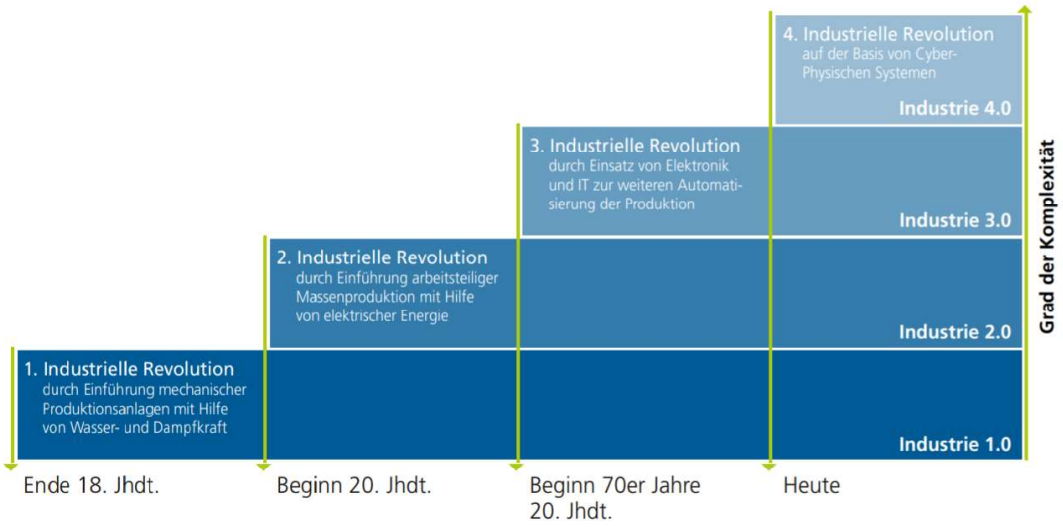

Abbildung 1: Vier Stufen der Industriellen Revolution [6]

Der Bedarfsbereich der Industrie 4.0 befasst sich mit der Veränderung von betrieblichen Rahmenbedingungen im Unternehmen, um bei der steigenden Anzahl an Konkurrenten auf dem Markt wettbewerbsfähig bleiben zu können. Beginnend mit dem Ziel die Time-to-Market zu reduzieren, über die sogenannten Losgröße 1 und der Individualisierung von Produkten[7], bis hin zu einer steigenden Flexibilität[8] in der Produktion. Aus diesen Punkten wird deutlich, dass der Kunde und sein Bedarf immer weiter in den Mittelpunkt rücken. Die Unternehmen müssen folglich immer schneller auf Veränderungen der Bedürfnisse reagieren und möglichst viele persönliche Lösungen bereitstellen können. Neben den oben genannten Veränderungen spielen die Aspekte der Dezentralisierung und Ressourceneffizienz eine wichtige Rolle.[9]

Im Bereich der Technologie ist auffällig, wie sehr Veränderungen IT-getrieben sind. Deutlich wird dies in der steigenden Mechanisierung und Automatisierung der Produktion bis hin zu einer autonomen Produktion, bzw. Produktionsreihe. Dazu kommt die zunehmende Digitalisierung, sowie Vernetzung, sowohl zwischen Menschen und Maschine, als auch zwischen Maschine und Maschine. Eine weitere Entwicklung ist die Miniaturisierung, welche besonders gut an der Entwicklung von Speichermedien oder Mobiltelefonen erkennbar ist.[10]

Es gibt ein beispielhaftes Konzept zur Umsetzung von Industrie 4.0, die Smart Factory. Die Fertigung der Smart Factory ist vollständig mit Sensoren und Aktoren ausgestattet, wodurch ein autonomes System vorliegt. Es verfügt über intelligente Technik, ganzheitliche Digitali-

[6] Vgl. Spath, D. (2013). Produktionsarbeit der Zukunft – Industrie 4.0. Frauenhofer IAO. Seite 23. Abb. 6.
[7] Vgl. Spath, D. (2013). Produktionsarbeit der Zukunft – Industrie 4.0. Frauenhofer IAO. Seite 35.
[8] Vgl. Spath, D. (2013). Produktionsarbeit der Zukunft – Industrie 4.0. Frauenhofer IAO. Seite 41.
[9] Vgl. Lasi, H. (2014). Industrie 4.0. In: Wirtschaftsinformatik 04/2014. Springer Verlag. Seite 261 f.
[10] Vgl. Lasi, H. (2014). Industrie 4.0. In: Wirtschaftsinformatik 04/2014. Springer Verlag. Seite 262. Spalte 1.

sierung und *Cyber-Physische-Systeme (CPS)*. Es orientiert sich am Menschen (Human-Focussed-Design), agiert nachhaltig und ressourceneffizient und ist dezentral gelagert.[11]

2.2 Zuverlässigkeit in der Industrie 4.0

Das Thema Zuverlässigkeit zieht sich komplett durch das Thema Industrie 4.0, auch wenn es nicht durchgehend betont wird. Zuverlässigkeit wird in der VDI 4003 definiert als ein zusammenfassender Ausdruck für Funktionszuverlässigkeit, Verfügbarkeit, Sicherheit und Instandhaltbarkeit. Was so viel bedeutet wie, dass eine Einheit zu einem gegebenen Zeitpunkt, eine geforderte Funktion unter gegebenen Bedingungen frei von unvertretbaren Risiken erfüllt und durch gegebene Anwendungen in einem Zustand erhalten oder zurückgesetzt werden kann.[12] Wie somit erkenntlich wird, ist Zuverlässigkeit ein sehr weitläufiger Begriff, der oftmals unterschiedlich verstanden werden kann. In wenigen Fällen kommt es auch vor, dass Zuverlässigkeit mit Qualität gleichgesetzt wird. Unter Qualität wird das Erfüllen von Anforderungen durch die Merkmale eines Objektes verstanden[13], was sich in gewisser Weise mit der Zuverlässigkeitsdefinition überschneidet, aber nicht dasselbe ist.

Das Zuverlässigkeitsmanagement führt, lenkt und koordiniert alle Zuverlässigkeitsaktivitäten in dem Unternehmen. Es werden ein Zuverlässigkeitsplan, ein Zuverlässigkeitsprogramm, Zuverlässigkeitsziele und Zuverlässigkeits-forderungen erstellt, bzw. definiert. Generell lässt sich das Zuverlässigkeits-management sehr gut mit dem Qualitätsmanagement vergleichen und überschneidet sich in gewissen Bereichen auch, wie dem Aufbau und dem Vorgehen. Während der Zuverlässigkeitsplan ausschließlich ein Dokument für die Zuverlässigkeits-verfahren, Mittel und Abfolgen von Tätigkeiten für ein Produkt ist, umfasst das Zuverlässigkeitsprogramm die Darstellung von Organisationsstruktur, Verantwortlichkeiten, Verfahren, Prozesse und Mittel.[14] Die Zuverlässigkeitsforderungen legen die zu erreichenden und nachzuweisenden Merkmale fest, welche oftmals ein Teil der Qualitätsforderungen sind.

2.2.1 Bedeutung von Zuverlässigkeit

Doch warum gewinnt die Zuverlässigkeit in der Industrie 4.0 noch mehr an Bedeutung? Ausgehend von der Prognose von Dr. Ferber aus dem Jahr 2013, wird davon ausgegangen, dass die Vernetzung sich auch auf die Dinge im Unternehmensumfeld und der Produktion auswirken wird. Das wiederrum bedeutet, dass die Kommunikation im Unternehmen steigt, wie in Abbildung 2 zu sehen ist. [15]
Durch die hochmodernen Anlagen gelingt es die Produktfehler zu reduzieren, wodurch jedoch die Anlagenfehler aufgrund der Komplexität der Anlagen deutlich zunehmen.[16] Hinzu kommt

[11] Vgl. Lasi, H. (2014). Industrie 4.0. In: Wirtschaftsinformatik 04/2014. Springer Verlag. Seite 262. Spalte 2f.
[12] VDI-Richtlinie 4003:2007. Zuverlässigkeitsmanagement. Seite 8.
[13] Vgl. DIN 9000-2015; Qualitätsmanagementsysteme – Grundlagen und Begriffe. Seite 39.
[14] Vgl. VDI-Richtlinie 4003:2007. Zuverlässigkeitsmanagement. Seite 8 + 9.
[15] Vgl. Spath, D. (2013). Produktionsarbeit der Zukunft – Industrie 4.0. Frauenhofer IAO. Seite 44.
[16] Ahrens D. (2016). Neue Anforderungen im Zuge der Automatisierung von Produktionsprozessen […]. In:AIS-Studien. Seite 47.

bei der steigenden Kommunikation und den damit verbundenen Datenmengen, dass die Daten richtig sind und wie geplant transportiert und analysiert werden. Somit lässt sich sagen, die Zuverlässigkeit der Maschinen und der Technik ist bedeutsamer als je zuvor und bildet die Grundlage für ein produzierendes Unternehmen.

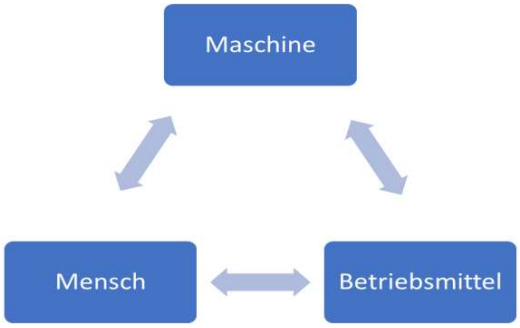

Abbildung 2: Mensch-Maschine-Betriebsmittel Vernetzung[17]

Besonders im Mittelpunkt steht hierbei der Lebenszyklus eines Produktes, welcher aus vier Lebensphasen besteht. Beginnend mit der Entwicklungsphase, über die Herstellungsphase, bis hin zu der Betriebsphase und abschließend mit der Entsorgungs- und Wiederverwertungsphase. In der Entwicklungsphase werden ausschließlich Konzepte und Konstruktionen erstellt und verifiziert. Während der Planung werden unter anderem auch Lieferanten mit einbezogen, da die Ressourcen, Halbzeuge oder Bauteile der Lieferanten relevant für die neuen Produkte sind. Nach der Prototypenproduktion beginnt die Herstellung des Produktes, die Serienproduktion. Nach der Produktion wird das Produkt an den Kunden verkauft und in Betrieb genommen. Während der Betriebsphase kann es vorkommen, dass das Produkt teilweise gewartet oder repariert werden muss. Wenn das Produkt irreparabel ist, wird es entsorgt und wenn es möglich ist für eine Wiederverwertung oder Recycling vorbereitet.[18]

2.2.2 Zuverlässigkeit Testen

Wenn Produkte, Bauteile oder Anlagen in Betrieb sind, lässt sich die Zuverlässigkeit testen, indem unterschiedliche Messungen durchgeführt werden. Diese Kennzahlen geben Aufschluss darüber, ob die Produkte die geforderten Merkmale, oftmals die geforderte Lebensdauer, erfüllen. Was bedeutet, dass die Qualität nicht der geforderten entspricht, sondern auch, dass durch die Nichterfüllung ein Sicherheitsrisiko bestehen kann, wenn es sich dabei um ein Sicherheitsrelevantes Bauteil handelt.
Daher gibt es für unterschiedliche Bereiche unterschiedliche Kennzahlen. Die bekanntesten sind wahrscheinlich die RPZ und die MTTF. Unter RPZ ist die Risikoprioritätszahl zu verstehen, welche vor allem in der Fehler-Möglichkeits-Einfluss-Analyse, kurz FMEA, Verwendung findet. Sie dient der Risikobewertung und wird durch die Multiplikation der

[17] Vgl. Ahrens D. (2016). Neue Anforderung im Zuge der Automatisierung von Produktionsprozessen […]. In: AIS-Studien. Seite 44.
[18] VDI-Richtlinie 4003:2007. Zuverlässigkeitsmanagement. Seite 5 + 6.

Bewertungszahlen B (Bedeutung des Fehlers), A (Auftretenswahrscheinlichkeit des Fehlers) und E (Entdeckungswahrscheinlichkeit des Fehlers) berechnet. Jede der Bewertungszahlen kann Werte von 1 bis 10 annehmen, wobei 1 der beste und 10 der schlechteste Wert ist. Folglich kann die RPZ im Bereich von 1 bis 1.000 liegen. Ab der vom Unternehmen festgelegten Schwelle (in der Literatur oftmals ca. 80) müssen Maßnahmen ergriffen werden, um eine oder mehrere der Bewertungszahlen zu senken und somit final auch die RPZ.[19]

In der EN ISO 13849-1 geht es um die Sicherheit von Maschinen. Eine große Bedeutung hat hier die Kennzahl MTTF. MTTF steht für Mean-Time-To-Failure[20], was zu Deutsch so viel bedeutet wie mittlere Zeit bis zum Ausfall, rechnerisch lässt sich die MTTF unkompliziert bestimmen, indem die summierten Betriebszeiten durch die in der Zeit genutzten Geräte dividiert werden.[21]

$$MTTF = \frac{\sum Betriebsstunden}{\sum genutzen\ Geräte}$$

Da die MTTF eine sehr allgemeine Kennzahl ist, gibt es die Möglichkeit mit der $MTTF_d$, der mittleren Zeit bis zu einem gefährlichen Ausfall, ein sicherheitsrelevantes Objekt oder System genauer zu beschreiben[22]. Die $MTTF_d$ wird wie folgt berechnet:

$$MTTF_d = \frac{B_{10D}}{0{,}1 * n_{op}}$$

B_{10D} die mittlere Anzahl von Zyklen bis 10 % der Bauteile gefährlich ausgefallen
n_{op} die mittlere Anzahl jährlicher Betätigungen [23]

Es gibt noch weitere Kennzahlen, die eine vergleichbare Bedeutung haben wie die MTTF. Für die Instandhaltung gibt es die MTBF oder für die Verkehrsmittel die MDBF ist. Die MTBF ist die mittlere Zeit zwischen den Ausfällen und unterscheidet sich nur von der MTTF durch die betrachteten Objekte. Bei MTTF wird immer nur ein vollständiger irreversibler Schaden betrachtet, bei der MTBF hingegen Objekte die reparierbar sind. Somit werden die Zeitabstände zwischen den Ausfällen bzw. Reparaturen betrachtet und wie folgt berechnet:[24]

$$MTBF = \frac{Gesamtbetriebsstunden}{Anzahl\ der\ Ausfälle}$$

[19] Vgl. Kaminske G. (2013). Handbuch QM-Methoden. Carl Hanser Verlag GmbH & Co.KG. Seite 227.
[20] Vgl. DIN EN ISO 61124:2013-01. Prüfungen der Funktionsfähigkeit. Seite 10.
[21] DIN EN ISO 13849-1: 2016-06. Sicherheit von Maschinen. Seite 19.
[22] DIN EN ISO 13849-1: 2016-06. Sicherheit von Maschinen. Seite 16.
[23] DIN EN ISO 13849-1: 2016-06. Sicherheit von Maschinen. Seite 70.
[24] Vgl. DIN EN ISO 61124:2013-01. Prüfungen der Funktionsfähigkeit. Seite 9.

Im großen industriellen Sektor der Verkehrsmittel, vor allem der Automobilindustrie, gibt es speziell auf diesen abgestimmte Kennzahlen. Die Kennzahl Mean Distance between Failure, kurz MDBF, zeigt die mittlere Distanz zwischen den Ausfällen. An sich ist die MDBF dieselbe wie die MTBF, nur bezogen auf die zurückgelegte Strecke.[25]

$$MDBF = \frac{Gesamtdistanz}{Anzahl\ der\ Ausfälle}$$

Eine andere für die Zuverlässigkeit relevante Kennzahl ist die Ausfallrate. Sie gibt an, welche Menge an Objekten in einem definierten Zeitintervall ausfallen. Die Einheit der Ausfallrate λ wird als FIT bezeichnet, was für Failure in Time steht. Die Einheit von FIT sind die Ausfälle pro 10^9 Stunden.[26] Anders gesagt es ist der Kehrwert der MTTF bzw. bei reparablen Objekten der Kehrwert der MTBF.

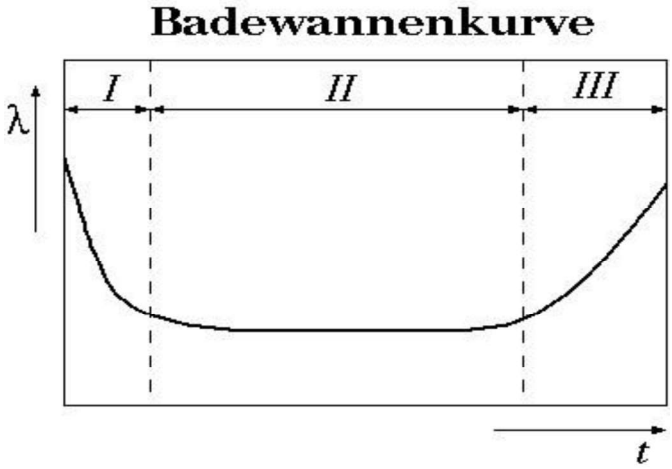

Abbildung 3: Ausfallverhalten Badewannenkurve [27]

Es gibt viele unterschiedliche Ausfallraten, von exponentiellen, über lineare, bis hin zu konstanten. Eine der bekanntesten Ausfallraten ist die sogenannte Badewannenkurve. Die Badewannenkurve zeigt die Ausfallraten in Abhängigkeit von dem Alter des Objektes an. Die Ausfallverteilung oder auch Ausfall-wahrscheinlichkeit, unterteilt sich wie in Abbildung 3 zu sehen ist in 3 Bereiche. Die erste Phase beschreibt die Frühausfälle, welche oft auch als Kinderkrankheiten bezeichnet werden. Hierbei handelt es sich oft um Objekte, die aufgrund von Produktions- oder Materialfehlern frühzeitig ausfallen. Wenn die Menge der Ausfälle sich stabilisiert hat und die Objekte nur noch in geringen Mengen ausfallen, bewegt man sich in der zweiten Phase. Mit zunehmender Lebensdauer der Objekte steigt die Beanspruchung und der Verschleiß dieser, aus diesem Grund steigt in der dritten Phase die Ausfallverteilung wieder an.[28]

[25] Vgl. Seite „Mean Distance Between Failure“(April 2016). In: Wikipedia – Die freie Enzyklopädie.
[26] Vgl. DIN EN 61709:2015-01. Elektrische Bauelemente – Zuverlässigkeit. Seite 72.
[27] Seite „Ausfallverteilung“(April 2019). In: Wikipedia – Die freie Enzyklopädie.
[28] Vgl. Tordeux A. (2019). Vorlesung: Technische Zuverlässigkeit. Veranstaltung: 3. Folie 5.

Solche grafischen Darstellungen der Ausfallverteilungen stellen vorliegende Probleme eindeutig dar. Wenn zum Beispiel die erste Phase „zu groß“ ausfällt sollte das eingesetzte Material bzw. die verwendete Produktionstechnik überdacht werden oder die Objekte sollten in der Produktion künstlich gealtert werden.

2.2.3 Zuverlässigkeit Planen

Um Probleme durch unzureichende Zuverlässigkeit zu vermeiden, ist es möglich in gewisser Weise Zuverlässigkeit zu planen, ähnlich wie bei der Qualitätsvorausplanung. Gut geeignet dafür ist die Qualitätsmanagementmethode Design for Six Sigma (DFSS), bei dem das Ziel verfolgt wird, dass die Produkte und Prozesse von Anfang an so entwickelt werden, dass sie robust gegenüber Störeinflüssen sind, Six Sigma Qualität erreichen und die Anforderungen erfüllen. Dabei werden Kunden und deren Bedürfnisse an die erste Stelle gestellt und systematisch in statistisch abgesicherte Produkt- und Prozesstoleranzen übersetzt.

Tabelle 1: DMADV-Phasen, Ziele, Methoden[29]

DMADV-Phase	Phasenziele	Methoden
Define	• Zielsetzung definieren • Risikobewertung des Projektes • Projektplanerstellung • Teambenennung	• Marktanalyse • Projektdefinition • Projektrisikoanalyse • Gantt-Chart
Measure	• Informationssammlung der Kundenanforderungen und Prioritäten • Überführung von Kundenanforderungen in funktionale Anforderungen	• Interviews • Kano-Modell • Benchmarking • ABC-Klassifizierung • QFD I

Analyze	• Erstellen, Evaluieren und Auswahl des Designkonzepts	• Kreativitätstechniken
		  ○ Brainstorming
		  ○ 6-3-5 Methode
		  ○ Mindmapping
		• System FMEA
		• QFD II

Design	• Entwicklung des detaillierten Designkonzepts • Test der Komponenten und des Systems • Vorbereitung der Pilotierung und Serienproduktion	• Simulation • Prozessanalyse • Design of Experiments
Verify	• Verifizierung der Designperformance • Pilotierung und Implementierung des finalen Designs • Übergabe der Prozess- & Produktverantwortung	• Designreview • Zuverlässigkeitsberechnung • Control Charts • Testplan

Im DFSS gibt es unterschiedliche Systematik Möglichkeiten zur Umsetzung dieser, drei der bekanntesten Systematiken sind DMADV (Define, Measure, Analyze, Design, Verify), DCOV (Design, Characterize, Optimize, Verify) und IDOV (Identify, Design, Optimize, Validate)Jede der drei Systematiken hat vergleichbare Methoden und Werkzeuge, dennoch un-

[29] Vgl. Kaminske G. (2013). Handbuch QM-Methoden. Carl Hanser Verlag GmbH & Co.KG. Seite 233f, Tabelle 8.3.

terscheiden sie sich anhand der Phasenreihenfolge.[30] Im Folgenden wird besonders die Systematik DMADV eingegangen.

Die **Define**-Phase beschreibt einen klassischen Projektbeginn, es werden die Ziele, die Ausgangssituation und Einflüsse beschrieben, sowie definiert. Hinzu kommt die Erstellung eines Projektplans, was unter anderem die Teamerstellung, die Zeitplanung im Gantt-Chart und die Vorgehensweise des Projektes beinhaltet.[31]

In der folgenden Phase der **Measure**-Phase werden die Anforderungen von Kunden, Gesetzen oder auch Normen gesammelt, priorisiert und mit Hilfe der QFD I in funktionale Anforderungen überführt. Die Überführung von Anforderungen in funktionale Anforderungen ist unvermeidbar, da dadurch die Anforderungen mit messbaren Kennzahlen beschrieben werden. Folglich lässt sich mit Abschluss des Projektes die Erfüllung der Anforderungen mit Hilfe der Kennzahlen besser bewerten.[32]

Die dritte Phase **Analyze** befasst sich mit der Erstellung, Evaluierung und dem Auswählen des Designkonzeptes des Zielproduktes/ Zielobjektes. Für die Erstellung und Ausarbeitung der Designkonzepte werden unterschiedliche Kreativitätsmethoden eingesetzt, die bekanntesten Kreativitätsmethoden sind das Brainstorming, das Mindmapping und die 6-3-5 Methode. Ziele der Methoden sind es neue Ideen zu sammeln, die Ideen anderer zu erweitern und vorliegende Gedankenbarrieren zu durchbrechen. Für die darauffolgende Evaluierung der Designkonzepte werden System-FMEAs und die QFD II (Quality-Function-Deployments) auf die jeweiligen Konzepte angewendet. Somit können durch die FMEA mögliche zukünftige Fehler durch Gegenmaßnahmen proaktiv vermieden werden und durch die QFD II die „Stimme des Kunden“ in die Planung und Entscheidung berücksichtigt werden.[33]

Mit der Wahl eines Designkonzeptes beginnt die **Design**-Phase. Das Designkonzept wird immer detaillierter weiterentwickelt und durch das Störgrößenmanagement soll das Design robuster gegenüber Störgrößen und äußeren Einflüssen werden. Während der kompletten Entwicklung werden Simulationen, Prozessanalysen und DoEs (Design of Experiments) eingesetzt. Wenn nach wenigen Wochen/ Monaten das Designkonzept fertig entwickelt wurde, werden alle Komponenten und Prozesse getestet, Prototypen für Kunden und die Pilotierung erstellt und die Produktion vorbereitet.[34]

Die letzte Phase **Verify** beginnt mit der Pilotierung des Prototyps. Hierbei wird dieser auf Performance und Robustheit getestet, indem Zuverlässigkeitswerte, wie MTTF oder MTBF ermittelt werden. Des Weiteren erfolgt eine Kontrolle der Einhaltung aller Regularien und Richtlinien, sodass bei erfolgreicher Abschlusskontrolle das Produkt freigegeben werden kann und eine Herstellung großer Stückzahlen fokussiert werden kann.[35]

[30] Vgl. Kaminske G. (2013). Handbuch QM-Methoden. Carl Hanser Verlag GmbH & Co.KG. Seite 219 + 220.
[31] Vgl. Kaminske G. (2013). Handbuch QM-Methoden. Carl Hanser Verlag GmbH & Co.KG. Seite 220f + 233.
[32] Vgl. Kaminske G. (2013). Handbuch QM-Methoden. Carl Hanser Verlag GmbH & Co.KG. Seite 221ff + 234.
[33] Vgl. Kaminske G. (2013). Handbuch QM-Methoden. Carl Hanser Verlag GmbH & Co.KG. Seite 224ff + 234.
[34] Vgl. Kaminske G. (2013). Handbuch QM-Methoden. Carl Hanser Verlag GmbH & Co.KG. Seite 226ff + 234.
[35] Vgl. Kaminske G. (2013). Handbuch QM-Methoden. Carl Hanser Verlag GmbH & Co.KG. Seite 231ff + 234.

2.2.4 Zuverlässigkeit Verbessern

Doch was soll gemacht werden, wenn im Unternehmen ermittelt wurde, dass die Zuverlässigkeit bestimmter Produkte nicht ausreichend ist? Um Probleme zu vermeiden kann das Produkt vom Markt genommen werden. Doch lohnt sich dieser drastische Schritt immer oder gibt es Möglichkeiten durch das Drehen an bestimmten Stellschrauben, dass das Problem der Zuverlässigkeit gelöst wird, darauf wird im folgenden Bezug genommen.

Es kann sein, dass durch die Kennzahlermittlung, Laborprüfungen oder übermäßig viele Reklamationen ein Problem in z.B. der Produktion auffällig wird. Es gibt unterschiedliche Maßnahmen, um ein solches Problem zu verbessern. Zunächst sollte, wenn das „Problemteil" noch nicht identifiziert ist, eine Schwachstellenanalyse durchgeführt werden, damit die kritischen Komponenten gezielt ausgebessert werden können. Neben der Schwachstellenanalyse sind Fehlerbäume oder FMEAs gut für die Komponentenanalyse geeignet. [36]

Es ist möglich die Komponente generell auszutauschen oder das System durch eine parallelgeschaltete vergleichbare Komponente vor einem Ausfall zu bewahren. Eine andere Option ist es, durch z.B. direktes Abschalten der Anlage beim Ausfall der Komponente die Anlage vor Folgeschäden zu schützen, ganz nach dem „Fail-Safe Prinzip[37]". Wenn aber Potential zur Verbesserung der Komponente besteht, sollte die Komponente durch die gewonnenen Erfahrungswerte so entwickelt werden, dass die Lebensdauer und Beanspruchbarkeit erhöht wird.

Wenn aber die Komponente weder verbessert noch ausgetauscht werden kann, sollten die Komponentenprüfungen vor dem Einbau verbessert und strenger werden, um somit schwächere Komponenten herauszufiltern. Mit strengeren Prüfungen sind Prüfungen mit kleineren Toleranzen beschrieben, folglich dürfen die Komponentenprüfwerte weniger vom Zielwert abweichen, wie in Abbildung 4 zu sehen ist. Hier wurde die erlaubte Toleranz von 5mm (± 2,5 mm) auf 4 mm (± 2 mm) gesenkt, wodurch mehr Komponenten als n.i.o. eingestuft werden, wie durch die rote Fläche gekennzeichnet wurde.

Final gibt es noch die Möglichkeit die Lebensdauer der Komponente zu erhöhen, indem die Komponente durch geschultes Instandhaltungspersonal gewartet wird.

[36] Vgl. Löwer, M. (2019). Vorlesung: Qualitätsvorausplanung in der Entwicklung. Veranstaltung: 6 + 7.
[37] Vgl. Wissen.de, Hrsg. (o.D.). Fail-safe-Prinzip.

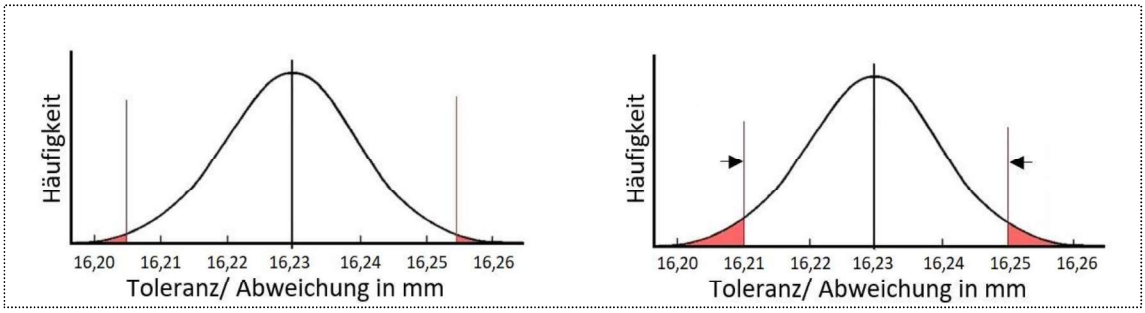

Abbildung 4: Normalverteilung, Anpassung der Toleranz / erlaubten Abweichung

2.3 Digitalisierung

Über Digitalisierung wird aktuell in vielen Bereichen berichtet und bietet Konfliktpotential. Einer der Hauptaspekte in den Medien ist die Digitalisierung in den Schulen, sie soll vorangetrieben werden, um moderne Unterrichtsmethoden durchführen zu können und auf dem neusten Stand der Technik lehren zu können. Der andere Hauptaspekt, in dem über den Begriff Digitalisierung berichtet wird, ist Industrie 4.0. Doch was genau unter Digitalisierung zu verstehen ist wird selten richtig aufgegriffen.

Eine der einfachsten allgemeinen Definitionen beschreibt Digitalisierung als die Umwandlung analoger Inhalte und Prozesse in eine digitale Form oder Arbeitsweise. Ein passendes Beispiel hierfür ist die Umwandlung von VHS-Kassetten oder Schallplatten in z.B. mp3- oder mp4-Dateien.[38] Für die Industrie muss die Definition der Digitalisierung ein wenig angepasst werden, um die optimal anwendbare Bedeutung zu haben. Digitalisierung in der Industrie beschreibt eine schnelle Verarbeitung von Daten, umwandeln, transportieren, verarbeiten, speichern und ausgeben sind ein relevanter Bestandteil der neuen automatisierten und vernetzen Produktion.[39] Was aus der Definition hervorgeht ist, dass Unternehmen technischer ausgestattet werden sollen, damit Prozesse über Computer geplant oder ausgeführt werden können. Ein wichtiger Part der Digitalisierung der Unternehmen ist das Condition Monitoring (siehe 2.4) und der Bereich der Big Data (siehe 2.6). Durch die steigende Verknüpfung im Unternehmen wird ebenfalls ein flexibles Arbeitssystem unterstützt. Arbeitszeiten können flexibler gestaltet werden, Arbeit von dezentralen Orten ist leichter zu vollziehen und Leute mit Mobilitätseinschränkung können leichter im Homeoffice arbeiten. Weitere Vorteile von Digitalisierung wären die Senkung von Transaktionskosten und die Steigerung der allgemeinen Effizienz[40]. Nachteile von Digitalisierung sind primär die damit zusammenhängenden Investitionen, sowohl zeitlich als auch finanziell und die Datensicherheit. Wenn im Unternehmen alle Dokumente und Daten, welche weiterhin zunehmen werden, auf Servern und Computern abgelegt und gesichert sind, besteht das Risiko bei Cyberangriffen einen deutlich höheren Schaden zu nehmen, als vor der Digitalisierung. Die Bedeutung und Verantwortung der IT nimmt somit deutlich zu. Ein

[38] Vgl. Tarkowski, P. (Februar 2018). Digitalisierung: Was ist das? Digitalmagazin.
[39] Vgl. Bundeszentrale für Politische Bildung (März 2020). Digitalisierung. Seite 4 + 37f.
[40] Vgl. Bundesministerium für Arbeit und Soziales (2017). Weißbuch- Arbeiten 4.0.Seite 74ff + 85.

weiterer Nachteil in Bezug auf die Digitalisierung ist die Datenschutzgrundverordnung, kurz DSGVO. Es sollen einerseits persönliche Daten geschützt werden, andererseits müssen bestimmte persönliche Daten gesammelt werden, um durch Nutzererfahrung das Produkt oder den Service verbessern zu können. In dem Fall den optimalen Mittelweg zu finden ist nicht immer einfach.[41]

2.4 Condition Monitoring

Wie schon in Kapitel 2.3 angeschnitten wurde, ist das Condition Monitoring ein wichtiger Part der Digitalisierung, bei der die kontinuierliche Zustandsüberwachung im Mittelpunkt steht. Die Zustandsüberwachung ist besonders auf die Überwachung von Maschinenparametern ausgerichtet. Hierzu können Vibrationen, Akustik, Durchfluss, Temperatur, Druck, Schwingungen oder Spannung gemessen werden.[42] Bis zu Beginn der vierten industriellen Revolution war es üblich Anlagen der Produktion und Fertigung nach Plan zu warten oder bei Ausfällen zu reparieren. Condition Monitoring ermöglicht es, wichtige Daten über die Anlagen während des Normalbetriebs zu sammeln, miteinander und mit Sollwerten zu vergleichen und zu analysieren. Die Analyse der Zustandswerte ergibt einen Anlagenzustand. Die Abweichung bestimmter Werte vom Sollzustand kann bereits bestehende Defekte aufzeigen oder ermöglicht es einen baldigen Ausfall vorauszusagen.[43]

Während die Datenerfassung der Zustandsüberwachung der Anlagen zugrunde liegt, kann die Instandhaltung mit den Ergebnissen der Zustandsüberwachung eine zustandsorientierte Instandhaltung einführen. Wie Zustandsorientierte Instandhaltung abläuft und was für Auswirkungen diese auf das Unternehmen hat, wird in Absatz 3.4.4 erklärt.

Um erfolgreiches Condition Monitoring einführen zu können, wird eine Investition von Unternehmensseite nötig sein. Es wird ein funktionierendes Softwaresystem benötigt, auf dem die großen Datenmengen abgelegt und gesichert werden. Außerdem werden Sensoren benötigt, die auf die Anlagen angepasst sind. Das bedeutet es werden Sensoren benötigt, die die richtigen physikalischen Messungen durchführen, an der ausgewählten Stelle messen können und den richtigen Messbereich haben. Somit muss in die zugrundeliegende IT und die Sensoren investiert werden. Zum Beispiel wird bei rotierenden Maschinen meistens die Lautstärke/ Akustik und die vorhandenen Schwingungen gemessen, denkbar wäre aber auch eine Temperaturüberwachung.[44]

2.5 Big Data

Wenn Themen wie Digitalisierung und Condition Monitoring besprochen werden, darf der Bereich der Big Data nicht vergessen werden, da dieser enorm wichtig für eine erfolgreiche Umsetzung der eben genannten Themen ist. Doch, bevor die Bedeutung von Big Data für

[41] Vgl. Wendl M. (April 2021). Nachteile Digitalisierung.
[42] Vgl. Verma N. (2016). Intelligend condition based Monitoring. Springer Verlag. Seite 292.
[43] Vgl. Seite „Condition Monitoring" (o.D.). In: Harmonicdrive.de.
[44] Vgl. Verma N. (2016). Intelligend condition based Monitoring. Springer Verlag. Seite 292f.

Digitalisierung und Condition Monitoring näher erläutert werden kann, muss zunächst einmal klar sein, was unter Big Data zu verstehen ist.

Vorab lässt sich sagen, dass der Begriff „Big Data“ seit der Veröffentlichung des Grün- bzw. Weißbuches im Jahre 2015 und 2016 regelrecht zu einem Modebegriff geworden ist. Viele Menschen verbinden mit Big Data Digitalisierung, große Maschinen und Hochleistungsserver. Allgemein beschreibt Big Data Datenmengen, deren Verarbeitung bzw. Analyse nicht mehr händisch durchgeführt werden kann[45] und sich aus der Verarbeitung großer Datenmengen mit Hilfe von Lernalgorithmen entwickelt hat.[46]

Es ist möglich Big Data durch die Differenzierung in unterschiedliche Bereiche detaillierter zu beschreiben und somit auch relevante Anforderungen an das IT-System formulieren zu können. Zur Unterscheidung wurden somit drei V´s festgelegt, Volume, Velocity und Variety, was so viel bedeutet wie, Volumen/Speicherkapazität, Geschwindigkeit und Vielfalt. Um aber weitere Eigenschaften besser differenzieren zu können, wurden die drei V´s um zwei weitere ergänzt, Value/ Mehrwert und Veracity/ Vertrauenswürdigkeit, siehe Abbildung 5.[47]

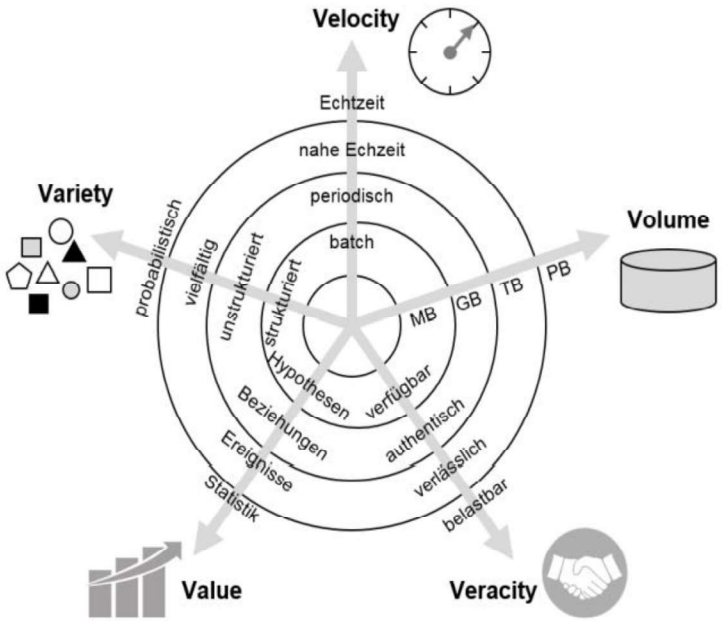

Abbildung 5: Differenzierung von Big Data [48]

[45] Vgl. Gölzer, P. (2017). Big Data in Industrie 4.0. doctoral dissertation an der FAU. Seite 62.
[46] Vgl. Yu, W. (2019). A global manufacturing Big Data Ecosystem for Fault Detection. In: IEEE. Seite 184.
[47] Vgl. Gölzer, P. (2017). Big Data in Industrie 4.0. doctoral dissertation an der FAU. Seite 63.
[48] Klein et al. (2013). Big Data. In: Informatik Spektrum 3/2013. Seite 320.

Diese fünf Punkte können noch um Datentransformationsfluss und Datensicherheitsinfrastruktur ergänzt werden, um den Aspekt des Datenschutzes mit einzubringen.[49]

Der Vorteil von Big Data Anwendungen ist der organisierte und koordinierte Umgang mit großen Datenmengen, wodurch z.B. Condition Monitoring oder vergleichbare Datenüberwachungs-/ Datensammlungsmethoden, die durch die steigende Digitalisierung an Bedeutung gewinnen, ermöglicht werden und somit auch die Mensch-Maschine-Interaktion fördert.

2.6 RUL

Bei der Anschaffung neuer Anlagen, Produkte, etc. wird oftmals die Nutzungsdauer der Neuanschaffung mit in Betracht gezogen. Vor allem in der Automobilindustrie wird häufig mit zu erwartenden Nutzungsdauern gearbeitet. Ein Keilriemen soll ca. 85.000 Km aushalten[50], eine Kupplung ca. 200.000 Km[51] und die Bremsbeläge je nach Qualität, 30.000 Km bis 60.000 Km bei günstigeren und 100.000 Km bis 150.000 Km bei hochwertigen.[52] Bei diesen Werten handelt es sich um die erwartete Nutzungsdauer der Teile. Anders gesagt können diese Werte auch als UL; usefull life (Nutzungsdauer); bezeichnet werden. Entweder können UL-Werte durch Durchschnittswerte aus Erfahrungen geschätzt werden oder sie werden nachträglich durch die genutzte Dauer ermittelt.

Wenn aber von der restlichen Nutzungsdauer während der Nutzung gesprochen wird, ist von der RUL; remaining usefull life; die Rede. Die RUL ist besonders für die Instandhaltung und die Wirtschaftlichkeitsrechnung von besonderer Bedeutung. Da durch die zu erwartende Nutzungsdauer, auf Basis von technischen, ökonomischen Daten und Erfahrungswerten, eine bessere und vor allem genauere Planung von Investitionen und Rüstungszeiten getätigt werden kann.[53] Ermittelt wird die RUL, indem von der UL die bis dato erreichte Nutzung subtrahiert wird. Diese Rechnung kann bespielhaft an dem Zahnriemen eines Autos gezeigt werden. Ein fabrikneues Auto wird 55.000 Km gefahren, somit beträgt die RUL des Keilriemens des Autos 30.000 Km.

2.7 Artifizielle Intelligenz

Als Artifizielle Intelligenz, kurz AI, oder auch künstliche Intelligenz ist die Erforschung von Rechenverfahren zu verstehen, die es ermöglichen, wahrzunehmen, zu schlussfolgern und zu handeln. Aktuelle Ziele der Wissenschaftler sind es, durch AI menschliche Wahrnehmung, Verhaltensweisen und denken simulieren zu können.[54]

[49] Vgl. Yu, W. (2019). A global manufacturing Big Data Ecosystem for Fault Detection. In: IEEE. Seite 184.
[50] Redaktion Focus Online (Februar 2018). Wann den Keilriemen wechseln?
[51] Farwer, L. (September 2018). Lebensdauer einer Kupplung.
[52] Aschermann, T. (Juni 2020). Lebensdauer von Bremsscheiben und Bremsbelägen.
[53] Vgl. Wirtschaftslexikon (Hrsg.) (o.D.). Restnutzungsdauer.
[54] Vgl. Görz, G. & Schneeberger, J. (2010). Handbuch der künstlichen Intelligenz. Walter de Gruyter. Seite 1.

AI nimmt an Wert für die Wissenschaft und Unternehmen immer weiter zu. Durch sehr große gespeicherte Datenmengen, immer größere werdende Arbeitsspeicher, Prozessorkapazitäten und zunehmende Vernetzung weltweit sind die Grundlagen für eine erfolgreiche Weiterentwicklung der AI gegeben. Des Weiteren wird die AI-Entwicklung durch milliardenschwere Investitions- und Förderprogramme unterstützt.[55]

Artifizielle Intelligenz hat viele unterschiedliche Anwendungsbereiche.

- Perfomance- und Metawissen durch formale allgemeine logikbasierte Systeme
- Heuristische Suchverfahren und Problemlösemethoden
- Maschinelles Lernen
- Komplexe Planungsaufgaben und dynamische Planrevision
- Neuronale Netze
- Verarbeitung natürlicher Sprache [56]

Wie aus der Aufzählung zu erkennen ist, hat AI Einfluss in viele unterschiedliche Bereiche des Lebens. Besonders relevant ist das maschinelle Lernen für die Industrie. Maschinelles Lernen kann definiert werden als ein Programm oder ein Algorithmus, der in der Lage ist, mit minimaler oder ohne Unterstützung zu lernen. Er kann dabei helfen Probleme wie Big Data, Bildverarbeitung, Spracherkennung oder Robotik zu vereinfachen oder gar zu lösen.[57]
AI kann bezogen auf die Lernmethoden weiter unterteilt werden (s. Abb. 4), maschinelles Lernen ist ein Bestandteil bzw. Anwendungsbereich der AI. Zum maschinellen Lernen zählt u.a. Deep Learning oder Reinforcement Learning.

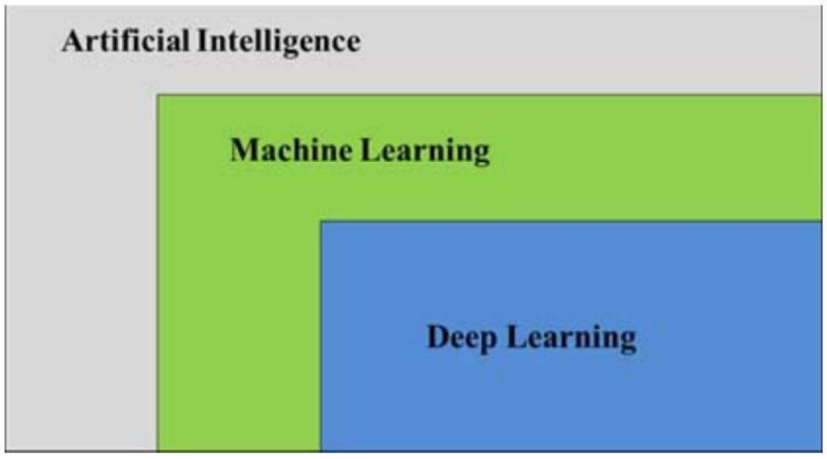

Abbildung 6: Struktur von Lernmethoden [58]

Doch es kann nicht jede Art von AI für alle oben genannten Einsatzbereiche angewendet werden. Eine AI kann im Groben in zwei Klassen eingeteilt werden. Zum einen gibt es

[55] Vgl. Görz, G. & Schneeberger, J. (2010). Handbuch der künstlichen Intelligenz. Walter de Gruyter. Seite 1.
[56] Vgl. Görz, G. & Schneeberger, J. (2010). Handbuch der künstlichen Intelligenz. Walter de Gruyter. Seite 20ff.
[57] Vgl. Amruthnath, N. (April 2018). A research study on unsupervised machine leaning[..]. IEEE. Seite 1.
[58] Amruthnath, N. (April 2018). A research study on unsupervised machine leaning[..]. IEEE. Seite 1, Figur 1.

schwache AI und zum anderen starke AI. Eine schwache oder auch spezialisierte AI, kann ausschließlich, eine vorabdefinierte Art von Informationsverarbeitung, ausführen und definierte Aufgaben lösen. Was so viel bedeutet wie, dass eine Texterkennungssoftware nur Texte verarbeiten und auswerten kann und eine Bildverarbeitungssoftware nur Bilder. Jede schwache AI hat ihren Einsatzbereich und diese sind im normalen Leben überall integriert[59,60], dazu zählen die Gesichtserkennungssoftware zum Entsperren des Smartphones, die Spamerkennung im E-Mail-Postfach oder virtuelle Assistenten wie Siri oder Alexa, die Text zu Sprache mit Sprache zu Text Anwendungen kombinieren. Außerdem ist die Anwendung von spezialisierter AI im Bereich des maschinellen Lernens äußerst relevant, da die Datenanalyse von sensorerfassten Maschinendaten immer mehr an Bedeutung gewinnt. [61] Im Vergleich zu der schwachen AI hat die starke, allgemeine AI einen offenen Einsatzbereich. Starke AI soll in der Vorstellung dem menschlichen Denken gleichen,[62] wodurch Probleme erkannt, bewertet, analysiert und final auch gelöst werden können. Die Grundlage für eine ausgeprägte starke AI ist es, dass diese lernt zu lernen und sich somit „weiterentwickelt". Beispiele für starke AI werden besonders in Science-Fiction-Filmen aufgezeigt, Humanoide (aus „I, Robot"), Hal 9000 (aus „Odyssee im Weltraum") und der Terminator (aus „Terminator"). Im realen Leben ist die Wissenschaft noch nicht so weit mit der Entwicklung von starker AI, stattdessen wird starke AI besonders im Bereich des maschinellen Lernens, bzw. „Reinforcement Learning" eingesetzt. Hier soll die AI lernen die Maschinen, Anlagen oder Produktionsprozesse zu verstehen und wie ein „Arbeiter" in / an einem Bereich zu arbeiten.[63]

Ein neuer Entwicklungsschritt wurde von dem IPA-Team am Zentrum für Cyber-Cognitive-Intelligence in Stuttgart entwickelt, welche AI zertifizierbar und transparent macht, sowie die Zuverlässigkeit des Maschinellen Lernens und der dazugehörigen AI bewertet. Es wurden fünf Kriterien herausgearbeitet, die die Grundlage einer künftigen AI-Zertifizierung sein können.

1. Alle AI-Entscheidungen müssen für den Menschen verständlich sein.
2. Die AI-Funktionen müssen vor dem Einsatz formal verifiziert werden.
3. Eine statistische Validierung der AI ist notwendig; z.B. Testläufe mit großen Datenmengen.
4. Die Unsicherheit neuronaler Netze muss vorab ermittelt und quantifiziert werden.
5. Die Systeme müssen während des Betriebs dauerhaft überprüft werden, besonders Input und Output; z.B. durch Online-Monitoring.

[59] Vgl. SimpliKI, Hrsg. (o.D.). Was sind schwache KI und starke KI?
[60] Vgl. Görz, G. & Schneeberger, J. (2010). Handbuch der künstlichen Intelligenz. Walter de Gruyter. Seite 6f.
[61] Vgl. Semmelmann, K. (August 2020). Was ist künstliche Intelligenz (KI)? Data Driven Company.
[62] Vgl. SimpliKI, Hrsg. (o.D.). Was sind schwache KI und starke KI?
[63] Vgl. Semmelmann, K. (August 2020). Was ist künstliche Intelligenz (KI)? Data Driven Company.

Des Weiteren gibt es drei Kriterien, die eine AI laut der zusammengefassten Richtlinien der HLEG AI erfüllen muss, um transparent zu sein.

1. Alle AI-Entscheidungen müssen nachvollziehbar sein.
2. Der Mensch muss AI-Entscheidungen auf umfassender Ebene des menschlichen Verständnisses erklären können.
3. AI-Systeme müssen mit dem Menschen kommunizieren und ihn informieren, was das System kann und wo die Systemgrenzen sind.[64]

[64]Vgl. O.A. (Juli 2021). KI fit machen für sicherheitskritische Anwendungen. In: Qualität und Zuverlässigkeit. Seite 6.

3. Analyse von Instandhaltungsstrategien

Nachdem die grundlegenden Informationen in Kapitel 2 betrachtet wurden, wird in Kapitel 3 das Thema Instandhaltung und der Vergleich der unterschiedlichen Strategien behandelt.

3.1 Was ist Instandhaltung?

Zuallererst werden in diesem Unterpunkt die Begriffe Instandhaltung und Instandhaltungsstrategien definiert, um eine Grundlage für die nachfolgenden Punkte generieren zu können. Als Instandhaltung wird die Kombination aller technischen und administrativen Maßnahmen, sowie Maßnahmen des Managements während des Lebenszyklus eines Objekts verstanden, die dem Erhalt oder der Wiederherstellung seines funktionsfähigen Zustands dient, sodass es die geforderte Funktion erfüllen kann.[65] Besonders stark beeinflusst wird die Instandhaltung durch die Instandhaltungsstrategie des Unternehmens bzw. des Managements. In der Instandhaltungsstrategie wird die Vorgehensweise des Managements zur Erreichung der Instandhaltungsziele festgelegt. Diese Ziele sind sehr unterschiedlich und können z. B. Verfügbarkeit, Kostenreduzierung, Produktqualität, Umweltschutz, Sicherheit, Brauchbarkeitsdauer, sowie Werterhaltung der Anlage einschließen.[66]

Doch was genau bringt Instandhaltung? Instandhaltung soll den funktionsfähigen Zustand einer Anlage maximieren und den anlagenbedingten Stillstand auf ein Minimum reduzieren.[67] Außerdem soll Instandhaltung Mitarbeiter sowie die Umwelt schützen, da Ausfälle oder Fehler von sicherheitsrelevanten Teilen schwerwiegende Konsequenzen haben können.

3.2 Untergliederung von Instandhaltung

Wenn von Instandhaltung gesprochen wird, ist allerdings nicht immer dasselbe gemeint, da der Begriff noch weiter untergliedert werden kann. Durch diese Untergliederung wird deutlich, wie unterschiedlich Instandhaltungsstrategien ausfallen können. Dazu ist die Instandhaltungsstrategie auch abhängig von den betrachteten Objekten. Doch mehr dazu im Abschnitt 3.3.

In Abbildung 7 ist dargestellt in welche vier Bereiche Instandhaltung unterteilt wird. Wartung, Inspektion, Instandsetzung und Verbesserung sind diese vier Bestandteile. Jeder der Bereiche liefert einen anderen Ansatz der Instandhaltung.

[65] DIN EN 13306:2018-02. Instandhaltung - Begriffe der Instandhaltung. Seite 8.
[66] Vgl. DIN EN 13306:2018-02. Instandhaltung - Begriffe der Instandhaltung. Seite 9.
[67] Vgl. DIN EN 13306:2018-02. Instandhaltung - Begriffe der Instandhaltung. Anlage B. Seite 66.

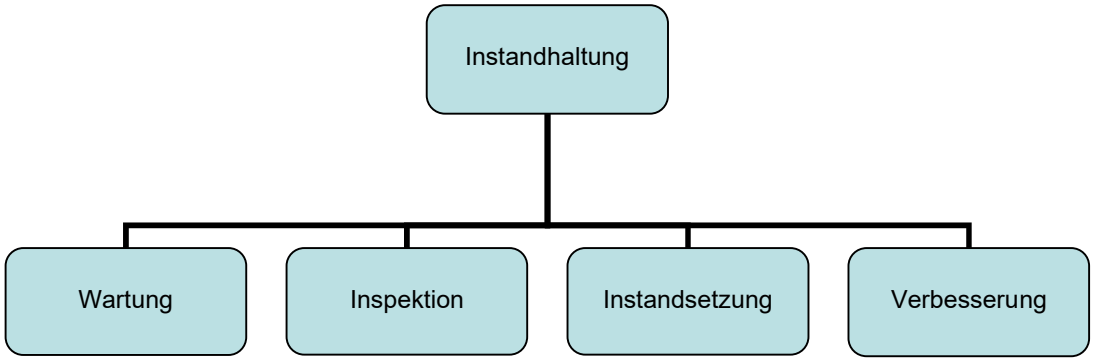

Abbildung 7: Unterteilung der Instandhaltung[68]

Die **Wartung** einer Anlage umfasst alle Maßnahmen zur Verzögerung des Abbaus des vorhandenen Abnutzungsvorrats.[69] Unter Abnutzungsvorrat ist z.B. die Dicke der Gummischicht von Autoreifen oder die Dicke von Bremsbelägen zu verstehen und im Allgemeinen als der Vorrat der Funktionserfüllung eines Objektes definiert.[70] Die wartungsumfassenden Maßnahmen reichen in dem Fall vom Nachfüllen, Schmieren und Ölen, bis hin zu Kontrollen vor Beginn der Arbeit von z.B. der Schutzausrüstung oder den vorbereiteten Bauteilen.[71]

Wenn die maßgeblichen Merkmale eines Objektes, durch Messung, Beobachtung oder Funktionskontrolle auf ihre Konformität geprüft werden, ist die Rede von einer **Inspektion**. Die Besonderheit von Inspektionen im Vergleich zu Wartungen ist die Bewertung des Istzustandes des betrachteten Objektes, die Analyse der Ergebnisse und das Aufzeigen von Lösungsmöglichkeiten bzw. Lösungsalternativen im Rahmen der betrieblichen und gesetzlichen Anforderungen.[72] Eins der bekanntesten Beispiele für Inspektionen, sind die Autoinspektionen. Hierbei wird zunächst der Zustand des Autos untersucht und bewertet. Die Zustände von sich abnutzenden Objekten werden mit den Abnutzungsgrenzen verglichen und beim Überschreiten dieser, bleibt nur die Option des Austauschs.

Bei der **Instandsetzung** hingegen ist die Funktion eines Objektes fehlerhaft, was zur Folge hat, dass diese durch physische Maßnahmen wiederhergestellt wird. Im alltäglichen Sprachgebrauch ist in dem Fall die Rede von einer Reparatur. Wenn eine fehlerhafte Funktion festgestellt wird, muss zunächst die genaue Fehlerursache ermittelt werden[73], bevor das fehlerhafte Teil bzw. die fehlerhaften Teile ausgetauscht werden. Um weiterhin beim Beispiel des Autos zu bleiben, lässt sich eine Instandsetzung bei diesem beispielhaft als Erneuerung einer kaputten Scheinwerferglühbirne oder der Austausch eines Reifens nach dem Platzen sein. Relevant für eine Instandsetzung ist, dass das kaputte Teil, die Lebensdauer komplett aufgebraucht hat und ein Austausch zwingend erfolgt.

[68] Vgl. DIN 31051:2019-06. Grundlagen der Instandhaltung. Seite 12. Bild 2.
[69] DIN 31051:2019-06. Grundlagen der Instandhaltung. Seite 5.
[70] DIN 31051:2019-06. Grundlagen der Instandhaltung. Seite 8.
[71] DIN 31051:2019-06. Grundlagen der Instandhaltung. Seite 5.
[72] DIN 31051:2019-06. Grundlagen der Instandhaltung. Seite 5f.
[73] DIN 31051:2019-06. Grundlagen der Instandhaltung. Seite 6.

Der vierte Bereich der Instandhaltung ist die **Verbesserung**, hierbei handelt es sich um die Kombination aller Maßnahmen, sowohl administrativ und technisch, als auch um alle Maßnahmen des Managements, welche eine Steigerung der Zuverlässigkeit, der Instandhaltbarkeit und/ oder der Sicherheit des Objektes, ohne dass die Funktion beeinträchtigt oder verändert wird, mit sich bringt.[74] Eine Verbesserung am Automobil ist der Austausch der Bremsen, bestehend aus Bremsschreiben und Bremsbelägen, bei dem hochwertigere Bremsen eingebaut werden. Diese neuen Bremsen sollen durch bessere Be-/Entlüftung höhere Temperaturen aushalten können, wodurch der Autofahrer mit kürzeren Bremswegen einem Unfall besser ausweichen kann.

3.3 Entwicklung von Instandhaltung

Bis Anfang der 50er Jahre gab es in Europa nur die reaktive Instandhaltung. Es wurde erst immer im Notfall reagiert, nach dem sogenannten Feuerwehrprinzip. Die Notfälle reichten von Stillständen der Anlagen, über die Zerstörung der Anlage bis hin zu menschlichen Schäden. Nach und nach wurden andere Instandhaltungsstrategien aus der japanischen Automobilindustrie übernommen, beginnend mit der vorbeugenden Instandhaltung, bei der kritische Bauteile frühzeitig, präventiv ausgewechselt wurden, wodurch es eine Reduktion der Stillstände gab. Einige Jahre später folgte die verbessernde Instandhaltung, die das Ziel verfolgte, bestehende Anlagen zu optimieren, sowohl in ihrer Leistung als auch in ihrer Zuverlässigkeit. Ab dem Jahr 1960 wurde die instandhaltungsgerechte Konstruktion für die Industrie relevanter, da es nun möglich war Anlagen leichter zu warten und instand zu setzen. Dieser Aspekt wurde ab dieser Zeit nun auch bei der Neuanschaffung von Anlagen berücksichtigt, es ließ sich Aufwand sparen, trotz höherer Verfügbarkeit und größerem Output.[75]

Es besteht zudem die Möglichkeit Instandhaltung individuell an die Produktion anzupassen. Das bedeutet die Kopplung der Instandhaltung mit der Produktion, wie beispielsweise erstmals bei Toyota.[76] Die Instandhaltungsaufgaben können jedoch auch umverteilt werden, wie bei der autonomen Instandhaltung. Eine autonome Instandhaltung liegt vor, wenn sich der Anlagenbediener neben dem Bedienen, Befüllen und Überwachen der Anlage nun auch um die Instandhaltung dieser kümmert. Das Gegenteil von autonomer Instandhaltung ist die gemeinschaftliche Instandhaltung, bei der sich der Maschinenbediener und die Instandhaltung die Verantwortung für die Anlage teilen.[77]

Heutzutage machen Instandhaltungskosten Studien zu Folge immer noch einen Produktionskostenanteil von 15 bis 70 Prozent je nach Branche aus. Das liegt unter anderem auch daran, dass die Industrie sich immer noch zu sehr auf die reaktive bzw. die falsche Instandhaltungsstrategie verlässt, die Folge ist die verringerte Produktionszeit und nachlassende Produktqualität. Ein gutes Beispiel dafür ist die Industrie in den USA, hier

[74] DIN 31051:2019-06. Grundlagen der Instandhaltung. Seite 6.
[75] Vgl. Heller T. & Prasse C. (2017). Total Productive Management. Springer Verlag. Seite 4.
[76] Vgl. Heller T. & Prasse C. (2017). Total Productive Management. Springer Verlag. Seite 5.
[77] Vgl. Heller T. & Prasse C. (2017). Total Productive Management. Springer Verlag. Seite 14.

werden im Schnitt von jedem Dollar, der in der Instandhaltung ausgegeben wird, 33 Cent verschwendet, aufgrund von unnötiger Wartungsarbeit.[78] Folglich ist eindeutig zu erkennen, dass das Thema Instandhaltung für die Unternehmen sehr bedeutsam ist und durch die richtige Optimierung und Anpassung noch einiges mehr möglich ist, was die Kosteneinsparung und hohe Verfügbarkeit betrifft.

3.4 Was für Instandhaltungsstrategien gibt es?

Aber es bleibt die Frage, ob es in jedem Unternehmen die gleiche Instandhaltungsstrategie gibt. Für die Beantwortung der Frage kann das folgende Kapitel dienen.

Zunächst einmal sollten die generellen Instandhaltungsstrategien betrachtet werden. Dazu kann die Abbildung 8 herangezogen werden. Hier werden vier unterschiedliche Strategien dargestellt, die „korrektive Instandhaltung“, „Redundanzen“, die „vorbeugende Instandhaltung“ und die „zustandsorientierte Instandhaltung“. Eine Besonderheit hierbei sind die Redundanzen, diese sind eine Mischung aus korrektiver und vorbeugender Instandhaltung und werden besonders in der Konstruktion und Planung verwendet. Im Detail wird in Kapitel 3.4.2 auf das Thema Redundanzen eingegangen.

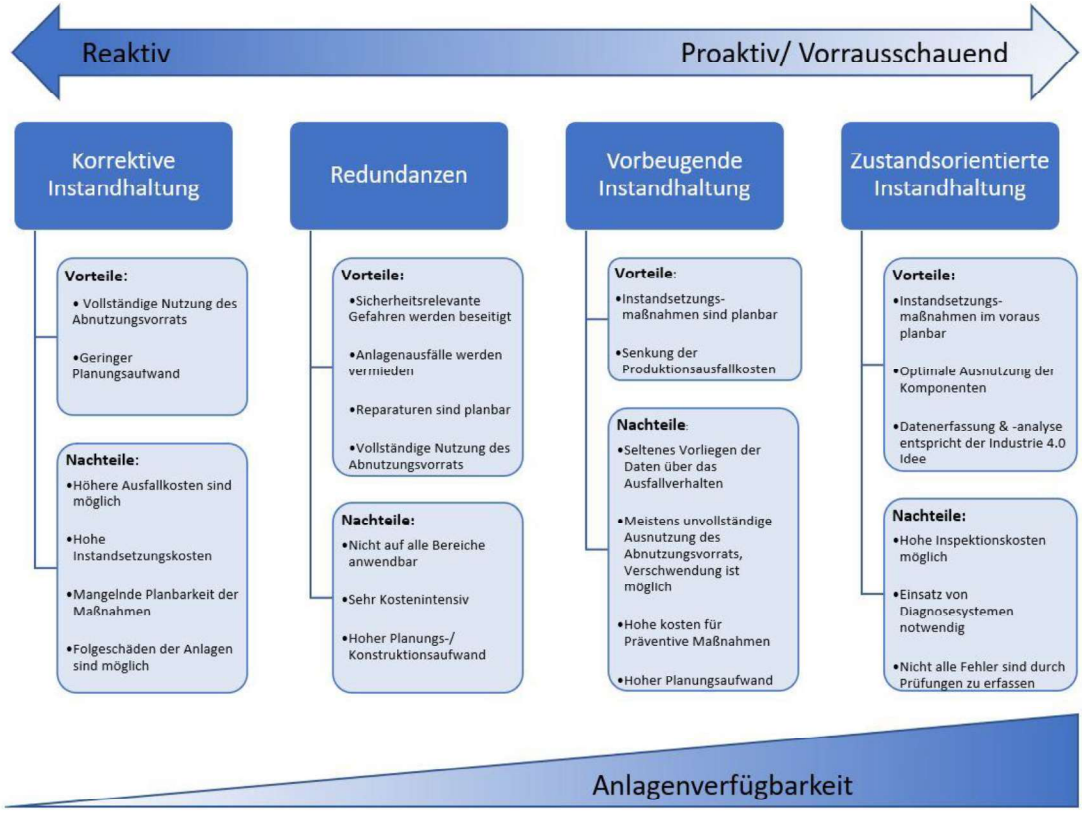

Abbildung 8: Instandhaltungsstrategien [79;80]

[78] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 1.
[79] Vgl. Heller T. & Prasse C. (2017). Total Productive Management. Springer Verlag. Seite.123 Abbildung 7.1.

3.4.1 Korrektive Instandhaltungsstrategie

Eine weit verbreitete Instandhaltungsstrategie ist die Korrektive oder auch Reaktive Instandhaltung. Hierbei wird erst reagiert, wenn Teile ausgefallen bzw. fehlerhaft sind, die Funktion von der Maschine nicht mehr erfüllt werden kann und eine Fehlererkennung ausgeführt wurde.[81] Trotzdem werden hierbei auch regelmäßig Wartungen durchgeführt, um die Ausfälle noch weiter in die Zukunft zu verschieben. Dennoch bestimmt die Maschine in gewisser Weise selbst, wann sie ausfällt und stillsteht.[82] Es wird daher auch nach dem Konzept „If it ain't broke, don't fix it" bzw. Run-to-Failure vorgegangen.[83]

Die reaktive Instandhaltungsstrategie bietet dem Unternehmen einige Vorteile, wie zum Beispiel durch die vollständige Ab-/Ausnutzung einer Komponente, hier wird der Abnutzungsvorrat zu 100% genutzt. Was aus Kostensicht optimal ist. Außerdem ist durch den unklaren Ausfallzeitpunkt, keine besondere Planung notwendig. Es wird einfach instandgesetzt, nachdem die Anlage ausgefallen ist. Ein Nachteil der plötzlichen Ausfälle der Anlage ist, dass zu keinem Zeitpunkt eine Anlagenverfügbarkeit garantiert werden kann, wodurch die Planung für die Produktion und den Versand deutlich erschwert wird. Weitere negative Auswirkungen können bei ungeplanten Ausfällen nicht ausgeschlossen werden, wie z.B. dass durch den Ausfall eines Bauteils Folgeschäden an der Anlage entstehen, wodurch die Instandsetzungskosten und Ausfallkosten, durch lange Standzeiten sehr stark ansteigen können. Hinzu kommt der Nachteil, dass die Planbarkeit der Instandsetzung nicht möglich ist. Wenn z.B. in der Produktion sieben Anlagen laufen und nur zwei Instandhaltungsmitarbeiter tätig sind, kann es, auch wenn es unwahrscheinlich ist, sein, dass mehr als 2 Anlagen gleichzeitig ausfallen und sich die Standzeiten deutlich verlängern.[84]

Bei den Vor- und Nachteilen der Strategie stellt sich final dennoch die Frage nach dem Einsatzgebiet der reaktiven Instandhaltungsstrategie. Oftmals sind der Sicherheits- und der Kostenfaktor hier die Antwort, ob die Strategie geeignet ist oder nicht. Bei dem Auto zum Beispiel werden die Scheinwerferglühbirnen erst bei einem Ausfall ausgetauscht, anders hingegen die Bremsen oder Autoreifen. Wenn die Bremsen oder Reifen ausfallen, besteht im Vergleich zu dem Glühbirnenausfall ein erhöhtes Risiko für einen Unfall mit Sach- und Personenschaden. Ebenso ist reaktive Instandhaltung für die Luftfahrt oder Kernenergie nicht geeignet, da die Folgen und Kosten bei einem Aus- bzw. Unfall immens ausfallen können. Wenn aber eine Industrieanlage aus Teilen besteht die sehr teuer, aber leicht zu wechseln sind, kann bei einem Ausfall das fehlerhafte Teil schnell und ohne großen Aufwand ausgetauscht werden und es wird kein Geld durch ungenutztes Potential des Bauteils verschenkt.[85]

---

[80] Vgl. Tordeux A. (o.D.). Lehrvortrag: Zuverlässigkeit von Kompressoren in Produktionsanlagen; Folie 16.
[81] Vgl. DIN EN 13306:2018-02. Instandhaltung - Begriffe der Instandhaltung. Seite 38.
[82] Vgl. Heller T. & Prasse C. (2017). Total Productive Management. Springer Verlag. Seite 123 f.
[83] Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 2.
[84] Vgl. Heller T. & Prasse C. (2017). Total Productive Management. Springer Verlag. Seite 122 f.
[85] Vgl. Heller T. & Prasse C. (2017). Total Productive Management. Springer Verlag. Seite 123 f.

3.4.2 Redundanzen

Die nächstaktivere Instandhaltungsoption ist die sogenannte Redundanz. Wie in Kapitel 3.4 schon aufgegriffen wurde, handelt es sich bei Redundanzen nicht direkt um eine Strategie, sondern um eine Anwendungsform. Oftmals werden Redundanzen in der Konstruktion angewendet, da eine Redundanz das Vorhandensein von mehr als einem Mittel zur Erfüllung der geforderten Funktion ist. Dennoch kann zwischen zwei unterschiedlichen Redundanzen differenziert werden. Zum einen gibt es aktive oder auch funktionsbeteiligte Redundanzen (siehe Abbildung 9), bei denen mehrere Funktionen gleichzeitig in Betrieb sind, um die geforderte Funktion zu erfüllen. Beispielweise können hierzu parallellaufende Systeme, paarige LKW-Reifen oder mehrere, voneinander unabhängige Stromquellen gezählt werden. Zum anderen gibt es die passiven und nicht funktionsbeteiligten Redundanzen, wie in Abbildung 10 dargestellt. Eine passive Redundanz ist eine Redundanz, bei der ein alternatives Mittel zur Erfüllung einer bestimmten Funktion nur dann in Betrieb genommen wird, wenn das aktive Mittel nicht verfügbar ist.[86] Ein Beispiel für passive Redundanz kann ein Reserverad oder auch ein Überdruckventil sein, welches aktiviert wird, wenn z.B. der Kühlkreislauf ausfällt.

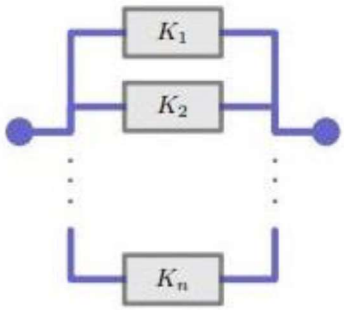

Abbildung 9: Aufbau aktive Redundanz[87]

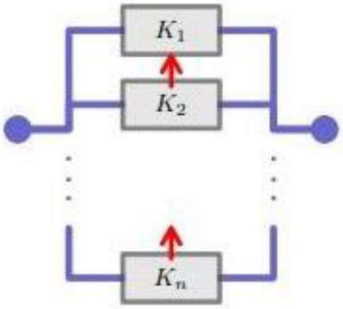

Abbildung 10: Aufbau passive Redundanz[88]

Zusammenfassend lässt sich sagen, dass Redundanzen besonders dafür geeignet sind die Verfügbarkeit und Zuverlässigkeit von Anlagen zu erhöhen, da Aus- und Unfälle egalisiert bzw. gedämpft werden.[89] Des Weiteren ist das vollständige Ausnutzen des Bauteilpotentials sehr bedeutsam und kann je nach Redundanz Kosten einsparen. Aktive Redundanzen hingegen können sehr kostenintensiv aufgrund ihrer Ineffizienz sein. Ein weiterer Nachteil ist, dass eine Redundanz nicht auf alle Bereiche angewendet werden kann, da es nicht möglich ist zwei Anlagen nebeneinander zu stellen und die zweite erst eingeschaltet wird, wenn die erste ausfällt. Außerdem ist es möglich, dass die Konstruktion bzw. Planung sehr aufwendig und kompliziert werden kann. Dennoch bieten Redundanzen die Möglichkeit Ausfälle und Stillstände zu verhindern und die Instandhaltung somit zu unterstützen.

[86] Vgl. DIN EN 13306:2018-02. Instandhaltung - Begriffe der Instandhaltung. Seite 21.
[87] Tordeux A. (2019). Vorlesung: Technische Zuverlässigkeit; Veranstaltung: 6. Folie 19.
[88] Tordeux A. (2019). Vorlesung: Technische Zuverlässigkeit. Veranstaltung: 6. Folie 21.
[89] Vgl. Tordeux A. (2019). Vorlesung: Technische Zuverlässigkeit. Veranstaltung: 1. Folie 17.

3.4.3 Vorbeugende Instandhaltung

Der nächste Schritt in Richtung proaktive Instandhaltung ist die vorbeugende Instandhaltung, welche sich normalerweise durch den hohen Grad an Planung der Maßnahmen auszeichnet.[90] Das Ziel von vorbeugender Instandhaltung ist es, durch die Beurteilung und Aufrechterhaltung des Zustandes der Anlage, dessen Ausfallzeiten zu reduzieren. Diese Maßnahmen können periodisch oder in gewisser Weise zustandsabhängig durchgeführt werden.[91] Aus diesem Grund ist die Planung der Maßnahmen sehr wichtig und muss immer auf dem aktuellen Stand sein, da sonst Maßnahmen doppelt oder verspätet durchgeführt werden. Dennoch muss die Produktion gestoppt werden, um die Instandhaltungsmaßnahmen durchführen zu können. Oftmals wird in Form von Inspektionen oder Wartungen vorgegangen, bei der die Bauteile kontrolliert, repariert oder ausgetauscht werden. Der Vorteil der vorbeugenden Instandhaltung ist die verbesserte Verfügbarkeit von Anlagen und das reduzierte Risiko von Ausfällen. Außerdem können die Kosten für Produktionsausfälle und Instandsetzung deutlich gesenkt werden, was auf den ersten Blick sehr positiv ist, dennoch steigen die Kosten für die frühzeitig gewechselten Ersatzteile und die Planung.[92,93]

Besonders effektiv ist, wenn vorbeugende Instandhaltungsmaßnahmen entweder direkt bei der Rüstung der Anlage oder bei ungeplanten Stillständen spontan durchgeführt werden, da somit nicht noch extra Stillstände entstehen. Beispielhaft können hierbei die stehenden Anlagen von „freien“ Mitarbeitern gewartet und deren Zustand beurteilt werden. Außerdem ist es möglich die Anlage oder bestimmte Bereiche zu säubern oder zu schmieren. Eben solche präventiven Maßnahmen sind jene, die anlagenabhängig auch während des Betriebs durchgeführt werden können und das unter anderem auch in einem periodischen Rhythmus.[94]

Bei Bauteilen mit hohen Lebensdauern und keinen kritischen Konsequenzen bei einem Ausfall lohnt sich die präventive Instandhaltung nicht. Zum einen aus Kostengründen, da die Bauteile nicht vollständig abgenutzt werden, zum anderen sind die Personalkosten dafür zu hoch. Wenn aber die Einsparungen an Stillstandzeiten und Reparaturkosten durch die vorbeugende Instandhaltung die Aufwandskosten übersteigen, sollte der präventive Austausch von Bauteilen in Betracht gezogen werden. Ein weiterer Bereich, in dem die präventive Instandhaltung relevant ist, ist der der Sicherheit. Ein geeignetes Beispiel hierfür sind Flugzeuge. Kein Passagier würde in ein Flugzeug steigen, wenn die Turbine, die Reifen oder ein Flügel während des Fluges versagen könnten.[95]

[90] Vgl. MaintMaster Systems GmbH (o.D.). Präventive und Korrektive Instandhaltung.
[91] Vgl. DIN EN 13306:2018-02. Instandhaltung - Begriffe der Instandhaltung. Seite 34 + 35.
[92] Vgl. MaintMaster Systems GmbH (o.D.). Präventive und Korrektive Instandhaltung.
[93] Vgl. Heller T. & Prasse C. (2017). Total Productive Management. Springer Verlag. Seite 123. Abbildung 7.1.
[94] Vgl. Heller T. & Prasse C. (2017). Total Productive Management. Springer Verlag. Seite 7 +122.
[95] Vgl. Heller T. & Prasse C. (2017). Total Productive Management. Springer Verlag. Seite 122, 124 + 129.

3.4.4 Zustandsorientierte Instandhaltung

Die letzte der vier Instandhaltungsstrategien aus der Abbildung 8 ist die zustandsorientierte Instandhaltungsstrategie, welche dazu auch die modernste ist. Zustandsorientierte Instandhaltung ist eine abgewandelte Form von präventiver Instandhaltung, mit der Besonderheit, dass der physische Zustand des Objektes überwacht, analysiert und bewertet wird und wenn nötig auch Instandhaltungsmaßnahmen durchgeführt oder geplant werden. Eine Erweiterung der zustandsorientierten Instandhaltung ist die vorausschauende Instandhaltung, hierbei ist es möglich, durch die Analyse von Daten und Parametern eine Vorhersage zu treffen wann und wie welche Eigenschaften des Objektes nachlassen, bis es zum Ausfall kommt.[96] Wenn von zustandsorientierter und/oder vorausschauender Instandhaltung die Rede ist, können diese auch mit prädiktiver Instandhaltung Umschrieben werden. Die Basis, um diese Instandhaltung erfolgreich durchführen zu können, bildet in den meisten Fällen das Condition Monitoring. Durch den Einsatz von hochentwickelten und genauen Sensoren kombiniert mit fortschrittlicher Prognosetechnik wird diese Strategie ermöglicht, was besonders dem Denken von Industrie 4.0 [97] mit zunehmender Digitalisierung entspricht.

Das Ziel was mit der zustandsorientierten Instandhaltung einher geht, ist die Steigerung der Effizienz der Produktionsanlagen. Diese Effizienzsteigerung kann sowohl durch die Reduktion von überflüssigen Wartungsaktivitäten als auch von Produktionsstillständen und Ausfällen erreicht werden. Somit ist das genaue Abschätzen der Restnutzungsdauer entscheidend, wie erfolgreich diese Strategie im Unternehmen ist. Je genauer die RUL prognostiziert werden kann, desto eher können zu frühe und zu späte Instandhaltungs-aktivitäten vermieden werden.[98] Die RUL Prognose kann nur erfolgreich durchgeführt werden, wenn die Sensoren genaue und verlässliche Daten ausgeben und somit ein Mindestmaß an Qualität aufweisen. Bei einigen Anlagen muss die Datenerfassung für die Zustandsüberwachung aber nicht über Sensoren laufen, sondern ist auch manuell möglich. Problematisch wird hierbei nur, dass es sich bei der manuellen Überwachung/ Wartung oft nur um eine Momentaufnahme handelt. Es kann vorkommen, dass in diesem Moment der oder die Überwachungsparameter einen positiven oder negativen Ausreißer haben.[99] „Der Ansatz erlaubt es nicht, Wartungsaktivitäten effizient im Voraus zu planen“[100], da der Ansatz das Problem hat, dass die Daten im Verhältnis zu kontinuierlicher Überwachung zu wenige und dazu dann auch noch sehr anfällig sind.[101]

Im Folgenden wird ein Framework für prädiktive Instandhaltung näher betrachtet. 2020 wurde dies unteranderem von Christian Krupitzer und Tim Wagenhals veröffentlicht. Dafür wurden während der Studie 150 andere Studien und Veröffentlichungen aus dem Zeitraum 1993 bis 2018, mit den Gesichtspunkten auf der Umsetzung von prädiktiver Instandhaltung,

[96] Vgl. DIN EN 13306:2018-02. Instandhaltung - Begriffe der Instandhaltung. Seite 35.
[97] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 1.
[98] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 1.
[99] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 2.
[100] Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 2.
[101] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 6.

zustandsorientiert und vorausschauend, analysiert.[102] Hierbei gabs es viele unterschiedliche Herangehensweisen, Umsetzungen und Techniken. In der Analyse konnte herausgearbeitet werden, dass die Veröffentlichungen aus zehn Bestandteilen bestehen, welche zu dem anwenderfreundlichen Framework ausgearbeitet wurden. Diese zehn Bestandteile, geben der Entwicklung bzw. Betrachtung von prädiktiver Instandhaltung einen Rahmen. Wobei nicht alle Bestandteile in jeder Veröffentlichung beinhaltet sind.[103] Somit lässt sich sagen, dass es wichtigere Bereiche gibt und tendenziell eher zweitrangige.

Im Folgenden wird dieses Framework, mit seinen Bestandteilen, etwas näher betrachtet, um das Vorgehen bei Kapitel 4 verständlich zu machen. Es ist möglich das Framework komplett theoretisch zu planen, es kann aber auch auf ein bestehendes System angewendet werden, was bedeutet, dass einige der Bestandteile des Frameworks schon vorgegeben sind.

In Abbildung 11 werden die zehn Bauteile des Frameworks aufgezeigt, dabei handelt es sich um die Ziele, das Condition Monitoring, den Umfang der Instandhaltung, die Ausfallkurven, die Fehlererkennung, die Systemgröße, die Terminierung, die Prognosetechniken, die Datenhandhabung und die Evaluation.[104]

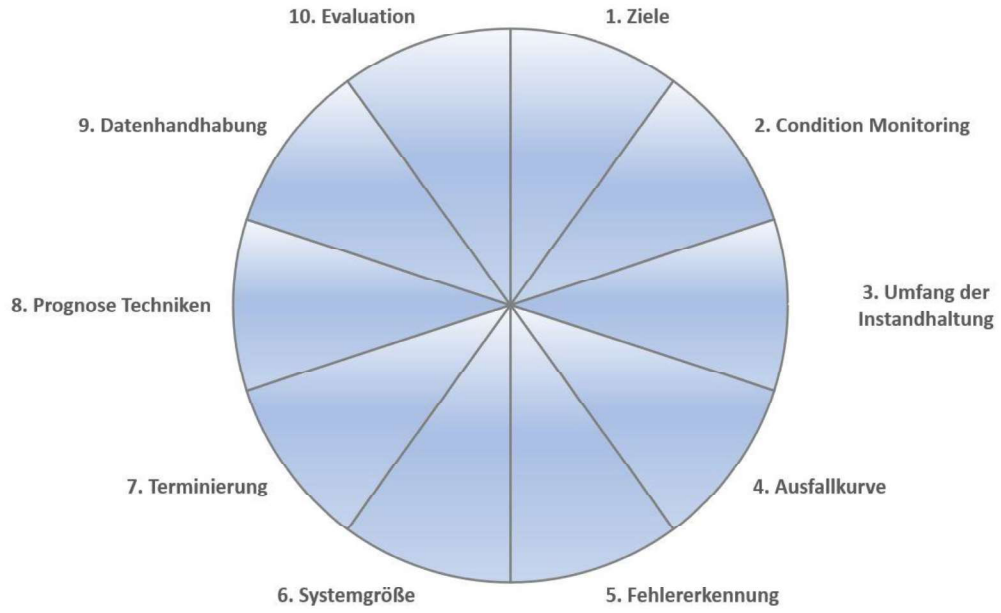

Abbildung 11: Framework für prädiktive Instandhaltung, 10 Bestandteile[105]

Zu Beginn jeder Entwicklung einer prädiktiven Instandhaltung müssen **Ziele** beschrieben werden, damit alle Beteiligten wissen in welche Richtung es gehen soll. Das Thema Ziele ist eins der relevanten Bausteine, was u.a. daran zu erkennen ist, dass in 139 der 150 Dokumente

[102] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 3.
[103] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 4.
[104] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 4.
[105] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 4.

das Thema Ziele behandelt wurde. Die Ziele unterteilen sich in acht mögliche Bereiche, wobei meistens der Aspekt des Geldsparens im Mittelpunkt steht.[106] Wenn Ziele definiert werden, lassen diese sich mindestens einem der folgenden Punkte zuordnen:

- Ersatzteilbestandsreduzierung
- Verlängern der Komponentenlebensdauer
- Kostenminimierung
- Ausfallzeiten reduzieren
- Verfügbarkeit erhöhen
- Produktivität erhöhen
- Zuverlässigkeit erhöhen
- Sicherheit erhöhen [107]

Wenn Ziele festgelegt werden, kann im weiteren Verlauf leichter ein Vorgehen entwickelt werden, wodurch ein erfolgreicherer Ausgang ermöglicht wird. Gut definierte Ziele werden nach dem SMART-Modell erarbeitet. Smart steht für speziell, messbar, attraktiv, realistisch und terminiert.[108] Wenn die Ziele nach SMART erstellt werden, bestehen diese auch aus Kennzahlen. Und durch Kennzahlen können Ziele besser dargestellt und der aktuelle Wert mit dem Zielwert verglichen werden.[109]

Der zweite Bestandteil des Frameworks ist die **Zustandsüberwachung** oder auch Condition Monitoring. Unter diesem Punkt wird das Sammeln von Daten über den Gesundheitszustand der Maschine beschrieben. Durch das Sammeln der Maschinendaten kann es ermöglicht werden die RUL zu prognostizieren oder zu schätzen. Doch wie genau das Condition Monitoring aufgebaut wird ist nicht genau definiert, es gibt dennoch vier unterschiedliche Möglichkeiten:

- Inspektionsbasierte Überwachung
- Sensorbasierte Überwachung
- Kontinuierliche Überwachung
- Onlineüberwachung[110]

Diese Überwachungsoptionen unterscheiden sich alle hauptsächlich durch zwei Punkte, zum einen durch den Aufenthaltsort des Überwachenden, zum anderen durch die Überwachungsintervalle. Die inspektionsbasierte Überwachung wird vom Instandhaltungspersonal, während der Wartung/ Inspektion an der Anlage durchgeführt, welche aber an den Zustand der Anlage angepasst werden. Die inspektionsbasierte Überwachung kann manuell als auch sensorgestützt durchgeführt werden. Eine der häufigsten Überwachungsmethoden ist die sensorbasierte Überwachung, denn hierbei werden Daten von den verschiedensten Sensoren für unterschiedliche Parameter gesammelt. Die sensorbasierte Überwachung bildet die Basis für die meisten Zustandsüberwachungsmethoden und wird durch die Entwicklung der Sensortechnik immer gefragter. Bei der sensorbasierten Überwachung können Intervalle angegeben werden, sie kann automatisch, auf Abruf oder dauerhaft durchgeführt werden. Der

[106] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 5 + 6.
[107] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 5.
[108] Vgl. Niermeyer R. (2007). Motivation: Instrumente zur Führung und Verführung. Haufe. Seite 80.
[109] Vgl. Niermeyer R. (2007). Motivation: Instrumente zur Führung und Verführung. Haufe. Seite 80.
[110] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 6.

Ort des Überwachenden hängt dann an der Verarbeitung und Darstellung der Daten. Bei der kontinuierlichen Überwachung ist die Sammlung der Daten mit folgender Analyse des Maschinenzustandes und des RULs maßgebend. Es gibt keine periodischen Messungen, da die Daten von Sensoren kontinuierlich überwacht, gesammelt und beispielsweise an ein Auswertungstool weitergeleitet werden. Die Grundlage für die kontinuierliche Überwachung ist die Online- oder Echtzeitüberwachung, da dadurch die Datenerfassung im laufenden Zustand möglich ist und somit auch eine Beschreibung der Gesundheit der Anlage.[111]

Bei Bestandteil drei handelt es sich um den **Wartungsumfang**, welches nur in 20 von den 150 analysierten Dokumenten vorgekommen ist und somit weniger oft behandelt wird. Bei dem Wartungsumfang wird zwischen drei Bereichen unterschieden:

- Perfekte Instandhaltung
- Unvollkommene Instandhaltung
- Gruppierte Instandhaltung[112]

Wenn von der perfekten Instandhaltung die Rede ist, wird die Anlage wieder in einen Neuzustand versetzt und müsste theoretisch wieder die maximale Lebensdauer und Funktion haben. Die Annahme hierbei ist, dass jedes Bauteil neuwertig ist. Wenn aber eine Instandhaltung unvollkommen ist, erreicht die Anlage wieder einen guten bis sehr guten gebrauchten Zustand. Es wird dabei davon ausgegangen, dass die Anlage noch gebrauchte Teile beinhaltet. Diese Annahme ist im Vergleich zu der perfekten Instandhaltung realistischer, da in den meisten Fällen nicht alle Bauteile einer Anlage ausgewechselt werden. Bei der gruppierten Instandhaltung werden Bauteile in Gruppen zusammengeschlossen, während die gesamte Anlage überwacht wird. Wenn nun eine Komponente aus der Gruppe gewechselt wird, werden alle der Gruppen angehörigen Teile ausgetauscht. Folglich kann der Austausch perfekt und unvollkommen sein, je nach Betrachtungsraum. Der Vorteil der Gruppeninstandhaltung ist die Kosteneinsparung, aufgrund von reduzierten Ausfallzeiten, aber auch nur wenn die gesamte Anlage überwacht und zustandsorientiert instandgehalten wird.[113]

Wenn die **Abbauprozesse** der Bauteile oder Anlagen betrachtet werden, können diese grafisch dargestellt werden. Die Darstellung kann durch unterschiedliche Annahmen oder Funktionen vollführt werden:

- Direkte Abbaumodellierung
- Exponentielle Abbaumodellierung
- Lineare Abbaumodellierung
- Weibullverteilungsabnahme[114]

Bei Betrachtung des Abbauprozesses soll zum einen der aktuelle Zustand der Anlagen grafisch dargestellt werden können und zum anderen soll der Abbauprozess als Basis für die

[111] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 6.
[112] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 7.
[113] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 7.
[114] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 7.

Prognose der RUL dienen. Die Daten für die Darstellung werden von den Sensoren in und an den Anlagen erfasst. Die möglichen Abbauprozesse wurden in den Abbildungen 12 bis 15 beispielhaft dargestellt. Auffällig sind hierbei, dass die Abbauprozesse unterschiedlich schnell und langsam abnehmende Bereiche haben.

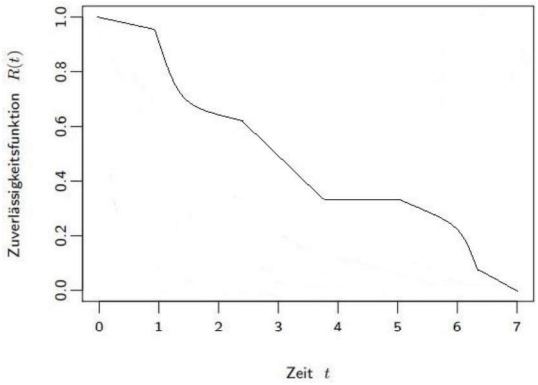

Abbildung 12: Direkte Abbaumodellierung

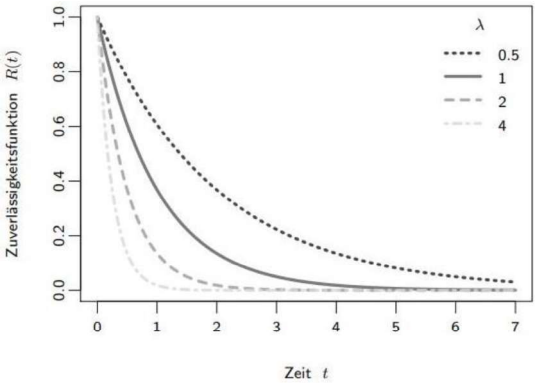

Abbildung 13: Exponentielle Abbaumodellierung[115]

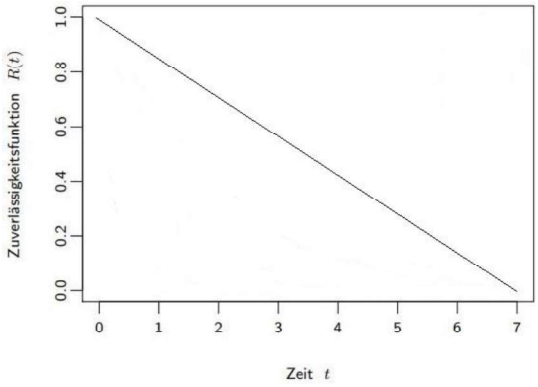

Abbildung 14: Lineare Abbaumodellierung

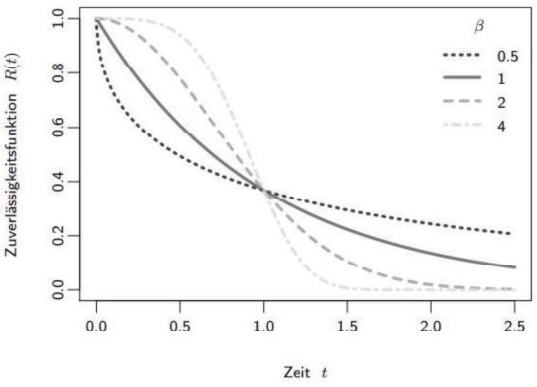

Abbildung 15: Weibullverteilung[116]

Wenn eine Anlage ausfällt, werden in vielen Fällen nur die beschädigten Komponenten ausgetauscht, aber eine detaillierte **Fehlererkennung** wird nicht durchgeführt. Die Aktion der Fehlererkennung wird nur in 36% der analysierten Veröffentlichungen erwähnt. Es gibt die Optionen der

- Hauptursachenanalyse
- Maschinen Diagnostik[117]

[115] Tordeux, A. (2018). Vorlesung: Technische Zuverlässigkeit. Veranstaltung: 2. Folie 21.
[116] Tordeux, A. (2018). Vorlesung: Technische Zuverlässigkeit. Veranstaltung: 2. Folie 28.
[117] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 8.

Abhängig von Unternehmen und Branche wird entweder der Begriff der Hauptursachenanalyse oder der der Maschinen Diagnostik verwendet, inhaltlich ist unter beiden Begriffen dasselbe zu verstehen. Die Maschinendaten werden ausgewertet, um nachträglich den Grund des Ausfalls zu bestimmen. Doch warum kann es für ein Unternehmen interessant sein die Gründe der Ausfälle zu bestimmen? Besonders lehrreich sind die Gründe für die Instandhaltung, diese kann dadurch nicht nur Erfahrungswerte sammeln, sondern können auch Schwachstellen der Anlage identifiziert werden, an denen eine Verbesserung durchgeführt werden muss. Ebenso interessant sind die Ausfallgründe für die zukünftige Instandhaltungsstrategie, eine Anpassung der Strategie kann die Anzahl der Ausfälle reduzieren. Es kann auch vorkommen, dass festgestellt wird, dass die bestehenden Sensoren nicht ausreichend sind und ein paar Wenige neu angeschafft oder bestehende anders ausgerichtet werden müssen. Müssen andere Komponenten überwacht, Maschinenparameter anders gemessen oder genauere Sensoren angebracht werden? Außerdem wirkt sich die Fehlererkennung final auch auf die genauere Abschätzung des RULs der Anlagen und deren Komponenten aus.[118]

Im sechsten Parameter des Frameworks wird die **Systemgröße** der angewendeten zustandsorientierten Instandhaltungsstrategie behandelt. Dieser Parameter ist als relevant einzustufen, da das Thema Systemgröße in 70% der Veröffentlichungen behandelt wurde. In der Analyse der Veröffentlichung wurde folgende Unterteilung festgestellt:

- Einkomponentensystem
- Mehrkomponentensystem
    - Strukturelle Komponentenabhängigkeit
    - Wirtschaftliche Komponentenabhängigkeit
    - Keine Komponentenabhängigkeit
    - Stochastische Komponentenabhängigkeit[119]

Zu einem Einkomponentensystem werden einzelne Komponenten gezählt. Neben den einzelnen Komponenten werden auch einzelne, separierte Maschinen in einer Einkomponentenanalyse untersucht. Das bedeutet, dass unter einem Mehrkomponentensystem mehrere Maschinen oder einzelne Komponenten, die im Zusammenspiel ein großes System bilden oder ein Teil eines großen Systems sind, verstanden werden. Im Vergleich von Einkomponentensystemen zu Mehrkomponentensystemen sind die Mehrkomponentensysteme deutlich schwerer vorrausschauend bzw. zustandsorientiert instandzuhalten, was sich nicht nur mit der Menge an Daten erklären lässt, sondern auch mit den möglichen Abhängigkeiten innerhalb des Systems. Wenn die Abhängigkeiten mit hineinbezogen werden, unterteilen sich diese in strukturelle, wirtschaftliche und stochastische Abhängigkeit. Bei der wirtschaftlichen Abhängigkeit wird oft eine Kostenreduktion angestrebt, da Ausfallzeiten z.B. durch die Instandsetzung mehrerer Komponenten gleichzeitig effizienter genutzt werden können. Wenn aber die stochastische Abhängigkeit betrachtet wird, werden Komponenten gruppiert, die sich und ihren Abbauprozess beeinflussen. Strukturelle Abhängigkeit bei Mehrkomponenten-systemen hat zur Folge, dass die Komponenten, die zusammenwirken, eine Einheit bilden, in

[118] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 8.
[119] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 8 + 9.

den meisten Fällen zusammen gewartet oder repariert werden. Die Abhängigkeiten können auch kombiniert auftreten, sodass es mehrere Gruppierungen in einem System gibt, in denen Komponenten mehrfach vorkommen.[120]

Ein ebenfalls relevanter Parameter ist die **Zeitplanung**. Logischer Weise ist die Zeitplanung relevant für die zustandsorientierte Instandhaltungsstrategie, da hierbei versucht wird die Wartungs- bzw. Instandhaltungszeitpunkte so effizient wie möglich zu organisieren. Dabei wird in drei Optionen unterteilt:

- Dynamische Maßnahmenplanung
- Dynamische Ersatzteilverfügbarkeit
- Keine Zeitplanung[121]

Der Vorteil der dynamischen Maßnahmenplanung ist die Anpassbarkeit des Wartungsplans an die Daten der zustandsüberwachenden Sensoren. Diese Planung ist aber nur möglich, wenn der zukünftige Maschinenzustand vorhergesagt werden kann. Die Wartungspläne können mit Algorithmen erstellt werden, damit diese kosteneffizient sind und an die neuen Maschinenzustände angepasst werden können. Wenn aber von einer dynamischen Ersatzteilverfügbarkeit die Rede ist, wird sowohl von den Instandhaltungsmaßnahmen gesprochen als auch von der erforderlichen Nachbestellung von Ersatzteilen, damit im Lager am besten immer ein Ersatzteil vorhanden ist, um die Maschinenausfälle relativ kurz zu halten. Diese Herangehensweise, der starken Vernetzung in dem Unternehmen, entspricht den Werten und Merkmalen der Smart Factory und dem Ansatz der Industrie 4.0.[122]

Nun nachdem die Grundlagen soweit abgesteckt wurden, behandelt der folgende Bereich einen der wichtigsten Punkte bei vorausschauender, zustandsorientierter Instandhaltung. Da die gesamte Strategie komplett Datengesteuert ist, sind die **Prognose Techniken** für die Maschinendaten elementar. Dabei unterscheiden sich die Prognose Techniken in zwei Bereiche, zum einen gibt es die datengesteuerten Ansätze und zum anderen die modellbasierten Ansätze. Das Ziel ist es aus den gewonnenen und gemessenen Daten wertvolle und relevante Informationen zu gewinnen, welcher Ansatz dafür am besten geeignet ist, muss immer erst getestet oder ermittelt werden. Es gibt sehr viele unterschiedliche Umsetzungsmöglichkeiten je Ansatz, daher werden im Folgenden vor allem nur die Ansätze aufgezählt, die auch in der Datenanalyse in Kapitel 4 genutzt und dort näher beschrieben werden.

- Datenbasierten Ansätze
    - Lineare Regression
    - Quadratische Regression
    - Exponentielle Regression
    - Random Forest

[120] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 9.
[121] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 9.
[122] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 9 + 10.

    - AI basierte Ansätze
        - Support-Vector-Machine
        - Neuronale Netzwerke
        - K-Nearest-Neighbours
    - Hidden-Markov-Model
- Modellbasierter Ansatz
    - Generische Algorithmen
    - Partikel Filterung
- Kein spezieller prognostischer Ansatz[123]

Der Unterschied zwischen den modellbasierten Ansätzen und den datenbasierten Ansätzen ist die Menge der benötigten Trainingsdaten. Die modellbasierten Ansätze benötigen wenige Trainingsdaten, um in der Praxis eingesetzt werden zu können, bei den datenbasierten hingegen werden große Mengen an Trainingsdaten benötigt.[124] Als Trainingsdaten werden Beispiele verstanden, auf dessen Basis die prognostischen Ansätze eigenständig lernen sollen, wie sie zum Beispiel Muster oder Zusammenhänge erkennen und verstehen können.[125] Genau aus diesem Grund werden die modellbasierten Methoden bei neuen, unerprobten Systemen eingesetzt, da diese Systeme nur wenige Daten vorliegen haben und noch nicht vollständig ausgereift sind. [126]

Nachdem die Maschinendaten gemessen wurden, stellt sich die Frage wie damit umgegangen wird. Dieser Punkt wird in dem Baustein **Datenhandhabung** behandelt. Die wichtigsten Aspekte, die darin behandelt werden, sind:

- Datenfusion
- Datenfilterung
- Speicherort
    - Lokal
    - Online in einer Cloud/ Fernspeicherung
- Datenzugriff
    - Lokal
    - Fernzugriff[127]

Die Datenhandhabung ist ein sehr wichtiger und nicht zu vernachlässigender Aspekt der zustandsorientierten Instandhaltungsstrategie, besonders beim Einsatz von kontinuierlicher Erfassung. Durch diese riesigen Datenmengen wird sehr viel Speicher beansprucht. Es ist möglich die Daten zu reduzieren. Eine Möglichkeit ist die Datenfusion, hierbei werden mehrere Datenquellen kombiniert, um zum einen zuverlässige Daten zu erhalten. Folglich erhält man eine fusionierte Wertereihe, anstelle von x-beliebigen. Eine weitere Option der Datenreduktion ist das Filtern von Daten. Beim Filtern ist relevant, welche Daten einen Mehrwert für die Analyse bieten und welche nicht. Wenn z.B. ein Sensor T durchweg dieselben Werte misst, obwohl die Anlage Fehler und Ausfälle hatte, kann diese Wertereihe des Sensors T gelöscht werden. Ein guter Indikator für solche Entscheidungen ist die

[123] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 10.
[124] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 10.
[125] Frenzel S. (Januar 2021). Guide: Machine Learning mit hochwertigen Trainingsdaten. Crowd-Guru.
[126] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 11.
[127] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 11.

Standardabweichung der Messreihe, je mehr die Standardabweichung gegen 0 geht umso weniger relevante Informationen sind in den Daten enthalten. Wenn die Daten nun vorgefiltert oder fusioniert wurden, stellt sich die Frage wie und wo werden die Daten abgespeichert und wer weiterhin Zugriff auf die Daten bekommt. Meistens werden die Daten noch lokal abgespeichert und daher gibt es auch hauptsächlich noch lokalen Datenzugriff. Durch die zunehmende Entwicklung Richtung Industrie 4.0 mit dem Internet der Dinge sollte sich das Verhältnis von hauptsächlich lokaler Speicherung und Zugriffe auf immer mehr Cloudbasierte Speicherung mit Fernzugriffen.[128]

Der letzte Baustein des Frameworks ist die **Evaluation** der Daten bzw. der Analyseergebnisse. Hierbei gibt es unterschiedliche Herangehensweisen, die den Ansatz der vorausschauenden Instandhaltung testen, die Realisierbarkeit beweisen und ggf. auch die Ergebnisse Verifizieren. Folgende Evaluationsmöglichkeiten wurden aufgezählt:

- (Numerische) Simulation
- Bewertung auf Basis der realen Daten
- Experiment
- Vergleich mit herkömmlichen Instandhaltungsstrategien.[129]

Bei der Evaluation ist es möglich, die unterschiedlichen Möglichkeiten zu kombinieren. Bei der Numerischen Simulation werden simulierte Daten bewertet. Die Monte-Carlo-Simulation generiert mögliche Ausfälle, die durch den vorrausschauenden Wartungsansatz identifiziert und verglichen werden müssen. Bei der Bewertung auf Basis von realen Daten wird das Unternehmen mit unterschiedlichsten Unternehmen verglichen, die auch diese Instandhaltungsstrategie verfolgen. Die Experimente werden oft bei Einkomponenten-systemen verwendet, da sie nur in kleinem Umfang möglich sind. Abschließend gibt es den Vergleich zur herkömmlichen Instandhaltung. Allgemein werden bei dem Vergleich hauptsächlich die Kosten und die Anlageneffizienz begutachtet und es wird die Frage beantwortet, ob die zustandsorientierte Instandhaltungsstrategie diese anvisierte Kostens-enkung und eine Steigerung der Effizienz ermöglichen kann oder nicht.[130]

3.4.5 Vergleich der Instandhaltungsstrategien: zustandsorientiert - herkömmlich

Wie in den vorherigen Kapiteln schon näher betrachtet wurde, ist es nicht möglich die beste Instandhaltungsstrategie zu bestimmen. Jede Instandhaltungsstrategie hat seine Berechtigung. Für die unterschiedlichen Einsatzbereiche gibt es besser geeignete Strategien und weniger geeignete Strategien. Um die geeignetste Strategie zu finden, müssen einige Aspekte berücksichtigt werden.

Einer der wichtigsten Aspekte für die Auswahl der geeigneten Instandhaltungsstrategie ist die Bewertung der Kritikalität von bestimmten Ereignissen beim Ausfall. Wenn das Ereignis nicht besonders ist und nur geringe bis keine negativen Auswirkungen hat, reicht eine reaktive

[128] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 11.
[129] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 12.
[130] Vgl. Krupitzer C. et al. (2020). A Survey of Predictive Maintenance for Industry 4.0. Seite 12.

Instandhaltung vollkommen aus. Sollte das Ereignis aber große finanzielle oder personelle Schäden nach sich ziehen muss proaktiv gehandelt werden und ein Ausfall verhindert werden. Die Ermittlung der Kritikalität kann z.B. in einer FMEA mit durch das Bestimmen der Risikoprioritätszahlen durchgeführt werden.

Ein weiterer ebenfalls bereits in den vorherigen Punkten beschriebener Aspekt ist der der Planung. Wie aufwendig und detailliert muss eine Planung für die Anlagen im Unternehmen sein? In welchem Verhältnis steht der Planungsaufwand im Vergleich zu dem Nutzen für das Unternehmen und die Produktion

4. Praktische Datenauswertung eines Turbinendatensatzes

Im folgenden Kapitel wird nun ein Datensatz der NASA von einem Triebwerk analysiert. Die Analyse erfolgt mit der Software R. Über die genaue Datengewinnung sind keine Informationen veröffentlicht, weder über die erfassten Parameter noch über die verwendeten Sensoren. Somit konnte eine reine Datenanalyse durchgeführt werden. Mit der Analyse wurde das Ziel verfolgt, den RUL des Triebwerks mit höchstmöglicher Genauigkeit vorherzusagen. Da keine Referenzwerte vorliegen, können nur die unterschiedlichen Algorithmen und Auswertungsmöglichkeiten miteinander verglichen werden.

4.1 Was ist die Software R?

Die Grundlage der Datenauswertung bildet die Software R, oder auch R-Project, welche im Jahr 2004 gelauncht wurde. R ist eine Softwareumgebung für statistische Berechnungen und Grafiken, die durch Pakete je nach Bedarf erweitert werden kann.[131] Das Programm ist eine „freie Software“, wodurch die Nutzung, Vervielfältigung und Anpassung durch den Nutzer uneingeschränkt möglich sind.[132] Die R-Software ähnelt der bekannten und weitverbreiteten S-Sprache, welche das Mittel der Wahl für die Erforschung und Entwicklung von statistischen Methoden ist, in vielen Bereichen sehr. Der Vorteil von R ist die Vielzahl an Einsatzbereichen, aufgrund der Erweiterung mit unterschiedlichsten Paketen, wodurch statistische Methoden, wie lineare und nicht lineare Modellierung, Zahlenreihenanalysen und Klassifizierung oder grafische Techniken leicht anzuwenden sind. Außerdem ist die Gestaltung von Diagrammen in R ohne viel Aufwand zu erlernen. Die Diagramme können mathematische Formeln, sowie Symbole beinhalten und weisen eine hohe Qualität auf, wodurch die Grafiken der R-Software publikationsfähig sind. Für Datenverarbeitungen, wie Berechnung oder Manipulation und Matrizenanalyse ist R sehr gut geeignet.[133]

4.2 Alternative Softwareanwendungen

Neben R gibt es viele weitere etablierte Softwareanwendungen auf dem Markt, die für die Datenauswertung der NASA-Triebwerksdaten geeignet wären. Unter anderem gibt es die Software SciDAVis oder Matlab.

Bei SciDAVis handelt es sich um eine Anwendung, die für die Datenanalyse und eine hochwertige Graphenerstellung entwickelt wurde. Die aktuellste Version SciDAVis wurde im Mai 2021 veröffentlicht. Es handelt sich dabei um die Version 2 und ist online kostenlos verfügbar. Um eine möglichst große Menge an Leuten ansprechen zu können ist die

[131] Vgl. R Foundation. (o.D.). What is R?
[132] Vgl. Seite „Was ist GNU? (Juli 2021). In: GNU Betriebssystem.
[133] Vgl. R Foundation. (o.D.). What is R?

Benutzeroberfläche SciDAVis intuitiv und einfach zu bedienen. Außerdem soll sich SciDAVis durch die freundliche und offene Umgebung von Vergleichbaren freien Anwendungen lösen und somit neue und erfahrene Nutzer akquirieren. Diese Absicht wird sowohl durch die Eignung der Anwendung für die gängigsten Plattformen (Windows, MAC, Linux), als auch von der Unterstützung der Nutzer durch Tutorials, ein ausgeprägtes Forum und Benutzerhandbücher unterstrichen. Beispielhafte Nutzungsmöglichkeiten für SciDAVis sind:

- Tabellen, Matrizen und Diagramme in 2D und 3D
- Dateneingabe oder Datenimport sind möglich
- Alle Tabellenwerte können mit Standard oder individuellen Formeln berechnet werden
- Anpassung linearer und nichtlinearer Funktionen an die Daten
- 2D-Diagramme in Publikationsqualität
- Export in viele gebräuchliche Formate wie PDF oder JPG[134]

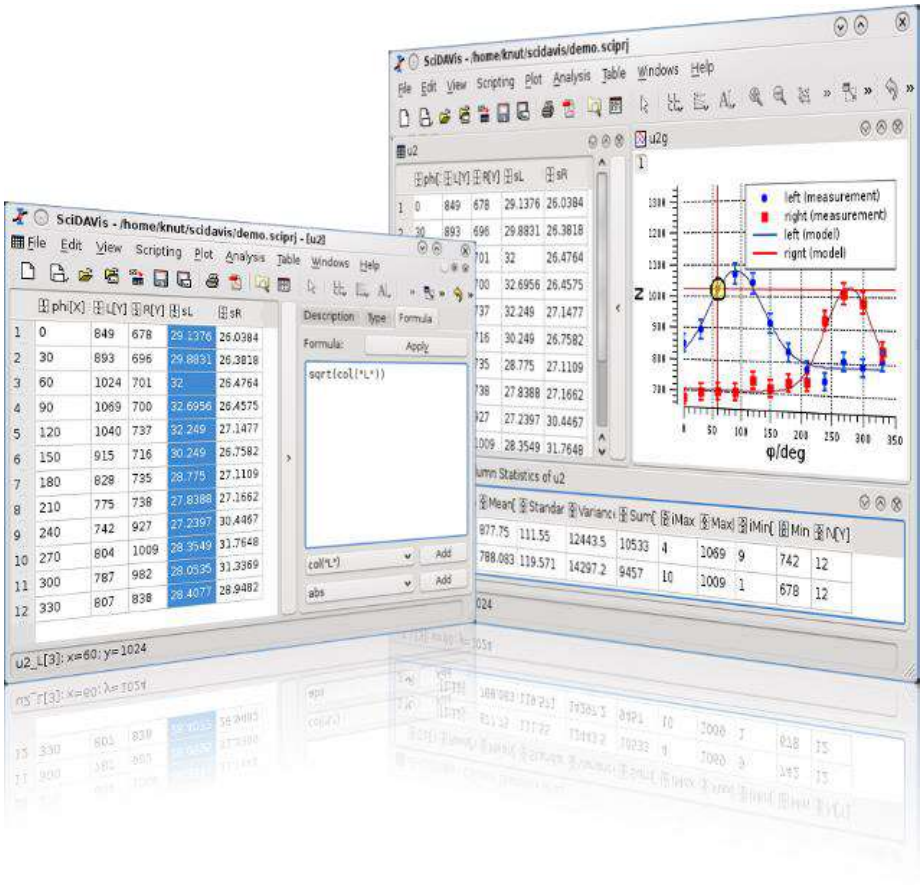

Abbildung 16: Beispielbild SciDAVis[135]

Eine weitere Softwareoption ist wie schon erwähnt Matlab. Matlab ist eine Software für mathematische Berechnungen, welche von dem Unternehmen MathWorks gelauncht wurde.

[134] Vgl. Standish R. (April 2017). SciDAVis - Mission Statement.
[135] Standish R. (April 2017). SciDAVis - What does it look like?

Im Vergleich zu SciDAVis ist Matlab eine kostenpflichtige Anwendung. Die Entwicklungsabsicht von Matlab war der Ansporn der Verknüpfung von Mathematik, Ingenieurswesen und die Algorithmenentwicklung, sowie die Schaffung einer Anwendung, um die Welt zu verändern. Matlab basiert auf einer leistungsstarken Programmiersprache, die mit Matrizen und Variablen umgehen kann. Außerdem sind in Matlab viele zusätzliche Tools implementiert, diese Tools reichen von Regelungssystemen, über Signalverarbeitung, bis hin zu computergestützten Finanzen und KI. Bei der Anzeige von Ergebnissen können diese auch visuell dargestellt werden. Als Zielgruppe für Matlab wurden Wissenschaftler und Ingenieure identifiziert, die Visionen haben. Die Nutzung von Matlab ist breit gefächert, wie im Folgenden zu sehen ist:

- Algorithmen und App Erstellung möglich
- Cloud Computing
- Deep- und Machine Learning
- Predictive Maintenance
- Signal- und Bildverarbeitung
- Robotik[136]

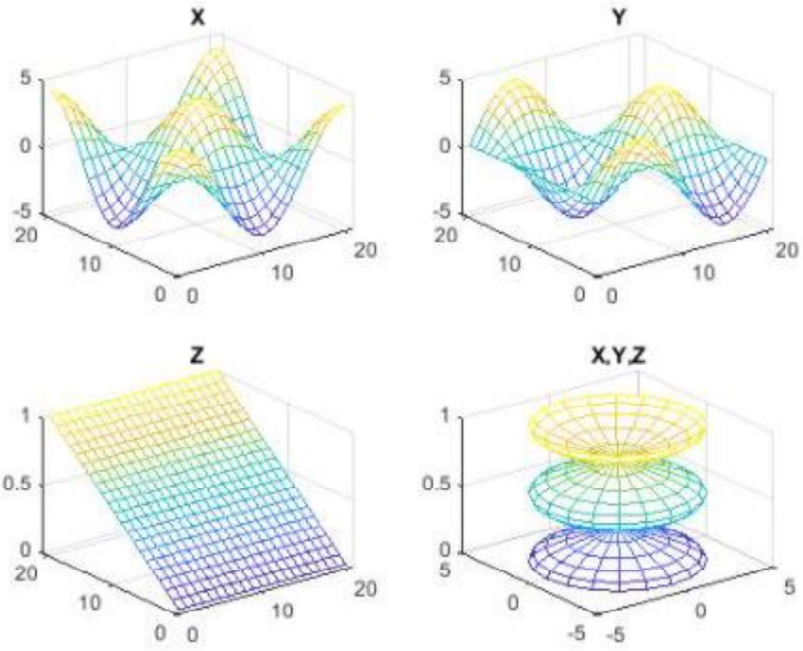

Abbildung 17: Beispielhafte Plots Matlap[137]

Im Vergleich der Anwendungssysteme fällt auf, dass diese sich grundsätzlich ähnlich sind, aber im Detail doch unterscheiden. R und SciDAVis sind kostenfreie Anwendungen, die sehr stark von dem Austausch der Nutzer profitieren. Die Nutzer können und sollen sich unterstützen, austauschen und Modifikationen bzw. Pakete entwickeln. Hinter den Softwareanwendungen steht kein Unternehmen, welches sich ausschließlich um die Software

[136] Vgl. The MathWorks, Inc (o.D.). MATLAB.
[137] The MathWorks, Inc. (o.D.). Hilfebereich_2D- und 3D-Plots.

kümmert. Anders sieht es bei Matlab aus. Diese Software wird von dem Unternehmen MathWorks vertrieben und wird von dem Unternehmen durchgehend weiterentwickelt bzw. verbessert. Wodurch auch der breite Anwendungsspielraum zu erklären ist. Dennoch kann nicht gesagt werden, dass eine der Anwendungen am besten ist, jede hat ihren Anwendungsbereich und ist stark abhängig vom Nutzer.

4.3 Datengewinnung

Für die Datengewinnung gibt es wie schon erwähnt viele unterschiedliche Messmethoden. Die Messgröße, welche die zu messende physikalische Größe eines Messobjektes ist[138], wird durch ein ausgewähltes Messverfahren mit passenden Sensoren gemessen. Bei dem Messverfahren wird die physikalische Grundlage der Messung (Messprinzip) mit einem darauf abgestimmten Vorgehen (Messmethode) praktisch angewandt[139]. Das wichtigste bei der Messung von Daten ist, dass die Wiederholbarkeit der Messungen gegeben ist. Somit hat die Messung unter vergleichbaren Bedingungen immer dieselbe systematische Messabweichung.[140] Die systematische Messabweichung ist definiert als Abweichung des Erwartungswert, bzw. dem gemessenen Wert vom wahren, dem Objekt innewohnenden Wert.[141] Wenn diese systematische Messabweichung immer dieselbe ist, haben die Messreihen auch eine Aussagekraft, wenn aber z.B. im 100stel Bereich gemessen wird und die systematische Messabweichung 10 hundertstell trägt und nicht gleichmäßig ist, liegen Schwankungen der Werte von mindestens 10% vor. Demensprechend müsste ein Messobjekt bei mehrmaligem Messen zu unterschiedlichen Zeitpunkten einen vergleichbaren Wert liefern.

Bei der Messung durch die Sensoren, dem sogenannten Messgrößenaufnehmer[142], werden die Daten in vorbestimmten Zeitintervallen gemessen. Diese Intervalle können stündlich, minütlich, alle 10 Sekunden, sekündlich oder auch 10-mal pro Sekunde ausfallen. Die erfassten Größen können viele physikalische oder chemische Eigenschaften widerspiegeln. Messbare Größen sind:

- Temperatur
- Geschwindigkeit
- Umgebungsdruck/ Betriebsdruck
- Schwingungen/ Vibrationen
- Gas-/Flüssigkeit-/Feststoffzusammensetzung
- Kraft
- Geometrische Daten
- Lautstärke
- Lichtstärke uvm.[143]

[138] DIN 1319-1:1995-01. Grundlagen der Meßtechnik. Seite 2 + 3.
[139] DIN 1319-1:1995-01. Grundlagen der Meßtechnik. Seite 7.
[140] DIN 1319-1:1995-01. Grundlagen der Meßtechnik. Seite 8.
[141] DIN 1319-1:1995-01. Grundlagen der Meßtechnik. Seite 3, 10 + 13.
[142] DIN 1319-1:1995-01. Grundlagen der Meßtechnik. Seite 19.
[143] Wünsche, V. (September 2021). Sensoren: Arten & Einsatz.

Nun werden zwei Messsensoren beispielhaft näher betrachtet. Ein optischer Sensor zum Beispiel funktioniert aufgrund von optischer Triangulation. Hierbei erzeugt eine Strahlungsquelle Rotlicht, Infrarotlicht oder einen Laser, der durch eine Optik gebündelt wird und auf den Prüfgegenstand ausgerichtet ist. Der reflektierte Strahl gelangt durch eine Optik in den Detektor, der mit Berücksichtigung der Streustrahlung den mittleren Abstand von Bezugsebene zum Prüfgegenstand ermitteln kann (siehe Abbildung 18). Eingesetzt wird eine solche Lasertriangulation innerhalb der Serienfertigung von Fahrzeug-Karosserien oder für die berührungslose Vermessung durch Koordinatenmessgeräte.[144]

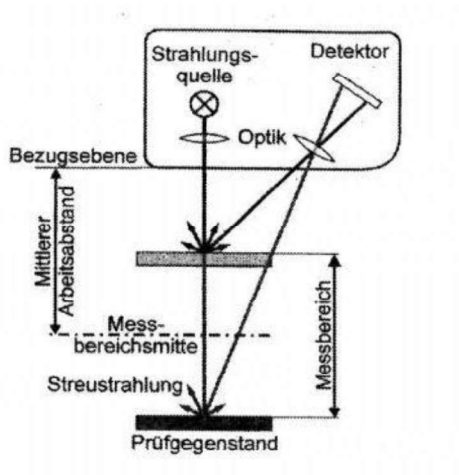

Abbildung 18: Aufbau Optischer Sensor[145]

Als weiterer Messsensor ist der Drucksensor, der unter anderem in Messgeräten wie dem Manometer oder einer Waage verbaut ist. Das Manometer soll den Druck eines Mediums erfassen, wo hingegen eine Waage den Gewichtsdruck/ die vertikal wirkende Gewichtskraft misst. In einem Drucksensor wird der physikalische Druck durch eine Kraftmessung ermittelt und darauffolgend in einen dazu proportionalen elektrischen Output transformiert.[146]

Der im folgenden betrachtete Datensatz ist einer von vier multivariaten Zeitreihen. Es werden die Datenreihen für unterschiedliche Triebwerke desselben Typs ermittelt. Dennoch unterscheiden sich die Triebwerke durch nicht genau definierte Fertigungsschwankungen. Diese erwähnten Fertigungsschwankungen sind aber noch in einem Bereich der nicht als Fehlerzustand angesehen wird. Außerdem ist zu betrachten, dass die Daten mit einem Sensorrauschen belegt sind, wodurch die Werte leicht verschlechtert sind. [147]

Die betrachteten Triebwerke laufen allesamt zu Beginn der Messungen normal und ohne Probleme. Mit zunehmender Betriebsdauer und folglich steigender Abnutzung stellen sich nach und nach Probleme und Fehler ein. Wenn diese zu sehr zugenommen haben wird das Triebwerk ausgewechselt. Das Ziel der Datenmessung und Zustandsüberwachung ist es, dass

[144] Vgl. Bracke S. (2012); Vorlesung: Methoden der Mess- und Prüfdatenanalyse. Veranstaltung: 2. Folie 4.
[145] Bracke S. (2012); Vorlesung: Methoden der Mess- und Prüfdatenanalyse. Veranstaltung: 2. Folie 4.
[146] Vgl. Seite „Druckmessgerät“ (April 2021). In: Wikipedia – Die freie Enzyklopädie.
[147] Vgl. Saxena, A. & Goebel K. (2008). Read me.

durch die Zustandsüberwachung gezielt Daten gemessen werden, mit denen es möglich ist, die noch übrigen Betriebszyklen bzw. die restliche RUL vorherzusagen.[148]

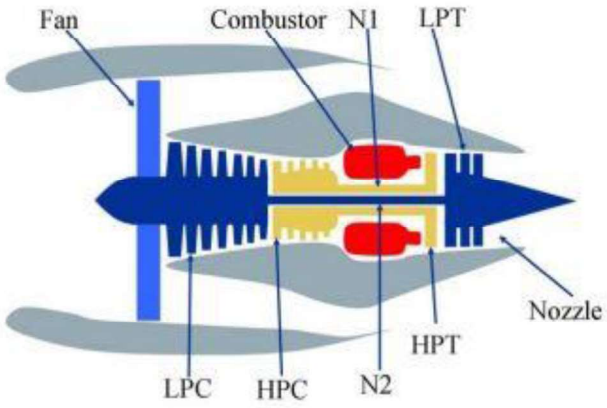

Abbildung 19: Darstellung des Triebwerkquerschnittes[149]

In der Abbildung 19 ist der Querschnitt des Triebwerkes dargestellt, dessen Daten im Folgenden näher betrachtet und analysiert werden. Es sind unterschiedliche Bauteile zu erkennen, die Abkürzungen aus der Abbildung lassen sich wie folgt übersetzen und erklären:

- LPC = Low-Pressure Compressor
- HPC = High-Pressure Compressor
- LPT =Low-Pressure Turbine
- HPT = High-Pressure Turbine
- Nozzle = Düse
- Combustor = Brennkammer
- Fan = Ventilator/Gebläse [150]

Die gemessenen Daten wurden in einer Zip-komprimierten Textdatei in 26 Spalten abgespeichert. Die Spalten sind wie folgt angeordnet:

- Nummer der Triebwerkseinheit in Spalte 1
- Zeit in Betriebszyklen in Spalte 2
- Die möglichen Betriebseinstellungen in Spalte 3 bis 5
- Sensormessungen in Spalte 6 bis 26[151]

Die einzelnen Zeilen der Spalten sind ein Auszug aus dem Betriebszyklus, somit sind Veränderungen zwischen den Datenwerten der Betriebszyklen möglich, aber nicht aus den Daten erkenntlich. Des Weiteren ist die genaue Dauer eines Betriebszyklus unbekannt. Die Messungen am Triebwerk für die Spalten 6 bis 26 sind in der Tabelle 2 zu sehen.

[148] Vgl. Saxena, A. & Goebel K. (2008). Read me.
[149] Saxena, A. & Goebel K. (2008). Damage Propagation Modeling. Seite 2.
[150] Vgl Saxena, A. & Goebel K. (2008). Damage Propagation Modeling. Seite 2.
[151] Vgl. Saxena, A. & Goebel K. (2008). Read me.

Tabelle 2: Messungen am Triebwerk[152]

Symbol	Description	Unit
T2	Total temperature at fan inlet	°R
T24	Total temperature at LPC outlet	°R
T30	Total temperature at HPC outlet	°R
T50	Total temperature at LPT outlet	°R
P2	Pressure at fan inlet	psia
P15	Total pressure in bypass-duct	psia
P30	Total pressure at HPC outlet	psia
Nf	Physical fan speed	rpm
Nc	Physical core speed	rpm
epr	Engine pressure ratio (P50/P2)	--
Ps30	Static pressure at HPC outlet	psia
phi	Ratio of fuel flow to Ps30	pps/psi
NRf	Corrected fan speed	rpm
NRc	Corrected core speed	rpm
BPR	Bypass Ratio	--
farB	Burner fuel-air ratio	--
htBleed	Bleed Enthalpy	--
NF_dmd	Demanded fan speed	rpm
PCNfR_dmd	Demanded corrected fan speed	rpm
W31	HPT coolant bleed	lbm/s
W32	LPT coolant bleed	lbm/s

Es werden in der folgenden Vorbereitung und Analyse alle Messreihen unabhängig von der Art der Messung, der Werte oder der Einheit betrachtet.

4.4 Vorbereitung des Datensatzes (deskriptiv)

Bevor der Datensatz analysiert wird, ist es von Vorteil den Datensatz und die Software für die Analyse vorzubereiten. Für die Software zum Beispiel bedeutet das, dass diese zu Beginn immer auf den Stand „0“ gesetzt wird und darauffolgend die benötigten Pakete installiert werden müssen. Diese Pakete sind wie eine Erweiterung der analytischen Fähigkeiten der Software R, da bestimmte Befehle, Modelle, Algorithmen unterschiedliche Pakete erfordern.

Die genutzten Pakete haben wie schon erwähnt unterschiedliche Funktionen. Als Beispiel kann das Paket *neuralnet* betrachtet werden. *Neuralnet* bietet die Möglichkeit in R neuronale Netze mit Rückführung, elastischer Rückführung mit oder ohne gewichteter Rücksetz-algorithmus oder der modifizierten, global konvergenten Version von Anastasiadis et al zu trainieren. Des Weiteren kann neben der individuell einstellbaren Aktivierungs- und Fehlerfunktion auch eine allgemeine Gewichtung kalkuliert werden. Das Paket *neuralnet* hat seit der erstmaligen Veröffentlichung im April 2008 viele Upgrades bekommen. Die aktuelle Version ist die jene mit der Nummer 1.44.2 aus Februar 2019. Viele Pakete sind erst

[152] Saxena, A. & Goebel K. (2008). Damage Propagation Modeling. Seite 3.

vollständig funktionsfähig, wenn andere Pakete ebenfalls installiert sind. Für das Paket *neuralnet* wird zum Beispiel unter anderem das Paket *MASS* benötigt.[153]

MASS ist die Abkürzung für „Modern Applied Statistics with S“ (zu Deutsch: Moderne angewandte Statistik mit S) und bietet unterstützende Funktionen und Datensätze. Das Paket *MASS* wurde im Mai 2009 das erste Mal veröffentlicht. Verantwortlich für das Paket sind Brian Ripley und Bill Venables, selbige hatten bereits 2002 ein Buch zu dem Thema mit dem gleichen Namen geschrieben. Das R-Paket hat mittlerweile viele Anpassungen bekommen, das letzte Update ist datiert auf den dritten März 2021. Hierdurch wird auch deutlich, dass an den R-Paketen weiterhin gearbeitet wird.[154]

Nachdem die Software nun „vorbereitet“ wurde, muss der Datensatz etwas genauer betrachtet werden. Wie in Abschnitt 4.3 bereits aufgezeigt gibt es 21 unterschiedliche Messungen pro Betriebszyklus. Die Betriebseinstellungen werden in der folgenden Analyse nicht näher betrachtet. Doch, bevor die Daten durch die Software R vorbereitet und aussortiert werden können, muss eine Verbindung zwischen dem Arbeitsordner, der Software und dem Datensatz erzeugt werden. Wichtig ist hierbei, dass der Datensatz in dem Arbeitsordner vorliegt.

Nun nachdem die Verbindung hergestellt wurde, wird festgelegt welche Datenspalte relevant ist und einen Mehrwert bietet. Wie schon erwähnt ist die erste Spalte die Triebwerksnummer, Spalte 2 der Betriebszyklus und die Spalten 6 bis 26 die gemessenen Daten. Dies wird abgesichert durch das Auslesen der Sensoren die Informationen bringen und sich dementsprechend mit der Zeit verändern. Der mathematische Begriff hierfür ist, dass die Standardabweichung ungleich 0 ist.

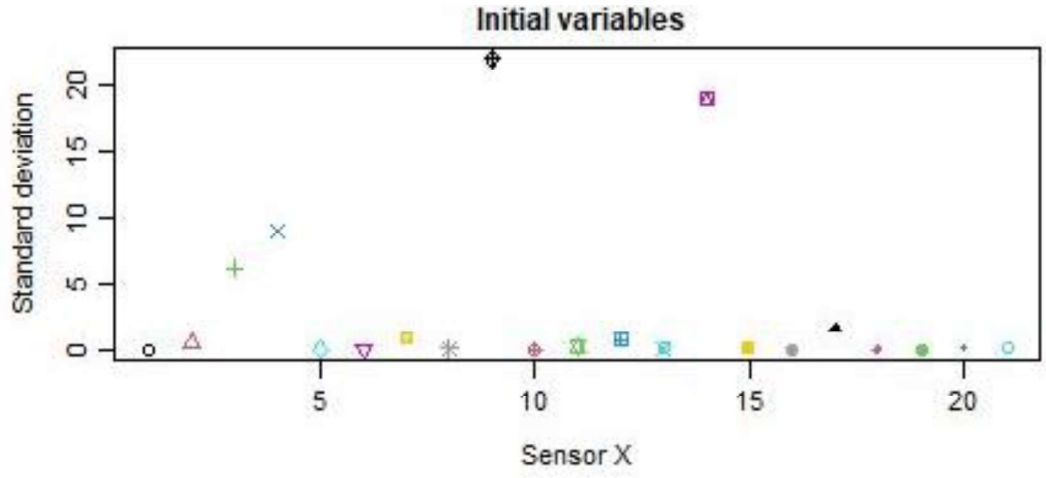

Abbildung 20: Standardabweichung der Messreihen

[153] Wright M. (Februar 2019). neuralnet: Training of Neural Networks.
[154] Ripley, B. (Mai 2021). CRAN-Package MASS.

Wie in Abbildung 20 „Standardabweichung der Messreihen“ zu sehen ist weisen die Messungen vieler Sensoren eine Standardabweichung im Bereich von null bis vier auf. Lediglich vier Sensoren haben eine Standardabweichung über 5, dazu gehören die Sensoren 3, 4, 9 und 14. Im nächsten Schritt werden die Sensoren mit einer Standardabweichung von null herausgefiltert, da eine Messreihe mit durchgehend gleichbleibenden Werten keinen Einfluss auf die Auswertung hat. Durch dieses Filtern fallen sechs Sensoren raus, dabei handelt es sich um Sensor 1, 5, 10, 16, 18 und 19.

Die noch übrigen Sensoren werden in Korrelation gesetzt und das Ergebnis wird grafisch in Abbildung 21 dargestellt. Bei genauer Betrachtung fällt auf, dass die Sensoren 6, 9 und 14 im Vergleich deutlich weniger mit anderen Sensoren und deren Ergebnissen korrelieren. Deutlich mehr Einfluss haben zum Beispiel die Sensoren 4, 7 und 12. Im Gesamtbild sind die Sensoren aber miteinander hoch korreliert.

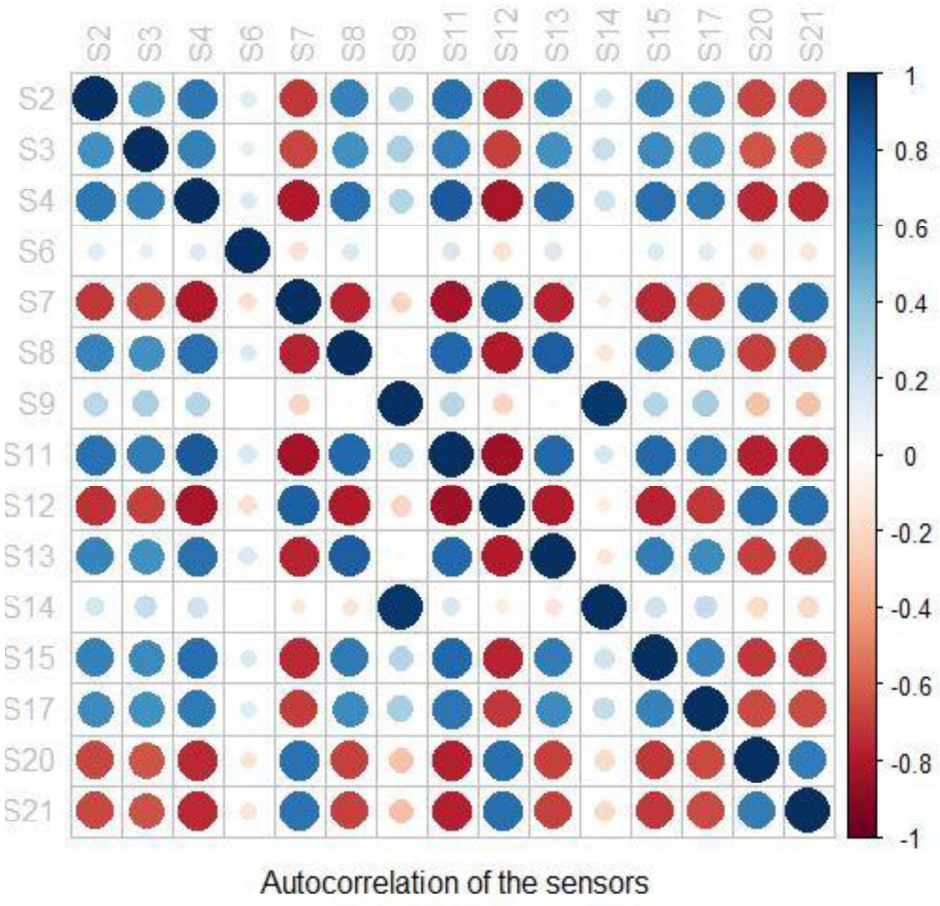

Abbildung 21: Darstellung der Korrelation der Sensoren

Der nächste Schritt bei dem Arbeiten mit den Sensordaten ist das in Korrelation setzen von den Sensormessungen mit der RUL. Durch diesen Schritt soll erkannt werden, welche der sich verändernden Sensormessungen einen negativen Einfluss auf die restliche Nutzungsdauer haben. Was so viel bedeutet wie, dass sich die Betriebsparameter durch die zunehmende Nutzungsdauer verändern.

Die erwähnte Korrelation zwischen Sensormessungen und RUL ist in der Abbildung 22 dargestellt. Vier der Messreihen liegen in einer positiven Korrelation mit der RUL, dazu gehören die Messreihen 7, 12, 20 und 21. Alle übrigen Messreihen stehen in einer negativen Korrelation mit der RUL. Den negativsten Zusammenhang weisen die Sensormessreihen 4 und 11 auf, mit ca. -0,7. Somit lässt sich sagen, dass von den zu Beginn 21 Messreihen nur 11 Messreihen mit ihren Veränderungen einen negativen Einfluss auf die RUL haben.

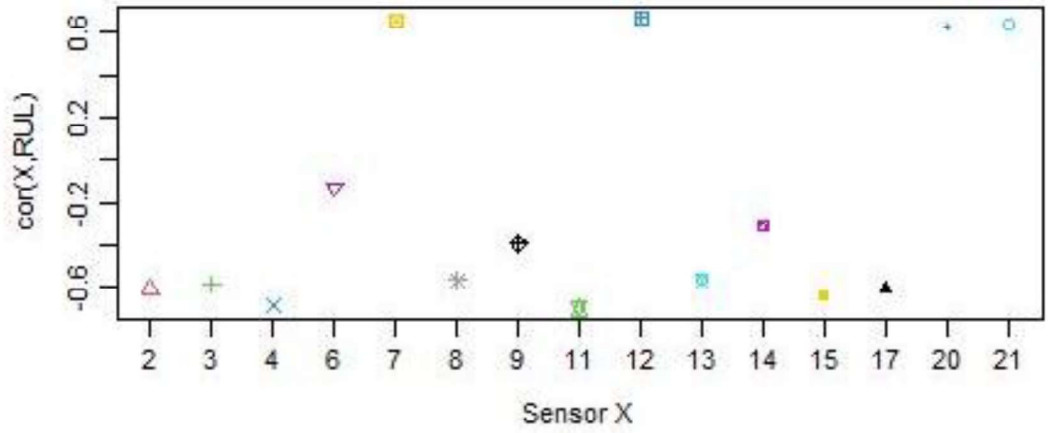

Abbildung 22: Korrelation Sensormessreihen und RUL

Nachdem im letzten Schritt der Einfluss der Sensormesswerte auf die RUL beschrieben wurde, kann als nächstes die Korrelation von zwei Sensoren bestimmt werden. Am besten bieten sich hierfür die zwei Sensoren an, die den stärksten negativen Einfluss auf die RUL haben und diese wären der Sensor 4 und der Sensor 11. Jede Kombination an Messwerten wurde in die Abbildung 23 eingetragen, beim Sensor 4 befinden sich die gemessenen Werte zwischen 1380 und 1440, bei dem Sensor 11 zwischen 46,7 und 48,6. Welche Einheit diese Werte aufweisen ist nicht aus den Daten und der dazugehörigen Quelle zu entnehmen. Die nun entstandene Datenwolke wurde in zwei Bereiche eingeteilt, einen blauen und einen roten. Der blau gefärbte Bereich ist jener, bei dem ein großer RUL vorliegt. Ein großer RUL zeigt einen „neuwertigen“ Zustand an. Der rote Bereich hingegen, bezieht sich auf einen kleinen RUL, der Zustand ist somit gebraucht. Es ist in der Grafik gut zu erkennen wie sich die Werte mit abnehmenden RUL langsam verschieben. Es gibt einen reinen blauen Bereich um (47,2/1400), einen Übergangsbereich (47,6/1410) und einen rein roten Bereich (48,0/1420).

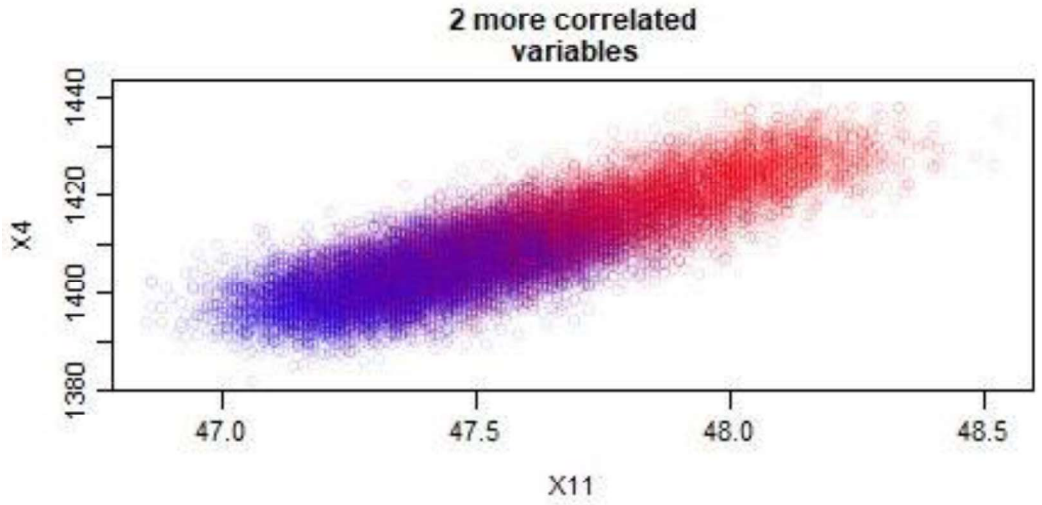

Abbildung 23: Korrelation der Sensoren 4 und 11 in zwei Phasen

Im nächsten Schritt sollen die Dimensionen des Datensatzes reduziert werden. Um das erfolgreich umsetzen zu können, wird eine Hauptkomponentenanalyse durchgeführt. Hierbei sollen die hochkorrelierten Sensormessungen durch die lineare Kombination der Variablen die Datensatzdimensionen reduzieren. Es werden im Folgenden die Hauptkomponenten durch die Linearkombination ermittelt und ausgedrückt. Damit die Daten leichter zu verstehen sind, wird für jede Hauptkomponente (Principal Component, kurz PC), die Standardabweichung und Varianz gebildet. Die Gesamtvarianz wird durch die kumulierte Summe der PCs ermittelt und liegt bei 15. Bei dem Anteil der PCs an der Gesamtvarianz fällt deutlich auf, dass die PCs 1 bis 3 über 80% der Gesamtvarianz im Datensatz erklären. Die gesamten Daten zu den Hauptkomponenten werden in den nachfolgenden Tabellen 3 und 4 aufgezeigt und grafisch in den Abbildungen 24 bis 26 dargestellt.

Tabelle 3: Daten zu den Hauptkomponenten Teil 1

	PC1	PC2	PC3	PC4	PC5
Standard-abweichung	3,0044	1,4486	0,9831	0,6434	0,6069
Varianz	9,0264	2,0985	0,9664	0,4139	0,3683
Kum.Summe der Varianz	9,0264	11,1249	12,0914	12,5053	12,8735
Anteil an der Gesamtvarianz	60,18%	13,99%	6,44%	2,76%	2,46%

Tabelle 4: Daten zu den Hauptkomponenten Teil 2

	PC6	PC7	PC8	PC9	PC10
Standard-abweichung	0,5902	0,5533	0,5383	0,5054	0,4541
Varianz	0,3438	0,3061	0,2898	0,2554	0,2062
Kum.Summe der Varianz	13,2219	13,5280	13,8178	14,0732	14,2794
Anteil an der Gesamtvarianz	2,32%	2,04%	1,93%	1,70%	1,37%
	PC11	PC12	PC13	PC14	PC15
Standard-abweichung	0,4427	0,4202	0,4154	0,3823	0,1711
Varianz	0,1960	0,1766	0,1725	0,1461	0,0293
Kum.Summe der Varianz	14,4754	14,6520	14,8245	14,9707	15,0000
Anteil an der Gesamtvarianz	1,31%	1,18%	1,15%	0,97%	0,20%

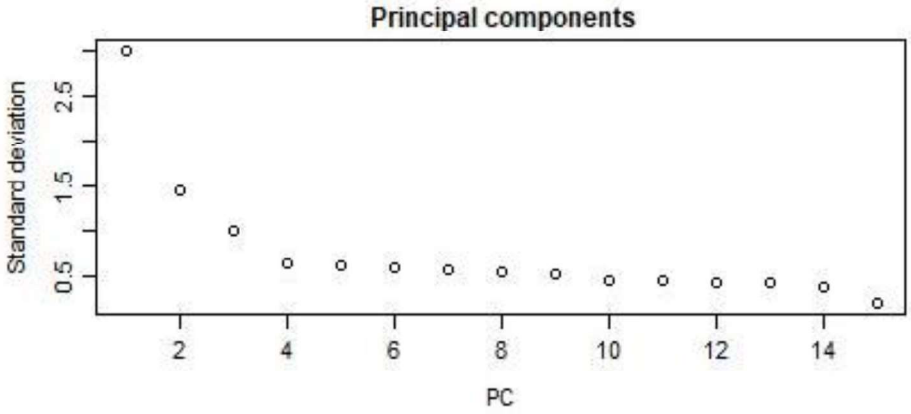

Abbildung 24: Standardabweichung der „Principal Components“

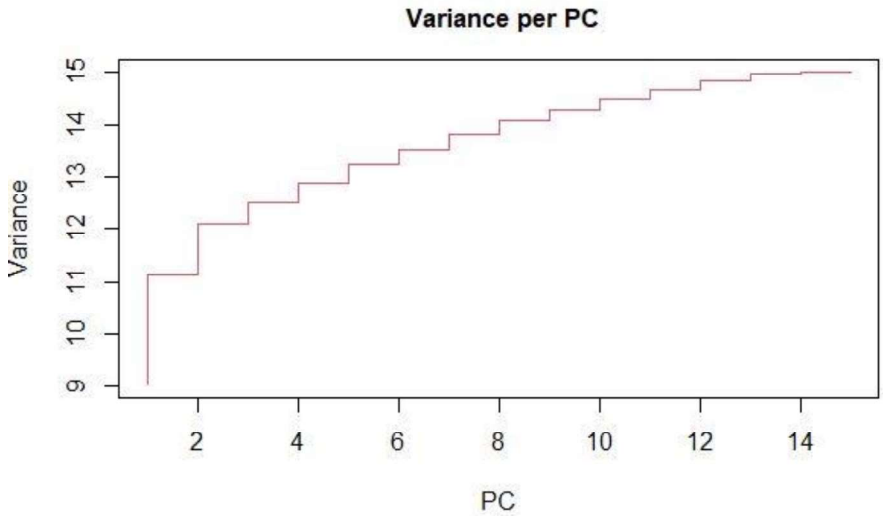

Abbildung 25: Varianz der „Principal Components“, kumuliert

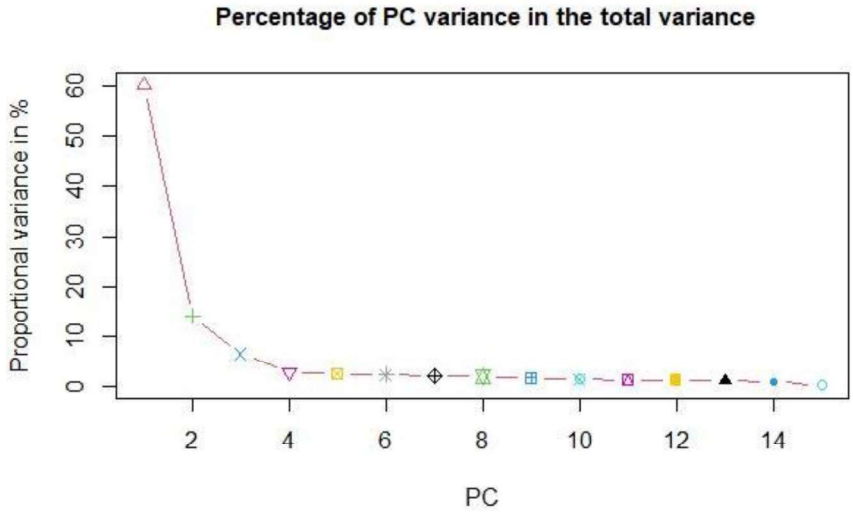

Abbildung 26: Anteil der PC-Varianz an der Gesamtvarianz

Nachdem die PCs und deren Standardabweichung, sowie Varianz bestimmt wurde, kann im nächsten Schritt wie bereits für die Sensormessungen eine Korrelation zwischen den PCs und der RUL hergestellt werden. Hierbei fällt auf, was bereits durch die Abbildung 26 visualisiert wurde, PC1 hat als Hauptkomponente 1 den stärksten Einfluss auf die RUL und somit den höchsten negativen Korrelationswert mit -0,741, wohingegen PC2 den höchsten positiven Korrelationswert mit 0,165 aufweist. Die restlichen Korrelationswerte von PC3 bis PC15 schwanken zwischen -0,018 und 0,024 und weisen somit einen fast vernachlässigbar geringen Wert auf.

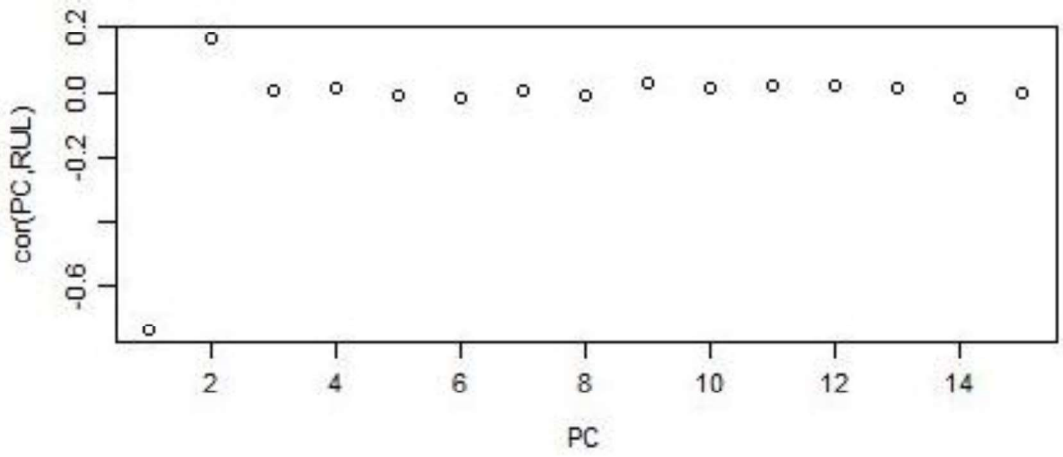

Abbildung 27: Korrelation PCs und RUL

Auf Basis des vorherigen Schrittes werden nun die zwei einflussreichsten PCs in Korrelation gesetzt, PC1 und PC2. Wie schon bei den Sensoren wird die entstandene Datenwolke durch einfaches Clustering in die zwei Phasen neuwertig; blau und gebraucht; rot unterteilt. Außerdem fällt in Abbildung 28 auf, dass erneut ein ellipsenförmiger blauer Bereich vorliegt, der große Unterschied ist aber, dass die rote Wolke, im Vergleich zu Abbildung 23, sehr großflächig ist und nicht nur eine Verschiebung der blauen Datenwolke in eine Richtung. Dieser Unterschied liegt daran, dass die Hauptkomponenten in gewisser Weise eine Matrix bzw. ein Vektor sind. Wodurch die Veränderungen andere Auswirkungen haben. Des Weiteren ist der Übergangsbereich vom Blauen zum Roten nicht so sehr ausgeprägt wie bei der Sensorkorrelation, diese Besonderheit würde das grobe Definieren eines Grenzbereichs ermöglichen, worauf aber mit Blick auf die folgende Analyse und das Ziel einer zustandsorientierten Instandhaltungsstrategie verzichtet wird.

Um die Aussagekraft und den Einfluss von PC1 und PC2, noch einmal zu betonen, wurden die zwei PCs mit der dritt- und vierthöchsten Korrelation zum RUL in Korrelation gesetzt. Die Werte hierfür lagen für PC9 bei 0,0241 und für PC 11 bei 0,0229. Das Ergebnis der Korrelation der zwei PCs ist in Abbildung 28 zu sehen. Es liegt eine fast runde Datenwolke vor, die sich mit einem Radius von ca. 1 um den Punkt 0;0 bewegt. Alle Phasen liegen übereinander, folglich ist keine Entwicklung, Verschiebung oder ähnliches zu erkennen.

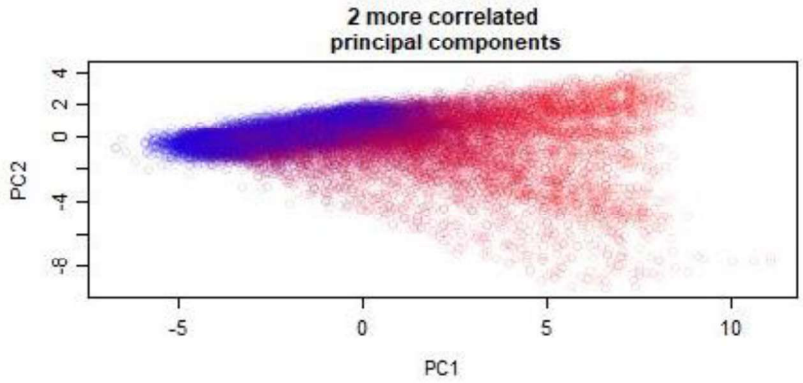

Abbildung 28: Korrelation PC1 und PC2

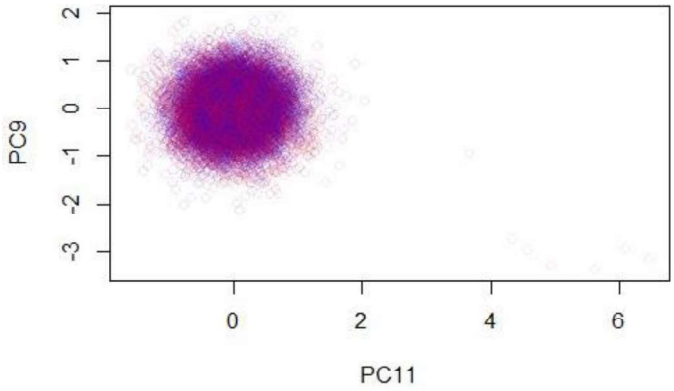

Abbildung 29: Korrelation PC9 und PC11

Der letzte vorbereitende Schritt dient der graphischen Darstellung der Entwicklung der Triebwerkzustands über die Zeit. Hierzu wird für jede Einheit eine Funktion erstellt. Diese Funktion wird durch die Hauptkomponente PC1 beschrieben. In der daraus resultierenden Grafik (Abbildung 30) sind zwei Besonderheiten sehr gut zu erkennen. Zum einen, sind die Startwerte der Einheiten unterschiedlich. Der Bereich, in dem die Einheiten nach dem Einbau starten erstreckt sich von -5 bis ca. 0. Damit macht der Startbereich circa ein Drittel der möglichen Werte der Funktion aus (-5 bis 10). Diese Auffälligkeit erschwert die Auswertung der Verläufe. Außerdem ist gut zu erkennen, dass unterschiedlichen Einheiten unterschiedliche Werte vor ihrem Ausfall aufweisen können. Somit ist es nicht möglich einen Wert als Grenz-bzw. Ausfallwert zu bestimmen. Dennoch lässt sich sagen, dass viele Einheiten ausfallen, wenn sie sich in einem Wertebereich von 7 bis 8 befinden.

Zur weiteren verständlicheren Darstellung wurden für jede Einheit in Abbildung 31 das Ausfallverhalten linear dargestellt. Hierbei wird erkenntlich, dass viele Einheiten einen vergleichbaren Verlauf haben und ein Großteil der Einheiten wie schon erwähnt zwischen 7 und 8 ausfallen. In beiden Abbildungen ist außerdem auch gut zu erkennen, dass es eindeutige Ausreißer nach oben, mit Werten bis zu 11, gibt.

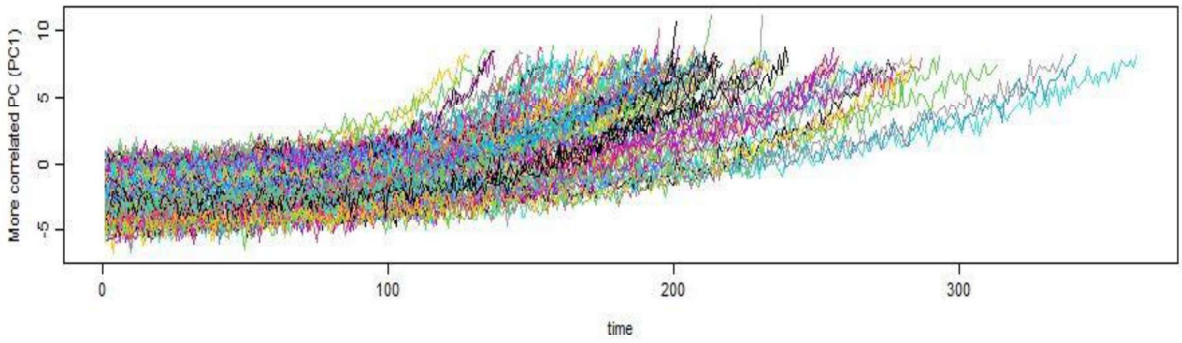

Abbildung 30: Funktion der Einheiten in Abhängigkeit von PC1 und t

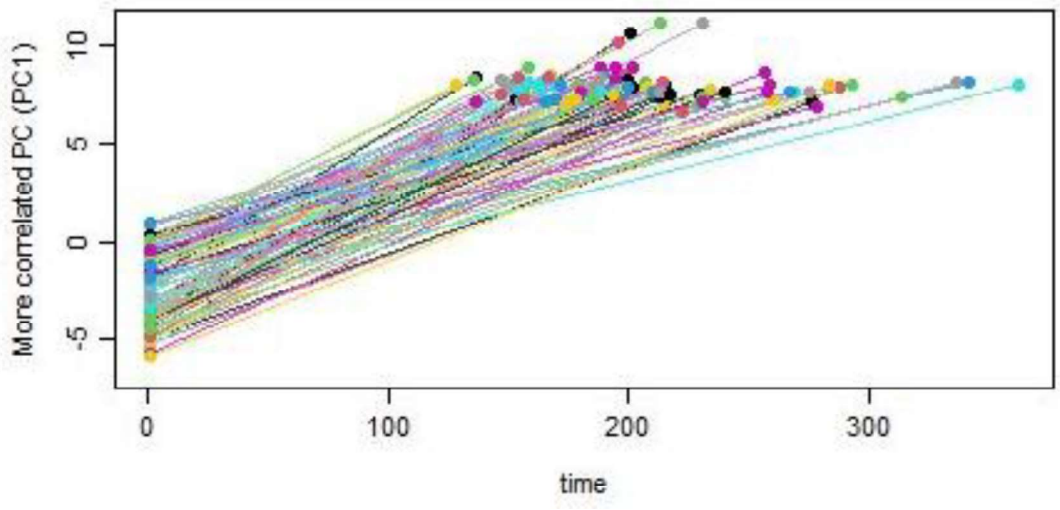

Abbildung 31: Lineare Funktion der Einheiten in Abhängigkeit von PC1 und t

4.5 Klassifizierung des Datensatzes

Nachdem der Datensatz bearbeitet und vorbereitet wurde, kann auf dessen Grundlage der nun übrige Datensatz analysiert werden. Hierzu werden unterschiedliche Algorithmen auf den Datensatz angewendet. Das Ziel ist weiterhin, durch die Datenanalyse eine Vorhersage über die restliche Lebensdauer zu einem bestimmten Zustandszeitpunkt treffen zu können.

Es werden im Folgenden drei unterschiedliche Algorithmen auf den Datensatz angewendet, bei denen es sich um das Hidden-Markov-Model, den k-Means-Algorithmus und die Hierarchische Clustermethode handelt.

Bei dem Hidden-Markov-Modell (HMM) handelt es sich um stochastische Modelle, die auf Markow-Ketten beruhen. Unter einer Markow-Kette sind spezielle stochastische Prozesse zu

verstehen, die die Wahrscheinlichkeiten aller zukünftigen Zustände nur vom momentanen Zustand des Objektes unabhängig vom Zeitpunkt darstellen können. Im Allgemeinen werden bei Markov-Modellen Zustände und Ereignisse verknüpft, wobei die Ereignisse die Beobachtungen widerspiegeln. Die Besonderheit des HMM liegt darin, dass die Zustände verborgen und somit nicht bekannt sind, das hat zur Folge, dass eine Ausgabe durch unterschiedliche Zustände erzeugt werden kann. Die Einsatzbereiche vom HMM sind das Herausfinden von Beobachtungswahrscheinlichkeiten, Ermittlung des besten Pfads und der Aufbau des besten Modells aus Trainingsdaten.[155]

Der zweite verwendete Algorithmus ist der k-Means -Algorithmus (KMA). Der KMA ist eine Clusteranalyse zu Gruppierung von Objekten und kann aus einer Vielzahl vergleichbarer Objekte, mit einer vorab definierten Anzahl an Gruppen, das jeweilige Zentrum eines Clusters bestimmen. Die Laufzeit des Algorithmus nimmt proportional zur Anzahl der Datenpunkte zu und es werden wenige Schleifendurchläufe benötigt. Eingesetzt wird der KMA im Big-Data-Bereich, in der Segmentierung von Bilddaten und besonders in der Analyse von Kundenverhalten und Marketing.[156]

Der übrige Algorithmus der schon erwähnten drei Algorithmen ist die agglomerative hierarchische Clustermethode (AHC), hierbei wird zu Beginn für jedes Objekt ein Cluster gebildet. Schritt für Schritt werden die vorhandenen Cluster zu größeren zusammengefasst. Dieses Bottom-up Zusammenfassen wird so lange fortgesetzt, bis die gewünschte Clusteranzahl erreicht wurde. Eine Rückverfolgbarkeit der Verbindungen sind nicht möglich.[157]

Mit den drei Algorithmen konnten die Abbildungen 28 und 30 in zwei Phasen unterteilt werden. Beginnend mit der Funktion der einzelnen Einheiten in Abhängigkeit von PC1 und der Zeit, welche in den Abbildungen 32 bis 34 verbildlicht sind. Die Phasen, in die die Funktionen unterteilt wurden, sind in dem Fall sehr guter Zustand (schwarz) und gebrauchter Zustand (rot). Bei genauer Betrachtung der drei unterschiedlichen Algorithmen fällt auf, dass der Grenz-/Übergangsbereich der zwei Phasen sehr unterschiedlich ausfällt. Bei dem HMM liegt kein fester Grenzwert zwischen den Phasen vor, sondern eher ein leicht schwankender Bereich. Des Weiteren gibt es zwischenzeitliche „Ausreißer“ aus den Phasen, zum Beispiel bei einer Zeit von 50 oder 170 liegt ein kurzzeitiger gebrauchter Zustand vor.

Im Vergleich zum HMM liegt beim KMA bei ungefähr 1 der Grenzwert zwischen den Phasen vor. Dieser Grenzwert zieht sich über alle Zeiten. Neben dem Grenzwert gibt es noch einen großen Unterschied der Phasen der zwischen der HMM und der KMA Abbildung. Wenn die Phaseneinteilung nach KMA als die wahre Einteilung angesehen wird, waren eingebaute Triebwerke, die eigentlich neuwertig sein sollten, schon dem gebrauchten Zustand zuzuordnen. Die Frage, die folglich daraus resultiert ist, ob das neuwertige Triebwerk schlecht verarbeitet ist oder es fehlerhaft installiert wurde.

[155] Vgl. Haenelt K. (2006). Kursfolien: Hidden Markov Models.
[156] Vgl. Luber S., Litzel N. (Juli 2018). Was ist der k-Means-Algorithmus? BigData-Insider.
[157] Vgl. Seite „Hierarchische Clustermethode“ (Mai 2021). In: Wikipedia – Die freie Enzyklopädie.

In der Abbildung 34 wird die Einteilung nach AHC dargestellt. Bei der Einteilung nach AHC gibt es wie schon bei dem HMM keinen festen Grenzwert zwischen den zwei Zuständen. Dennoch ist ab einer Zeit von ca. 130 bis 225 ein Mischbereich der Zustände zu erkennen. Des Weiteren ist bei allen Einheiten der Triebwerke bis zur Zeit von ca. 90 ausschließlich ein neuwertiger Zustand zu erkennen.

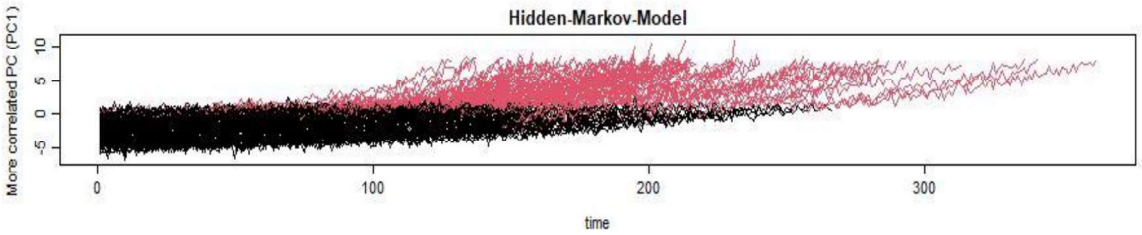

Abbildung 32: Funktionen der Einheiten nach HMM-Einteilung

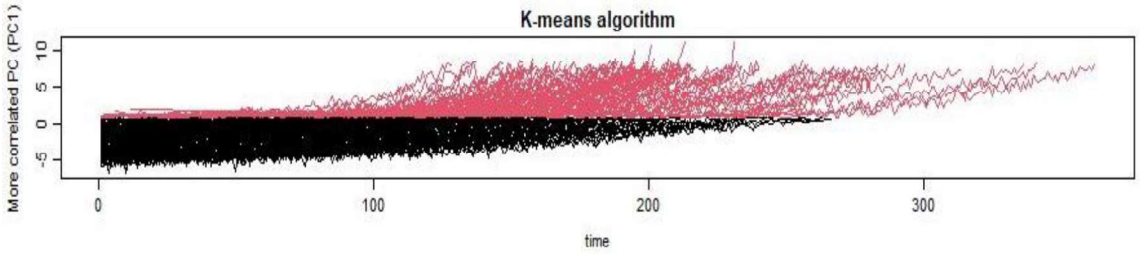

Abbildung 33: Funktion der Einheiten nach k-Means-Einteilung

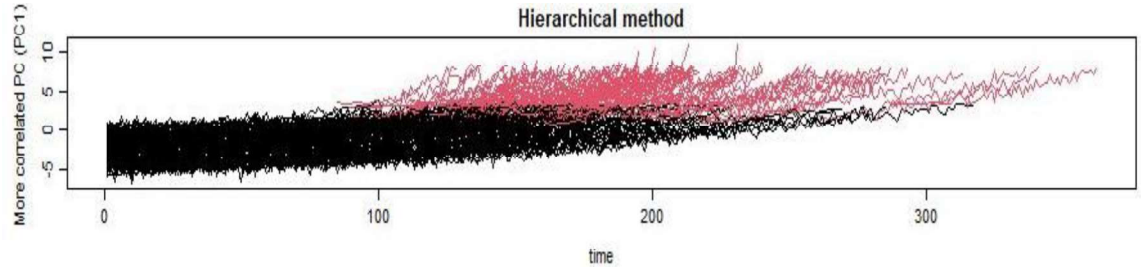

Abbildung 34: Funktion der Einheiten nach hierarchischer Clustermethode

Nachdem die Funktionen der Triebwerke in die zwei Zustände eingeteilt wurden, werden die Algorithmen ebenfalls die Korrelationswolke von PC1 und PC2 in die zwei Zustände einteilen. Die Ergebnisse der Zustandseinteilung der Algorithmen ist in den Abbildungen 35 bis 37 zu sehen.

Bei der Betrachtung der Abbildung 35, der Zustandseinteilung nach HMM fällt auf, dass diese Einteilung der Clustereinteilung in Abbildung 28 sehr ähnlich ist. Der Zustand „Neuwertig“ liegt ellipsenförmig im linken Bereich der Wolke und erstreckt sich bis auf wenige Ausreißer von PC1 gleich -6 bis 1,25. Ab PC1 gleich -2,5 beginnt der gebrauchte Bereich am unteren Ende des neuwertigen Zustands und ab -1,5 am oberen Bereich

Wie schon bei der Zustandseinteilung der Funktionen der Triebwerke mit dem KMA hat die Zustandseinteilung der Korrelation zwischen PC1 und PC2 einen Grenzwert, der die zwei Zustände teilt. In dem Fall liegt der Grenzwert bei PC1 gleich 1. Alles was sich links von dem Grenzwert befindet ist der neuwertige Zustand und alles auf der rechten Seite ist der gebrauchte Zustand. Im Vergleich zu der Einteilung mit dem HMM lässt sich ein größerer neuwertiger Zustand bei der KMA-Einteilung erkennen, des Weiteren weist der Grenzbereich bei dem KMA nur minimal eine Mischung der Zustände auf.

Zu guter Letzt wird die Einteilung nach dem AHC näher betrachtet. Hierbei fällt auf, dass die Größe des neuwertigen Zustands im Verhältnis zur KMA-Einteilung nochmal gewachsen ist, da der neuwertige Bereich zwei weitere Ausprägungen hat. Eine Ausprägung ist am unteren Bereich der Datenwolke, wie ein kleiner Arm und die zweite Ausprägung ist am oberen Ende und könnte als ein großer Arm beschrieben werden. Wie auch schon bei der Betrachtung der Funktion der Einheiten ist der neuwertige Bereich der Größte der drei Algorithmen.

Bei der Gesamtbetrachtung der Abbildungen der Korrelationen fällt auf, dass der neuwertige Zustand wie schon erwähnt immer weiter zu nimmt. Außerdem kann man den Abbildungen 32 bis 37 gut entnehmen, dass die drei Algorithmen ähnliche Einteilungen haben, die sich zum Teil dennoch signifikant unterscheiden, wie z.B. durch den Grenzwert oder die Einordnung bei dem k-Means, dass neue Triebwerke schon beim Einbau gebraucht sind.

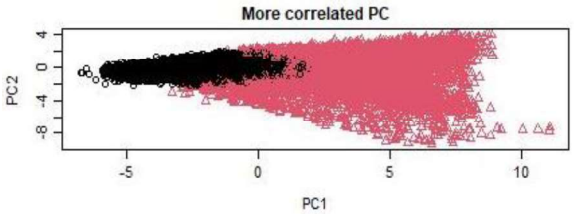

Abbildung 35: Korrelation von PC1 und PC2 nach HMM Einteilung

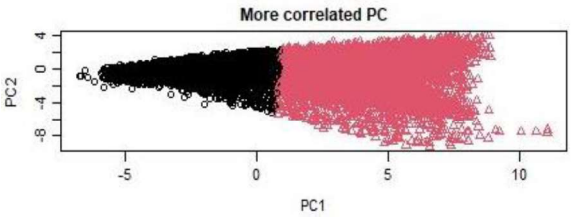

Abbildung 36: Korrelation von PC1 und PC2 nach k-Means-Einteilung

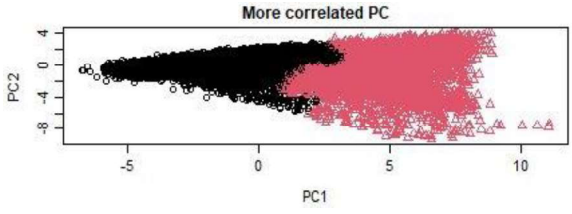

Abbildung 37: Korrelation von PC1 und PC2 nach hierarchischer Clustermethode

4.6 Vorhersage des RUL

Um die RUL vorherzusagen, können unterschiedliche Algorithmen verwenden werden. Es ist möglich die Algorithmen in unter anderem vier unterschiedliche Arten zu gliedern. Dabei handelt es sich um folgende vier Arten: lineare Methoden, nicht lineare Methoden, neuronale Netze und Support-Vector-Machinen.[158]

4.6.1 Anwendung linearer Methoden

Wenn die Daten mit einer linearen Methode analysiert werden, wird in der Regel von linearer Regression (LR) gesprochen. Bei einer LR handelt es sich um ein Analyseverfahren des Einflusses von Variablen auf eine metrisch abhängige Variable. Das bedeutet, dass eine Variable abhängig ist und durch unabhängige Variablen beschrieben wird. Etwas vereinfacht gesagt wird eine Gerade durch eine Punktewolke gelegt, die eine Art Vorhersage der Werte ist und die Entwicklung bzw. Verschiebung der Daten anzeigt.[159]

Gut geeignet ist die LR für die Darstellung von Wachstumsfunktionen, Prognosen von z.B. die Anzahl von Zuschauern oder Zusammenhängen von z.B. Marketingaufwand und Vertriebsmengen.[160]

4.6.2 Anwendung nicht linearer Methoden

Die nicht linearen Methoden sind gut mit der linearen Regression zu vergleichen. Denn wie bei der LR auch, wird bei der nicht linearen Regression (NLR) eine anhängige Variable durch eine oder mehrere unabhängige Variablen beschrieben. Der entscheidende Unterschied ist aber, dass bei der NLR keine Gerade ermittelt wird, sondern ein nicht lineares mathematisches Modell, dazu gehören Funktionen mit Potenzen oder der e-Funktion.[161]

Ein typischer Einsatzbereich für NLR ist in der Medizin und Biologie. Hier werden zum Beispiel die Durchlässigkeit von Membrantechnologien, der Sauerstoffverbrauch oder die Wirkung von Medikamenten dargestellt.[162] Andere Bereiche sind in der Industrie, wo Fehler in der Produktion in Abhängigkeit von unterschiedlichen Faktoren, wie Geschwindigkeit, Komplexität, Größe oder Art der Schicht veranschaulicht werden sollen.

4.6.3 Anwendung künstliche neuronaler Netze

Bei den künstlich neuronalen Netzen (NN) handelt es sich um Neuronen, die lineare Regressionen von Eingaben mit einer Gewichtung berechnen und dann nicht lineare Aktivierungsfunktionen berechnen, um Ausgänge zu erzeugen. Die Netzparameter werden in

---

[158] Vgl. Serradilla O. et al. (2020). Deep learning models for predictive maintenance: […]. Association for Computing Machinery. Seite 7.
[159] Vgl. Lois D. (2015). Kursfolien: Lineare Regression: Grundlagen und BLUE-Annahmen; Folie 3 + 5.
[160] Vgl. Grünwald R. (Dezember 2020). Der Klassiker lineare Regression einfach erklärt. Novustat Statistik-Blog.
[161] Vgl. IBM, Hrsg. (o.D.). Nicht lineare Regression.
[162] Vgl. Ruckstuhl A. (2008). Nichtlineare Regression. Zürcher Hochschule für angewandte Wissenschaften.

der Regel nach dem Zufallsprinzip initialisiert, wodurch folglich bei der Anwendung des Gradienten-Abstiegs-Algorithmus ein Lernprozess erfolgt. Durch diesen Lerneffekt können die Neuronen in Bezug auf deren gemachte Fehler angepasst werden und somit die Fehlerquote reduziert werden. Am besten bekommen NN Trainingsdaten, um zu lernen und sich zu entwickeln.[163]

Das NN hat eine sehr gute Eignung für die Verwertung von großen Datenmengen mit ungenauer Dateninformation mit dem Ziel eines genauen Ergebnisses. Das bedeutet, es kann verwendet werden für Bild-, Sprach- und Schrifterkennung, die Prognose und Steuerung komplexer Systeme oder auch komplexe Simulationen.[164] Der Einsatzbereich des NNs ist vielfältig und ist durch das Lernen auch auf spezielle Anwendungen und Probleme anwendbar.

4.6.4 Anwendung der Support-Vector-Machine

Der vierte Bereich umfasst die Support-Vector-Machine (SVM). Die SVM werden oftmals auch beim „Machine Learning“ eingesetzt, da es sich hierbei um eine mathematische Methode handelt, die es ermöglicht mit dem statistischen Ansatz Objekte zu Klassifizieren. Die Klassifizierung mit der SVM ist sowohl linear als auch nicht linear möglich. Die Anwendung von SVM verfolgt das Ziel, aufgrund von Training und vorhandenen Daten, neue Daten den richtigen Klassen zuordnen zu können. Besonders bei der Bild-, Text- und Schrifterkennung wird die SVM eingesetzt.[165]

4.6.5 Analyse der Daten

Die finale Analyse der Daten wird mit sieben unterschiedlichen Algorithmen durchgeführt. Hierbei werden die Daten des Zustandes 2 „gebraucht“ von dem HMM verwendet. Folglich wird die RUL erst bestimmt, wenn das Objekt von dem neuwertigen in den gebrauchten Zustand gewechselt ist. Dazu kommt die Einteilung in zwei Bereiche für die Kreuzvalidierung, den Trainingsdatenbereich und den Testingdatenbereich. Der Trainingsdatenbereich umfasst 80% der Daten, hier sollen die Algorithmen trainiert werden, das heißt die Koeffizienten und Parameter der Algorithmen werden ermittelt. Mit den anderen 20% der Daten wird die Testingphase durchgeführt. Die angepassten Algorithmen werden auf die Daten angewendet, wodurch kontrolliert werden kann, ob diese überangepasst sind. Von dem sogenannten Overfitting spricht man, wenn die Modelle und Algorithmen zu sehr auf die Trainingsdaten spezialisiert sind, was zur Folge hat, dass ein kleiner Fehler im Training einen großen Fehler im Testing zur Folge hat.[166] Die Einteilung in die zwei Bereiche (80/20) erfolgt zufällig, wodurch die Ergebnisse bei Auswertung leicht variieren können.

---

[163] Vgl. Serradilla O. et al. (2020). Deep learning models for predictive maintenance: […]. Association for Computing Machinery. Seite 8.
[164] Luber S., Litzel N. (März 2019). Was ist ein neuronales Netz? BigData-Insider.
[165] Luber S., Litzel N. (November 2019). Was ist eine Support Vector Machine? BigData-Insider.
[166] Vgl. Grünwald R. (Juli 2019). So vermeidet man Overfittung und Underfittung! Novustat Statistik-Blog.

Bei den Algorithmen handelt es sich um die lineare Regression, das quadratische Modell, das exponentielles Modell, die Support-Vector-Machine, der Random-Forest, K-nearest neighbors und das künstlich neuronale Netz. Die Wirksamkeit und Genauigkeit der Algorithmen werden durch den MSE, der mittleren quadratischen Abweichung der Vorhersage und dem Gütewert der Prognose RMSE beschrieben. RMSE (Root mean square error), ist die Wurzel der mittleren Fehlerquadratsumme, liegt zwischen 0 und 1, sollte möglichst klein sein und zeigt die Güte des Algorithmus an. Bei MSE (Mean square error) handelt es sich bei Regressionsanalysen um die Abweichung des Schätzers/ vorhergesagten Wertes vom realen Wert, je kleiner dieser ist, umso genauer sind die vorhergesagten Werte. Doch, bevor die Daten von den Algorithmen analysiert und ausgewertet werden, muss die Darstellung der Algorithmen festgelegt werden. Die Ergebnisse werden in zwei Diagrammen dargestellt. Zum einen wird die vorhergesagte RUL ins Verhältnis mit der realen RUL gesetzt, inklusive der Idealwertegeraden mit der Funktion y=x und zum anderen wird in einem Histogramm die Häufigkeit von vorhergesagten RULs (Yb) abgebildet. Auf Grund der Diagramme und der Algorithmen bezogenen Kennzahlen ist es möglich die Algorithmen und deren Vorhersagen vergleichen zu können.

Da die Ergebnisse und Kennzahlen der Algorithmen wie vorab schon beschrieben aufgrund der Kreuzvalidierung leicht schwanken können wurden zehn Prognosedurchläufe gemacht. Hierdurch soll sichergestellt werden, dass die gebildeten Mittelwerte über die Durchläufe genauer sind als bei nur einem Durchlauf. Das Verhalten der Prognosen über diese mehrfachen Durchläufe wird im Abschnitt 4.7 näher betrachtet. Im Folgenden werden zunächst alle Ergebnisse des ersten Durchlaufs analysiert.

Vor der Analyse der Algorithmen wird zunächst einmal das Histogramm mit der Werteverteilung der realen RULs erstellt. Das Histogramm ist abhängig von der Einteilung der Werte und kann je nachdem immer leicht unterschiedliche aussehen. Es ist zu erkennen das die meisten Werte in dem Bereich zwischen 0 und 60 liegen. Weniger Werte gibt es zwischen 120 und 220.

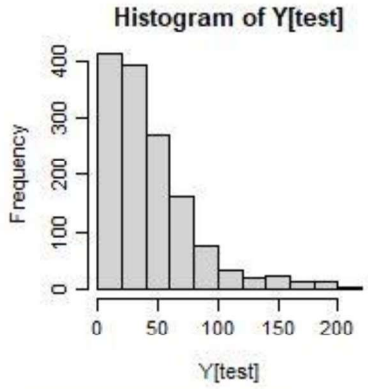

Abbildung 38: Histogramm der Verteilung der realen RULs

Der zuerst angewendete Algorithmus ist die lineare Regression. Hierbei handelt es sich um einen weniger komplexen Algorithmus. Bei der Betrachtung der Auswertungsdiagramme fällt eine Besonderheit auf, die bei keinem anderen Algorithmus vorkommen wird. Die Vorhersage des RUL vom Zeitpunkt 0 kann Werte bis -50 annehmen. Aus dem Histogramm ist zu entnehmen, dass sich dabei um ungefähr 100 Werte handelt. Diese den Werten zugehörigen Objekten müssten demzufolge bereits ausgefallen sein. Ebenfalls eine große Abweichung hat die lineare Regression bei der Vorhersage von einer RUL über 100. Fast alle Objekte mit einer realen RUL von über 100 haben laut der Vorhersage eine RUL von 50 bis 100. Trotz der Abweichungen bei der Vorhersage von unter 0 und über 100 ist die Vorhersage zwischen diesen beiden Werten, deutlich besser als vor und nach diesem Bereich. Im Mittel hat die lineare Regression einen MSE von 24,21, was so viel Bedeutet, wie das die vorhergesagte RUL im Mittel um 24,21 Zeiteinheiten von der realen RUL abweicht. Des Weiteren hat die lineare Regression einen RMSE von 0,5519. Der Gütewert ist mit 0,5519 ein leicht unterdurchschnittlicher Wert, der noch deutlich verbessert werden muss, damit dieser akzeptabel ist. Bei der Betrachtung des Histogramms fällt auf, dass die Werte vor allem im Bereich von 0 bis 80 liegen und in dem Betrachtungsbereich beinahe linear zunehmen. Wodurch keine Ähnlichkeiten zwischen der Realen RUL-Verteilung und den vorhergesagten Werten von der linearen Regression sind.

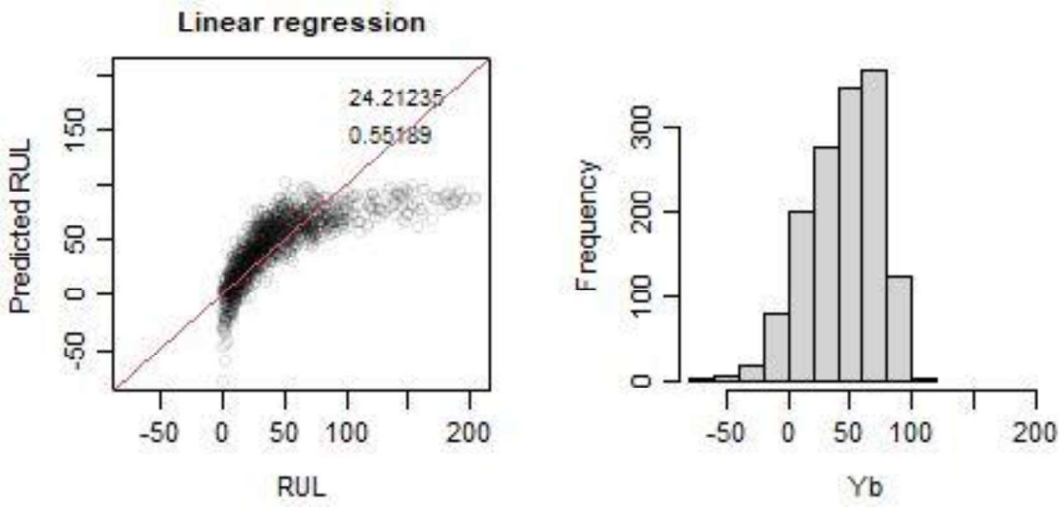

Abbildung 39: Ergebnisse lineare Regression

Der zweite verwendete Algorithmus ist das quadratische Model. Bei der Darstellung der Ergebnisse ist zu erkennen, dass die Werte bis 40 sehr gebündelt um den Idealwert sind, ab diesem Zeitpunkt fangen die Werte an zu streuen, sowohl positiv als auch negativ. Ab dem realen RUL von 100 steigt die Vorhersage nicht proportional mit an, sondern bleibt circa 100 bis maximal 110 stehen. Durch das Streuen um den Idealwert hat das Quadratische Modell einen MSE von 22,66, welcher besser ist als der der linearen Regression, ebenso wie der geringere und somit bessere RMSE-Gütewert von 0,5164. Das Histogramm zeigt auch, dass sich der Großteil der vorausgesagten Werte im Bereich von 0 bis 70 sammeln und die Werte

ab diesem Zeitpunkt wie eine quadratische Funktion abnehmen. Diese Auffälligkeit ist als eine leichte Übereinstimmung zu der realen Werteverteilung zu erkennen.

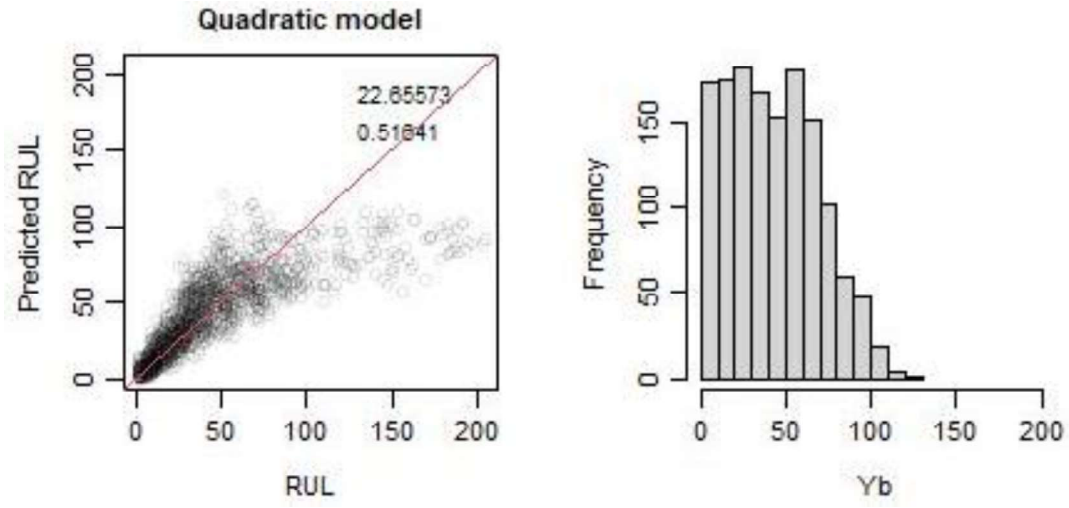

Abbildung 40: Ergebnisse des quadratischen Modells

Das Ergebnis des Algorithmus exponentielles Modell ist im Bereich 0 bis 40 sehr gut mit dem Ergebnis des quadratischen Modells zu vergleichen, da die Werte sehr geschlossen um den Idealwert liegen. Danach beginnt auch bei dem exponentiellen Modell die Streuung der Werte. Die Streuung ist sehr breit gefächert, da die Werte bis zu 200/90 bis 50/200 streuen. Somit ist die Vorhersage um 110 Zeiteinheiten zu niedrig bzw. 150 Zeiteinheiten zu hoch. Durch die große Streuung hat das exponentielle Model einen MSE von 24,35 und einen RMSE von 0,555, diese zwei Kennzahlen sind sehr ähnlich zu denen von der linearen Regression. Das Histogramm hat einen Verlauf von einem exponentiellen Zerfall. Es beginnt mit über 400 Werten nur im Bereich von 0 bis 20 und endet mit einer sehr geringen Anzahl zwischen 200 und 220.

Vom Verlauf her ist das Histogramm des exponentiellen Modells ähnlich zu dem der realen RUL-Werte, besonders durch die wenigen Werte zwischen 180 und 220.

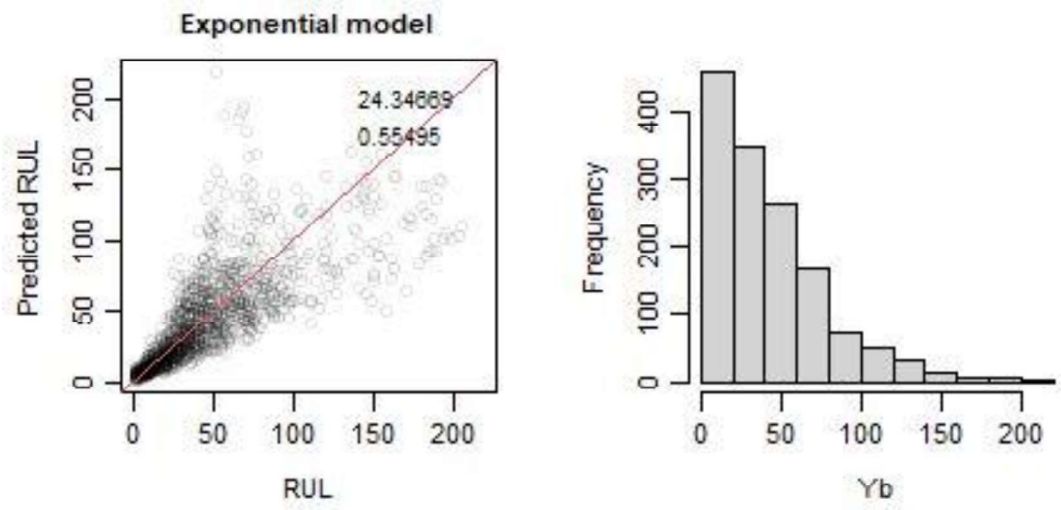

Abbildung 41: Ergebnis exponentielles Modell

Der vierte Algorithmus die SVM liefert bis zu einem RUL von 50 eine sehr gute Genauigkeit mit wenig Streuung und einer hohen Wertedichte. Im Bereich von 50 bis 100 verringert sich die Wertedichte etwas und der Schwerpunkt weicht langsam vom Idealwert ab, da die vorhergesehene RUL nicht hoch genug prognostiziert wird. Auch wenn in dem Bereich der Schwerpunkt der Werte nach rechts abweicht (vorhergesehene RUL ist zu niedrig), verlagert sich die Streuung eher in die linke Richtung, wodurch die reale RUL zu hoch eingeschätzt wird. Es werden maximal Werte bis zu 150 vorhergesagt. In dem Histogramm ist zu sehen, dass die ersten drei Klassen, mit Werten von 0 bis 60 einen Großteil der Gesamtmenge ausmachen, über einer RUL von 100 gibt es nur in etwa 100 Werte. Im Vergleich zu der realen Verteilung der RUL hat das SVM Histogramm starke Übereinstimmungen, die vielen Werte im Bereich bis einer RUL von 60, die ca. 170 Werte zwischen 60 und 80 und die stark abnehmenden Werte von der RUL 80 an. Final lässt sich sagen, dass die SVM den geringsten MSE mit 19,16 hat und den besten RMSE Gütewert des Durchlaufes mit 0,4366 aufweisen kann.

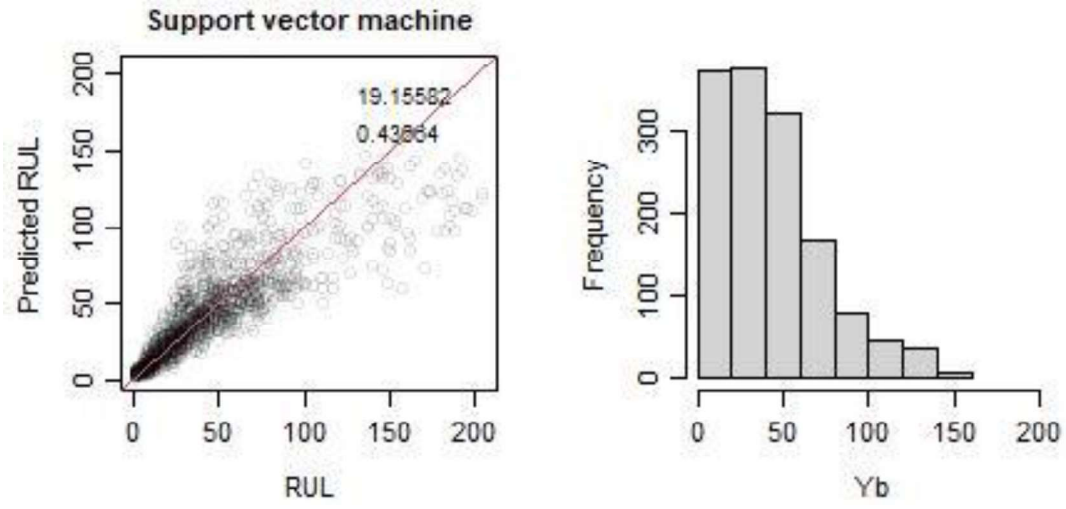

Abbildung 42: Ergebnis Support-Vector-Machine

Der nächste angewendete Algorithmus ist der sogenannte Random Forest. Bei der Betrachtung der Ergebnisse fällt auf, dass diese stark den Ergebnissen der SVM ähneln. Es gibt bis 50 eine Bündelung der Werte um den Idealwert, danach beginnt die Streuung der Werte mit der leichten Verschiebung vom Idealwert aus nach rechts. Der Unterschied zwischen den Ergebnissen der zwei Algorithmen ist, dass bei dem Random Forest die Werte bis ca. 150 gleichmäßiger streuen und ab diesem Wert fast ausschließlich nur noch zu niedrig eingeschätzt werden. Außerdem ist bei den vorhergesehenen Werten 130 bis 140 eine erhöhte Anzahl im Diagramm zu erkennen und das im linken, als auch rechten Bereich von der Idealwertelinie. Diese Auffälligkeit ist auch in dem Histogramm zu erkennen. Dennoch unterscheidet sich das Histogramm nicht groß von dem der realen RUL-Werte. Da die graphischen Ergebnisse sich nur geringfügig voneinander unterscheiden sind die Kennzahlen der Algorithmen ebenfalls sehr vergleichbar. Der Algorithmus Random Forest weist einen MSE von 19,93 und einen RMSE von 0,454 auf. Im Vergleich der beiden Werte mit den der SVM ist der Random Forest in beiden minimal schlechter.

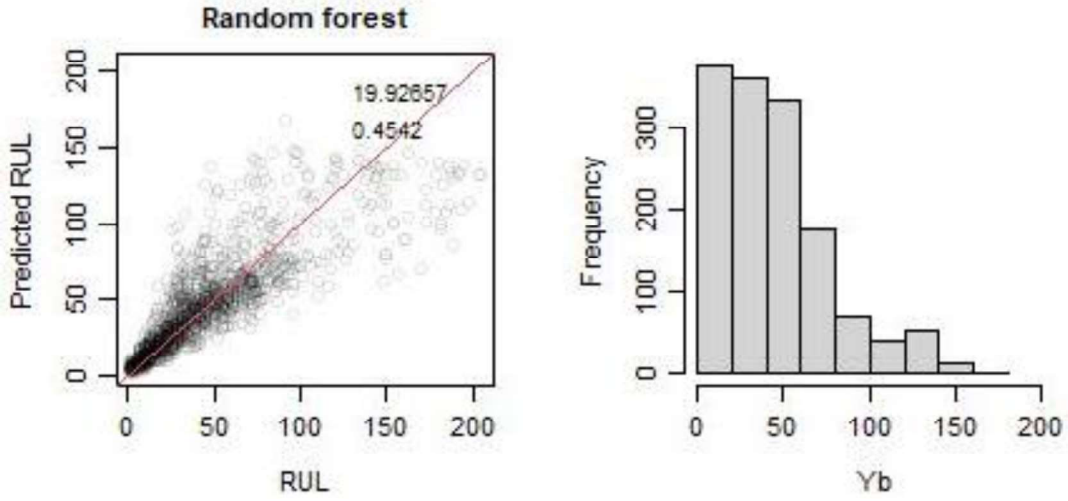

Abbildung 43: Ergebnis Random Forest

Die nächste Analyse erfolgte mit dem K-Nearest Neigbors Algorithmus. Auf dem ersten Blick ist gut zu erkennen, dass sich die Werte auch bei den vorhergesagten RULs bis 200 erstrecken. Nachdem die Werte auch hier bis ca. 50 sehr gebündelt um den Idealwert liegen, beginnen die Werte darauffolgend zu streuen. Auffällig ist hierbei aber, dass die Werte relativ gleichmäßig verteilt sind. Der MSE beläuft sich hier auf 25,82 und diese erhöhte Kennzahl hat einen einfachen Grund. Da die Werte hierbei zu beiden Seiten weiter abweichen, sind die Differenzen zwischen der Schätzung und dem realen Wert deutlich höher. Auch wenn die Streuung sehr ausgeglichen ist, zeigt der Gütewert RMSE mit 0,589, dass dieser Algorithmus weniger gut für diese Prognose geeignet ist. Bei dem Histogramm sind zudem keine Ausreißer in einer Klasse zu erkennen, was zudem auch die gleichmäßige Verteilung ohne Grenzwert belegt. Außerdem unterscheidet sich das Histogramm kaum von dem der realen RUL-Werte.

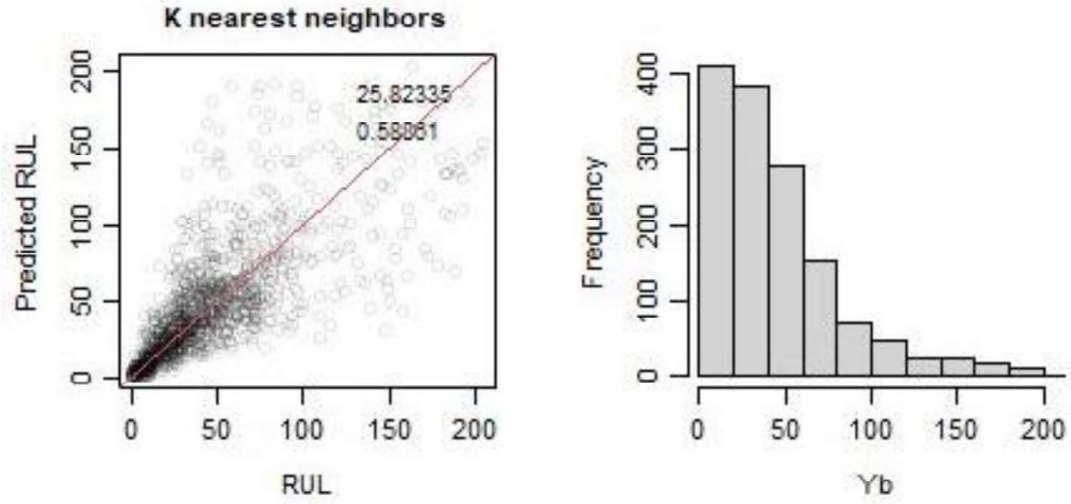

Abbildung 44: Ergebnis K-Nearest Neighbors

Der letzte angewandte Algorithmus ist das Neuronale Netzwerk, welches in den letzten Jahren immer weiterentwickelt wird. In dem Fall hat das neuronale Netzwerk einen niedrigen Standardfehler (MSE) mit 19,80 und den RMSE-Gütewert von 0,451. Diese Werte lassen sich gut in der Grafik nachvollziehen. Zum einen läuft der gebündelte Bereich zu Beginn bis 60 und in dieser Wolke liegen mehr als zwei Drittel aller Werte. Zum anderen, wenn die Streuung ab ca. 60 beginnt, ist diese nicht sehr weitläufig, oftmals ist die Abweichung unter 50. Dennoch gibt es nur wenige vorhergesagte RUL-Werte im Bereich um 150 und diese sind meistens auch niedriger als real prognostiziert. Außerdem ist im Prognosebereich um 50 eine leichte Grenze zu erkennen, wie im Histogramm auch. Das Histogramm unterscheidet sich zu dem realen Werte-Histogramm besonders durch die verstärkte Ausprägung im Bereich 40-60.

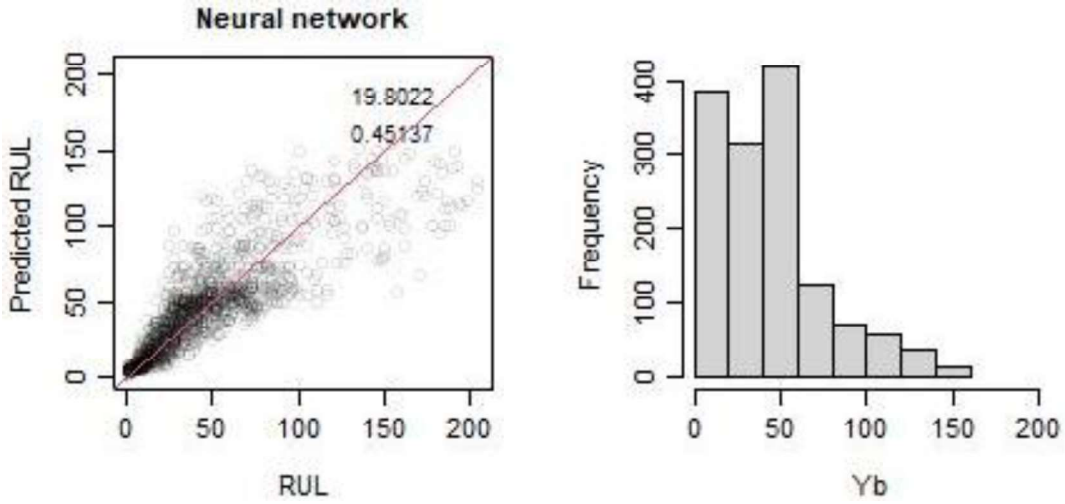

Abbildung 45: Ergebnis Neuronales Netzwerk

Über alle Algorithmen kann auch gut ein Mittelwert gezogen werden, welcher sehr viel Ähnlichkeit mit dem neuronalen Netzwerk hat. Es gibt eine Bündelung der Werte bis zu einer vorhergesagten RUL von 60. Die danach auftretende Streuung ist nicht sehr weitläufig und die Streuung beläuft sich tendenziell eher auf den rechten Bereich der Idealwertegeraden. Mit einem MSE-Wert von 19,21 und einem RMSE-Wert von 0,438 ist der Mittelwert aller Algorithmen um einiges genauer als viele der einzelnen Algorithmen. Wie auch bei dem NN sind ein Großteil der Werte im Bereich von 0 bis 60 angesiedelt und die restlichen Werte liegen, bis auf wenige Ausreißer, im Bereich bis 140. Wenn das Histogramm, mit dem der realen Werte verglichen wird, ist der eine große Ähnlichkeit zu erkennen, bis auf zwei auffällige Unterschiede. Zum einen ist der Bereich 40 bis 60 zu stark ausgeprägt, wie bei dem NN auch schon. Zum anderen liegen die Werte höchstens im Fenster 140 bis 160 und bei dem realen Histogramm reichen die Werte bis in den Bereich 200/220.

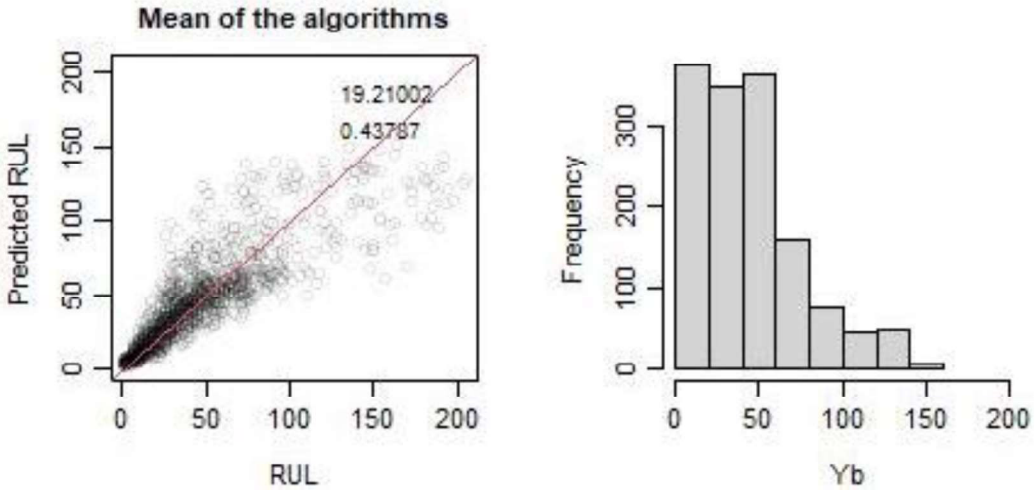

Abbildung 46: Ergebnis Mittelwert der Algorithmen

4.7 Auswertung der Ergebnisse

Nachdem der erste Durchlauf der Algorithmen etwas näher betrachtet wurde, wird in den folgenden Tabellen 5 und 6, sowie in den Abbildungen 47 und 48 der Gütewert aller Algorithmen analysiert.

Zunächst einmal wurden alle Kennzahlen (MSE und RMSE) der zehn Durchläufe der Algorithmen in den Tabellen 5 und 6 gesammelt. Es ist gut zu erkennen, dass die Werte aller Algorithmen schwanken. Die MSE-Werte schwanken teilweise bis zu 30 zehntel. In den meisten Fällen aber nur um wenige zehntel. Bei den RMSE-Werten sind die Schwankungen deutlich geringer, mit Schwankungen im tausendstel Bereich. Die größte Veränderung gab es bei dem Algorithmus K-nearest-neighbors mit 4 hundertstel. In den nachfolgenden Tabellen 5 und 6 sind die durchschnittlichen RMSE-Werte aufgeführt. Im Vergleich mit dem ersten

Durchlauf ist zu erkennen, dass dieser Durchlauf bei vielen Algorithmen den höchsten RMSE erzeugt hat und somit der schlechteste Durchlauf war.

Tabelle 5: Kennzahlen der 10 Durchläufe; Algorithmen 1-4

	linear		quadratisch		exponentiell		SVM	
	MSE	RMSE	MSE	RMSE	MSE	RMSE	MSE	RMSE
Durchlauf 1	24,212	0,552	22,656	0,516	24,347	0,555	19,156	0,437
Durchlauf 2	24,119	0,536	22,689	0,505	24,631	0,548	18,757	0,417
Durchlauf 3	22,618	0,534	20,771	0,491	24,220	0,572	17,567	0,415
Durchlauf 4	25,464	0,535	24,300	0,510	27,161	0,570	19,424	0,408
Durchlauf 5	24,671	0,545	23,262	0,514	25,463	0,562	18,986	0,419
Durchlauf 6	24,442	0,540	23,096	0,510	25,596	0,566	18,903	0,418
Durchlauf 7	24,319	0,537	23,004	0,508	25,456	0,562	18,909	0,418
Durchlauf 8	24,265	0,538	22,883	0,507	25,206	0,559	18,934	0,420
Durchlauf 9	24,446	0,540	23,075	0,510	25,263	0,558	19,023	0,420
Durchlauf 10	24,595	0,544	23,165	0,512	25,428	0,562	18,947	0,419
Durchschnitt	**24,315**	**0,540**	**22,890**	**0,508**	**25,277**	**0,561**	**18,861**	**0,419**

Tabelle 6: Kennzahlen der 10 Durchläufe; Algorithmen 5-8

	RF		KNN		NN		Mittelwert	
	SF	GW	SF	GW	SF	GW	SF	GW
Durchlauf 1	19,927	0,454	25,823	0,589	19,802	0,451	19,210	0,438
Durchlauf 2	20,059	0,441	27,329	0,600	19,483	0,433	19,554	0,429
Durchlauf 3	18,511	0,437	24,619	0,581	18,423	0,435	17,741	0,419
Durchlauf 4	20,027	0,421	27,610	0,580	20,326	0,427	19,505	0,410
Durchlauf 5	19,438	0,439	26,265	0,593	19,697	0,435	18,702	0,422
Durchlauf 6	19,654	0,443	26,972	0,607	19,686	0,435	18,848	0,424
Durchlauf 7	19,330	0,436	25,124	0,567	19,673	0,435	18,810	0,425
Durchlauf 8	19,455	0,433	26,067	0,581	19,631	0,435	18,903	0,421
Durchlauf 9	19,057	0,430	26,418	0,597	19,723	0,436	18,520	0,418
Durchlauf 10	18,811	0,423	25,133	0,565	19,659	0,435	18,593	0,418
Durchschnitt	**19,427**	**0,436**	**26,136**	**0,586**	**19,610**	**0,436**	**18,839**	**0,422**

Doch daraus stellt sich eine Frage, verhalten sich alle Algorithmen gleich? Dazu wurde in der Abbildung 47 die Entwicklung der RMSE-Werte dargestellt. Hierzu wurden die Differenzen der Durchläufe gezogen, um vergleichbare Reaktionen der Algorithmen auf veränderte Datenbereiche zu beweisen bzw. widerlegen zu können. Ins Auge fällt hier besonders die Entwicklung des KNN. Zum einen ist die Reaktion des Algorithmus zum Teil gegenläufig zu den meisten anderen Algorithmen und zum anderen sind die Veränderungen des RMSE deutlich größer als die der anderen, siehe alles ab Durchlauf 6 bis 10. Ebenfalls sehr starke Ausreißer haben die Algorithmen mit dem quadratischen und exponentiellen Modell, zwischen den Durchläufen 2&3 bzw. 3&4. Einen sehr unauffälligen durchschnittlichen

Verlauf haben die Algorithmen SVM und NN sowie der Mittelwertalgorithmus. Durch den durchschnittlichen Verlauf liegen die Algorithmen im Trend. Ob die Algorithmen statistisch signifikant sind, wird in dieser Masterthesis nicht näher betrachtet.

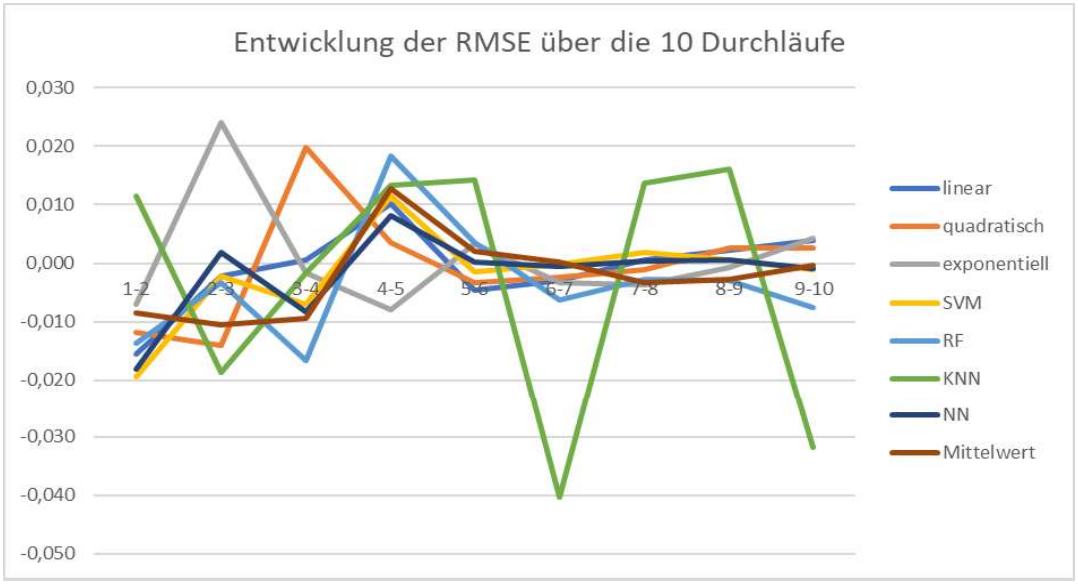

Abbildung 47: Entwicklung der RMSE über die zehn Durchläufe

Neben der Entwicklung der RMSE-Werte in den zehn Durchläufen wurde für jeden Algorithmus ein vereinfachtes Boxplot Diagramm erstellt, wie in der Abbildung 48 zu sehen ist. In dem Boxplot sind die Spannweite und Lage der RMSE Werte eines Algorithmus abgebildet, sowie der Mittelwert. Es ist gut zu erkennen, dass Werte sich in zwei Bereiche gliedern lassen, indem bei einem RMSE-Wert von 0,480 eine Grenze gezogen wird. Vier Boxplots befinden sich unterhalb und vier oberhalb der Grenze. Die vier Algorithmen mit höheren RMSE-Werten sind der lineare, quadratische und exponentielle Algorithmus, ebenso wie der KNN. Systematisch gesehen sind die anderen drei Algorithmen und der Mittelwert deutlich besser und genauer. Welcher der Algorithmen aber der beste und genauste ist, lässt sich nicht eindeutig vorhersagen. Die Ergebnisse liegen sehr nah beieinander, somit könnte nach dem besten Mittelwert oder der niedrigsten Spannweite entschieden werden. Nach den Aspekten wäre die SVM oder das NN am besten.

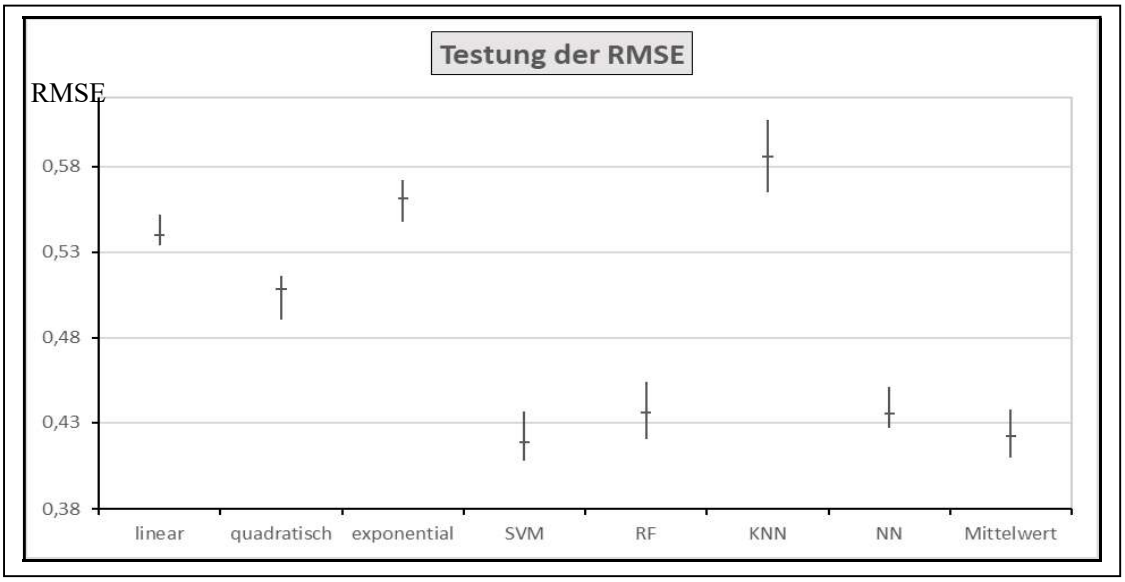

Abbildung 48: Testung der RMSE über zehn Durchläufe

Es gibt noch weitere Faktoren, die es ermöglichen würden, die Ergebnisse der Algorithmen zu verbessern. Einer der Faktoren ist eine relevante Stellschraube direkt zu beginn. Es gab insgesamt 21 unterschiedliche Messsensoren, von denen aber nur 15 verwendet werden konnten, da die restlichen 6 immer denselben Wert geliefert, und dadurch keine Information geliefert haben. Wenn diese 6 Sensoren umpositioniert, anders ausgerichtet oder genauer messen würden, würden mehr Daten vorliegen und Einflüsse eingefangen werden. Allein nur durch die Stellschraube Sensoren ist ein verbessertes Ergebnis möglich. Ein weiterer Faktor ist die Datenmenge. Die Daten waren in dem Fall auf einen Datensatz begrenzt, aus diesem Grund war die Kreuzvalidierung mit der Datensatzaufteilung notwendig. Wenn aber diese zunächst ungenauere Auswertung in der Praxis angewendet wird, nimmt der Datensatz immer weiter zu und die Algorithmen lernen besser, aufgrund von mehr Trainingsdaten. Der dritte Faktor ist jener der sich mit den Algorithmen beschäftigt. Es wurden in dieser Masterthesis nur die herkömmlichen Standardalgorithmen und Regressionen angewandt. Es ist folglich möglich, dass ein anderer Algorithmus genauer ist. Außerdem ist die Kombination mehrerer Algorithmen in anderer Form als hier, mit dem Mittelwert, möglich.

5. Fazit

Für die industriellen Unternehmen ist der Bereich der Instandhaltung unabdingbar und die Bedeutung nimmt immer weiter zu. Neben steigenden Preisen für Ersatzteile und Rohstoffe nehmen die Lohnkosten immer weiter zu. Umso teurer wird ein Ausfall einer Anlage. Durch effiziente Prozesse bleibt ein Unternehmen wettbewerbsfähig. Wenn ein Unternehmen dann auch dem Standard der Industrie 4.0 folgt, nimmt die Vernetzung und Menge an Daten immer weiter zu. Weshalb sich für diese Unternehmen eine zustandsorientierte Instandhaltung gut anbietet unter Berücksichtigung unterschiedlicher Faktoren und Einflüsse. Zum einen ist wichtig, dass es sich lohnt diesen Aufwand zu haben und diese Investitionen für die zustandsorientierte Instandhaltungsstrategie zu tätigen. Aus diesem Grund ist diese Strategie nicht immer die Strategie der Wahl, sondern eher die reaktive, bei kleineren unkritischen Bauteilen oder auch vorbeugende Instandhaltungsstrategie bei kritischen Objekten oder Bauteilen. Wenn Ausfälle generell vermieden werden wollen, ist der Einsatz von Redundanzen eine gute Option. Nachdem diese Aspekte betrachtet wurden, müssen die Rahmenbedingungen für die zustandsorientierte Instandhaltung gesetzt werden, damit diese möglichst effektiv und gut anwendbar ist.

Ein wichtiger Bereich der zustandsorientierten Instandhaltung ist die Datenanalyse und Auswertung. Doch, bevor die Daten überhaupt ausgewertet werden können, müssen brauchbare, informative und auch ausreichend Daten erfasst werden. Wenn diese Anforderungen erfüllt werden, geht es an die Datenanalyse, welche hier an dem Beispiel eines Datensatzes zu Triebwerken mit der Software R durchgeführt wurde. Es hat sich gezeigt, dass die Datensatzanalyse und die darauf basierende RUL Prognose komplex ist. In der Analyse kamen Ergebnisse mit dem niedrigsten Standardfehler von 19 Betriebszyklen und einem Gütewert von 0,437 bei der Support-Vector-Machine heraus. Für den Anfang sind solche Werte gut, für die Praxis in der Industrie nicht, hier sollten die Vorhersagen um einiges genauer sein. Um das Ziel zu erreichen, müssen die Einflussfaktoren bestimmt und angepasst werden. Hierzu zählen das bessere Einstellen von Sensoren, das Sammeln von mehr Daten, die kontinuierliche Anpassung und Optimierung der Analyse und die Zeit. Je länger Daten gesammelt werden, desto mehr Daten liegen vor und der Erfahrungsschatz steigt.

Wie nun verdeutlicht wurde ist die zustandsorientierte Instandhaltungsstrategie eine sehr gute Option, die wenn sie funktioniert die Effizienz in der Produktion steigern und Kosten einsparen kann. Doch dieser Prozess ist nicht von heute auf morgen möglich, es werden Zeit, Geduld, Anpassungen und bedingt auch Investitionen benötigt, um den größtmöglichen Nutzen aus der zustandsorientierten Instandhaltungsstrategie zu bekommen.

Literatur- und Quellenverzeichnis

Literatur

Ahrens, Daniela. Neue Anforderungen im Zuge der Automatisierung von Produktionsprozessen: Expertenwissen und operative Zuverlässigkeit, in: AIS-Studien, 9(1), 2016. S.43-56.

Amruthnath Nagdev, Gupta Tarun. A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance, 2018 5th International Conference on Industrial Engineering and Applications (ICIEA), 2018. S. 355-361. DOI:10.1109/IEA.2018.8387124.

Bracke S. Vorlesung Methoden der Mess- und Prüfdatenanalyse. Veranstaltung 2. 2012. Folie 4. Bergische Universität Wuppertal.

Bundesministerium für Arbeit und Soziales. Weißbuch - Arbeiten 4.0. 2017.

Gölzer, P. (2017). *Big Data in Industrie 4.0-Eine strukturierte Aufarbeitung von Anforderungen, Anwendungsfällen und deren Umsetzung* (Doctoral dissertation, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)).

Görz, G., & Schneeberger, J. (Eds.). (2010). *Handbuch der künstlichen Intelligenz*. Walter de Gruyter.

Haenelt, Karin. *Hidden Markov Models*, Kursfolien 2006.Centrum für Informations- und Sprachverarbeitung - Universität München.

Hahn, Hans-Werner. *Die industrielle Revolution in Deutschland*. De Gruyter Oldenbourg. Berlin. 2011.

Heller, T., & Prasse, C. (2017). *Total Productive Management-ganzheitlich: Einführung in der Praxis*. Springer-Verlag.

Informationen zur politischen Bildung /izpb, *Digitalisierung,* Nr. 344, 3/2020.

Kaminske, Gerd F. *Handbuch QM-Methoden*; Carl Hanser Verlag GmbH & Co. KG. München. 2012.

Klein, D., Tran-Gia, P., & Hartmann, M. (2013). Big Data. Informatik-Spektrum, 36(3), 319-323.

Krupitzer, C., Wagenhals, T., Züfle, M., Lesch, V., Schäfer, D., Mozaffarin, A., ... & Kounev, S. (2020). A survey on predictive maintenance for industry 4.0. *arXiv preprint arXiv:2002.08224*.

Lasi Heiner, Fettke Peter, Feld Thomas; Hoffmann Michael. Industrie 4.0, in: Wirtschaftsinformatik 04/2014. Springer Verlag. Wiesbaden. 2014.

Lois, D. Lineare Regression: Grundlagen und BLUE-Annahmen, Kursfolien Universität der Bundeswehr München, Fakultät für Humanwissenschaften, Sozialwissenschaftliche Methodenlehre, 2015.

Löwer Manuel. Vorlesung: Qualitätsvorausplanung in der Entwicklung. 2019. Veranstaltung: 6 FMEA – Failure Mode & Effects Analysis und Veranstaltung: 7 FTA – Fault Tree Analysis. Bergische Universität Wuppertal.

Niermeyer, R. (2007). *Motivation: Instrumente zur Führung und Verführung* (Vol. 195). Haufe-Lexware.

O.A. KI fit machen für sicherheitskritische Anwendungen, in: Qualität und Zuverlässigkeit, Seite 6. Bd. 7/2021.

Serradilla, O., Zugasti, E., & Zurutuza, U. (2020). Deep learning models for predictive maintenance: a survey, comparison, challenges and prospect. *arXiv preprint arXiv:2010.03207*.

Spath, Dieter. *Produktionsarbeit der Zukunft-Industrie 4.0*. Fraunhofer Verlag. Stuttgart. 2013.

Tordeux Antoine. Zuverlässigkeit von Kompressoren in Produktionsanlagen. Lehrvortrag - W3-Professur für Reliability Engineering an der TU Dortmund. Folie 16

Tordeux Antoine. Vorlesung Technische Zuverlässigkeit. 2019. Bergische Universität Wuppertal

Verma N. K., A. Salour. *Intelligent Condition Based Monitoring.* Springer Verlag 2016.

Yu, W., Dillon, T., Mostafa, F., Rahayu, W., & Liu, Y. (2019). A global manufacturing big data ecosystem for fault detection in predictive maintenance. *IEEE Transactions on Industrial Informatics*, *16*(1), 183-192.

DIN:

DIN 1319-1; 1995-01. Grundlagen der Meßtechnik – Teil 1: Grundbegriffe. Beuth Verlag. Berlin

DIN EN 13306:2018-02. Instandhaltung – Begriffe der Instandhaltung. Beuth Verlag. Berlin.

DIN EN 31051:2019-06. Grundlagen der Instandhaltung. Beuth Verlag. Berlin

DIN EN 61709:2015-01. Elektrische Bauelemente – Zuverlässigkeit – Referenzbedingungen für Ausfallraten und Beanspruchungsmodelle zur Umrechnung. Beuth Verlag. Berlin.

DIN EN ISO 13849-1:2016-06. Sicherheit von Maschinen – Sicherheitsbezogene Teile von Steuerungen – Teil 1 Allgemeine Gestaltungsansätze. Beuth Verlag. Berlin.

DIN EN ISO 61124:2013-01. Prüfungen der Funktionsfähigkeit – Prüfpläne für konstante Ausfallrate und konstante Ausfalldichte. Beuth Verlag. Berlin

DIN EN ISO 9000: 2015. Qualitätsmanagementsysteme – Grundlagen und Begriffe. Beuth Verlag. Berlin.

VDI 4003: 2007. Zuverlässigkeitsmanagement. Beuth Verlag. Berlin

Internetquellen:

2D- und 3D-Plots - MATLAB & Simulink - MathWorks Deutschland. (o. D.). MATLAB - MathWorks. Abgerufen am 1. September 2021, von https://de.mathworks.com/help/matlab/learn_matlab/plots.html.

Aschermann, T. (2020, 25. Juni). Lebensdauer von Bremsscheiben und Bremsbelägen - alle Infos. FOCUS.de. Abgerufen am 25. August 2021, von https://praxistipps.focus.de/lebensdauer-von-bremsscheiben-und-bremsbelaegen-alle-infos_99291.

Comprehensive R Archive Network (CRAN). (2019, 7. Februar). *CRAN - Package neuralnet*. Abgerufen am 1. September 2021, von cran.r-project.org. https://cran.r-project.org/web/packages/neuralnet/index.html.

Condition Monitoring. (o. D.). Harmonic Drive SE. Abgerufen am 4. Oktober 2021, von https://harmonicdrive.de/de/glossar/condition-monitoring%20/.

Deppe, Kerstin. (2021, 09. April). Arbeit 4.0 – Geschichte der Arbeit. Abgerufen am 20. Juli 2021, von https://www.planet-wissen.de/gesellschaft/arbeit/die_geschichte_der_arbeit/arbeit-vier-punkt-null-100.html.

Farwer, L. (2018, 9. September). Lebensdauer einer Kupplung: Einschätzung und Ursachen. FOCUS.de. Abgerufen am 04. Oktober 2021, von https://praxistipps.focus.de/lebensdauer-einer-kupplung-einschaetzung-und-ursachen_97737#:%7E:text=Kupplung%20im%20Auto%20-%20so%20lange%20k%C3%B6nnen%20Sie,falsche%20Nutzung%20der%20Kupplung.%20. . .%20Weitere%20Artikel. . .%20.

Frenzel, S. F. (2021, 11. Januar). *Machine Learning mit unserer Crowd | Trainingsdaten*. Crowd Guru. Abgerufen am 25. August 2021, von https://www.crowdguru.de/blog/guide-machine-learning-mit-hochwertigen-trainingsdaten/.

Grünwald, R. (2019, 29. Juli). *So vermeidet man Overfitting und Underfitting! - NOVUSTAT* Statistik-Blog. Statistik Service. Abgerufen am 1.September 2021, von https://novustat.com/statistik-blog/overfitting-und-underfitting.html.

Grünwald, R. (2020, 2. Dezember). *Lineare Regression einfach erklärt - NOVUSTAT Statistik-Blog*. Statistik Service. Abgerufen am 15. September 2021, von https://novustat.com/statistik-blog/klassiker-lineare-regression-einfach-erklaert.html.

IBM. *Nicht lineare Regression*. (o. D.). ibm. Abgerufen am 4. Oktober 2021, von https://www.ibm.com/docs/de/spss-statistics/SaaS?topic=regression-nonlinear.

Instandhaltung: Präventive und Korrektive Instandhaltung. (o. D.). MaintMaster Systems. Abgerufen am 4. Oktober 2021, von https://maintmaster.com/de/instandhaltungshandbuch/instandhaltung/

Kagermann Henning, Lukas Wolf-Dieter& Wahlster Wolfgang. (2021, April). Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution. Abgerufen am 01. August 2021, von https://web.archive.org/web/20130304101009/https://vdi-nachrichten.com/artikel/Industrie-4-0-Mit-dem-Internet-der-Dinge-auf-dem-Weg-zur-4-industriellen-Revolution/52570/1.

Luber, S. (2018, 18. Juli). *Was ist der k-Means-Algorithmus?* BigData-Insider. Abgerufen am 15. September 2021, von https://www.bigdata-insider.de/was-ist-der-k-means-a-734637/.

Luber, S. (2019, 19. März). *Was ist ein Neuronales Netz?* BigData-Insider. Abgerufen am 15. September 2021, von https://www.bigdata-insider.de/was-ist-ein-neuronales-netz-a-686185/.

Luber, S. (2019, November 6). *Was ist eine Support Vector Machine?* BigData-Insider. Abgerufen am 15. September 2021, von https://www.bigdata-insider.de/was-ist-eine-support-vector-machine-a-880134/.

MATLAB - MathWorks. (o. D.). MATLAB & Simulink. Abgerufen am 1. September 2021, von https://de.mathworks.com/products/matlab.html.

R-Project. *R: What is R?* (o. D.). www.r-Project.Org. Abgerufen am 4. Oktober 2021, von https://www.r-project.org/about.html.

Ripley, B. (2021, 3. Mai). *CRAN - Package MASS*. cran.r-project.org. Abgerufen am 1. September 2021, von https://cran.r-project.org/web/packages/MASS/index.html.

Ruckstuhl, A. (2008) Einführung in die nichtlineare Regression. ZHAW Zürcher Hochschule für angewandte Wissenschaften.

Saxena, A. & Goebel K. (2008). Turbofan Engine Degradation Simulation Data Set. Zip-Ordner. Textdatei: Read me. PDF-Datei: Damage Propagation Modeling. Datensatz: TrainFD001. Abgerufen am: 1. Juni 2021, von ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/.

Seite „Ausfallverteilung“. In: Wikipedia – Die freie Enzyklopädie. Bearbeitungsstand: 3. April 2019, 04:57 UTC. Abgerufen am 4. Oktober 2021, von https://de.wikipedia.org/w/index.php?title=Ausfallverteilung&oldid=187179128.

Seite „Druckmessgerät“. In: Wikipedia – Die freie Enzyklopädie. Bearbeitungsstand: 11. April 2021, 01:12 UTC. Abgerufen am 4. Oktober 2021, von https://de.wikipedia.org/w/index.php?title=Druckmessger%C3%A4t&oldid=210804021.

Seite „Hierarchische Clusteranalyse“. In: Wikipedia – Die freie Enzyklopädie. Bearbeitungsstand: 28. Mai 2021, 08:54 UTC. Abgerufen am 4. Oktober 2021, von https://de.wikipedia.org/w/index.php?title=Hierarchische_Clusteranalyse&oldid=212450898 ().

Seite „Industrielle Revolution“. In: Wikipedia – Die freie Enzyklopädie. Bearbeitungsstand: 11. September 2021, 22:05 UTC. Abgerufen am 4. Oktober 2021, von https://de.wikipedia.org/w/index.php?title=Industrielle_Revolution&oldid=215512074.

Seite „Mean Distance Between Failure“. In: Wikipedia – Die freie Enzyklopädie. Bearbeitungsstand: 22. April 2016, 13:46 UTC. Abgerufen am 4. Oktober 2021, von https://de.wikipedia.org/w/index.php?title=Mean_Distance_Between_Failure&oldid=153702684.

Semmelmann, K. (2020, 29. August). *Was ist künstliche Intelligenz (KI)? Definition und Beispiele*. Data Driven Company. Abgerufen am 25. August 2021, von https://datadrivencompany.de/was-ist-kuenstliche-intelligenz-ki-definition-und-beispiele/.

Standish, R. S. (2017, 6. April). *SciDAVis: About*. Abgerufen am 24. Juli 2021, von http://scidavis.sourceforge.net & http://scidavis.sourceforge.net/about.html.

Tarkowski, P. (2018, 28. Februar). Digitalisierung: Was ist das? Eine Definition. digital-magazin.de. Abgerufen am 24. Juli 202, von https://digital-magazin.de/digitalisierung-definition/.

The GNU Operating System and the Free Software Movement. (o. D.). Gnu.Org. Abgerufen am 22. Juni 2021, von http://www.gnu.org/.

Redaktion Focus Online. (2018, 19. Februar). Wann den Keilriemen wechseln? Das müssen Sie beachten. FOCUS.de. Abgerufen am 04.10.2021, von https://praxistipps.focus.de/wann-den-keilriemen-wechseln-das-muessen-sie-beachten_99698.

Was sind Schwache KI und Starke KI? Künstliche Intelligenz Erklärung. (2019, 3. April). SimplyKI. Abgerufen am 4. Oktober 2021, von https://simplyki.de/was-sind-schwache-ki-und-starke-ki-kuenstliche-intelligenz-erklaerung/.

Wendl, M. (2021, 20. April). Nachteile Digitalisierung – diese 6 Punkte sollten Sie beachten. faktur:digital. Abgerufen am 4. Oktober 2021, von https://fakturdigital.de/nachteile-digitalisierung/.

Wirtschaftslexikon24. *Restnutzungsdauer - Wirtschaftslexikon*. (o. D.). Wirtschaftslexikon24. Abgerufen am 4. Oktober 2021, von http://www.wirtschaftslexikon24.com/d/restnutzungsdauer/restnutzungsdauer.html.

Wissen.de. Fail-safe-Prinzip. (o.D.). Wissen.de. Abgerufen am 24. August, von wissen.de/lexikon/fail-safe-prinzip.

Wright M. (2019, 07 Februar) neuralnet: Training of Neural Networks. R-Project.com. Abgerufen am 25. September 2021, von cran.r-project.org/web/packages/neuralnet/index.html.

Wünsche Veikko (o. D.). *Sensoren: Arten & Einsatz*. wiki.induux.de. Abgerufen am 4. Oktober 2021, von https://wiki.induux.de/Kategorie:Sensoren.

Anlage

```
## Alle Variablen löschen
rm(list())

##Pakete installieren und/oder aktivieren
library(neuralnet)
library(MASS)
library(corrplot)
library('depmixS4')
library(smooth)
library(rpart)
library(e1071)
library(class)
library(randomForrest)
library(LPCM)
```

```
## Feststellen des Arbeitsverzeichnisses
setwd(„C:/Users/LischkaNico/Desktop/Masterthesis/R)

##Lesen der Daten
X=read.table("Datensatz.txt")
exp="Datenanalyse" ## später benutzt für die graphische Darstellung
```

Zu der Abbildung 20

```
##Triebwerksnummer
unit=X[,1]
##Betriebszyklusnummer
time=X[,2]
##Sensormessungen
X=X[,6:26]
names(X)=paste("S",1:21,sep=" ")

##Auslesen der Sensoren, die sich mit der Zeit ändern
select=(1:ncol(X))[apply(X,2,var)>0]
X=X[,select]

##Grafische Darstellung
layout(matrix(c(1,2,3,4,5,6,7,7,8),by=T,ncol=3))
par(mar=c(4,4,3,2))
par(mgp=c(2.5,.75,0))

##Graphische Darstellung der Standardabweichungen der Sensorvariablen
plot(apply(X,2,sd),main=paste("\n Initial
variables"),pch=1:ncol(X),col=1:ncol(X),cex.main=1.1,ylab="Standard
deviation",xlab="Sensor X")
```

Zu der Abbildung 21

```
##Grafische Darstellung der Autokorrelation der Sensorvariablen
par(mfrow=c(1,1))
par(mar=c(4,4,3,2))
corrplot(cor(X),tl.cex=1,tl.col="grey",title=paste(" "))
text(8,0,"\n Autocorrelation of the sensors")
Sys.sleep(5)
```

```
##Totale Betriebsdauer (UL) und verbleibende Betriebsdauer (RUL) bestimmen
ULv=NULL
UL=NULL
RUL=NULL
For(u in unique(unit)){
        ULv=c(ULv,max(time[unit==u]))
        UL=c(UL,rep(max(time[unit==u]),length(time[unit==u])))
        RUL=c(RUL,max(time[unit==u])+1-time[unit==u])}
```

Zu der Abbildung 22

```
##Graphische Darstellung der Korrelation zwischen Sensormessungen und der RUL
Plot(cor(X,RUL),main=paste(“”),pch=select,col=select,cex.main=1.1,ylab=”cor(X,RUL)”,
xlab=”SensorX”,xaxt=’n’,yaxt=’n’);axis(2);(1,1:ncol(X),select)
```

Zu der Abbildung 23

```
##Graphische Darstellung der korrelierten Variablen 4 und 11
ordX=order(abs(cor(RUL,X)),decreasing=T)
cc=sample(1:nrow(X),nrow(X))
plot(X[,ordX[1]][cc],X[,ordX[2]][cc],col=rgb(1-
RUL[cc]/UL[cc],0,RUL[cc]/UL[cc],.1),main=”2 more correlated\n
variables”,cex.main=1.1,xlab=paste(“X”,select[ordX[1]],sep=””),
ylab=paste(“X”,select[ordX[2]],sep=””))
```

Zu den Abbildungen 24 bis 26

```
##Dimensionen durch eine Hauptkomponentenanalyse reduzieren
PC=prcomp(X,scale=T)$x
head(PC)
apply(PC,2,sd)^2
K=apply(PC,2,sd)^2
Sum(K)
J=cumsum(K)
Gesamtvarianz=K/sum(K)*100

##Grafische Darstellung
plot(apply(PC,2,sd),main=”\n Principal components”,cex.main=1.1,ylab=Standard
deviation”,xlab=”PC”)
plot(Gesamtvarianz, main=\n Percentage of PC variance in the total variance”, type=”b”,
pch=select, col=select, cex.main=1.1, ylab=”Proportional variance in %”,xlab=”PC”)
plot(J, main=”\n Variance per PC”, type=”S”,pch=select, col=select, cex.main=1.1,
ylab=”Variance”, xlab=”PC”)
```

Zu der Abbildung 27

```
##Graphische Darstellung der Korrelation zwischen den PCs und der RUL
plot(cor(PC,RUL), cex.main=1.1, ylab=”cor(PC,RUL)”, xlab=”PC”)
```

Zu den Abbildungen 28 und 29

```
##Graphische Darstellung der Korrelation von PC1 und PC2
ord=order(abs(cor(PC,RUL)),decreasing=T)
cc=sample(1:nrow(PC),nrow(PC))
plot(PC[,ord[1]][cc],PC[,ord[2]][cc],col=rgb(1-RUL[cc]/UL[cc],0,RUL[cc]/UL[cc],.1),
main="2 more correlated\n principal components",cex.main=1.1,
xlab=paste("PC",ord[1],sep""),ylab=paste("PC",ord[2], sep=""))

##Graphische Darstellung der Korrelation von PC9 und PC11
plot(PC[,ord[4]][cc],PC[,ord[3]][cc],col=rgb(1-RUL[cc]/UL[cc],0,RUL[cc]/UL[cc],.1),
cex.main=1.1, xlab=paste("PC",ord[4],sep""),ylab=paste("PC",ord[3], sep=""))
```

Zu den Abbildungen 30 und 31

```
##Höchste korrelierte Hauptkomponente als Funktion der Zeit für jede Einheit
par(mar=c(4,4,1,2))
plot(NA,xlim=range(time),ylim=range(PC[,ord[1]]),xlab="time",ylab=paste("More
correlated PC (PC",ord[1],")",sep=""))
for(u in unique(unit))
lines(PC[unit==u,ord[1]],col=u)

##Höchste korrelierte Hauptkomponente als Funktion der Zeit für jede Einheit
plot(NA,xlim=range(time),ylim=range(PC[,ord[1]]),xlab="time",ylab=paste("More
correlated PC (PC",ord[1],")",sep=""))
for(u in unique(unit)){
        lines(c(1,length(PC[unit==u,ord[1]])),PC[unit==u,ord[1]]
[c(1,length(PC[unit==u,ord[1]]))],col=u)
        lines(c(1,length(PC[unit==u,ord[1]])),PC[unit==u,ord[1]]
[c1,length(PC[unit==u,ord[1]]))],col=rgb(1,1,1,.25))}
for(u in unique(unit))
lines(c(1,length(PC[unit==u,ord[1]])),PC[unit==u,ord[1]]
[c1,length(PC[unit==u,ord[1]]))],col=u,type='p',pch=16)
```

Zu den Abbildungen 32 bis 37

```
nb_states=2
k=3
PC=as.data.frame(PC)

##Hidden-Markov-Model
f=list(as.formula(paste(names(PC)[ord[1]],"~1")))
fam=list(gaussian())
if(k>1)
        for(i in 2:k){
        f[[i]]=as.formula(paste(names(PC[,ord])[i],"~1"))
        fam[[i]]=gaussian()}
algo_HMM=depmix(f,nstates=nb_states,data=as.data.frame(PC),ntimes=ULv,family=fam)
HMM=fit(algo_HMM)
HMM=posterior(HMM)$state
plot_fct(HMM,"Hidden-Markov-Model")

##K-Means
KM=kmeans(PC[,ord[1:k]],nb_states)$cluster
plot_fct(KM,"K-means algorithm")

##Agglomerative hierarchical method (unsupervised)
H=cutree(hclust(dist(PC[,ord[1:k]]),method="ward.D"),nb_states)
plot  fct(H,"Hierarchical method")
```

```
##Vorhersage der RUL für den zweiten(gebrauchten) Zustand auf Basis des HMM
cc=HMM==1
if(mean(RUL[cc])>mean(RUL[!cc])) cc=HMM==2
X=as.data.frame(PC[cc,ord[1:3]])
Y=RUL[cc]
time=time[cc]

##MSE (Mittlere Quadratische Fehler) und RMSE (Wurzel der mittleren
Fehlerquadratsumme)
error=function(Yb,Y)
        c(sqrt(mean((Yb-Y)^2)),sqrt(mean((Yb-Y)^2))/mean(Y))

##Kreuzvalidierung (80% Trainingsdaten; 20% Testdaten)
train=NULL
train[1:nrow(X)]=T
train[sample(1:nrow(X),nrow(X)/5)]=F
test=!train
```

Zu den Abbildungen 38 bis 46 (1/2)

```
## Linear regression
model_LM=glm(Y[train]~.,data=X[train,])
LMtr=predict(model_LM,as.data.frame(X[train,]))
LMte=predict(model_LM,as.data.frame(X[test,]))
error(LMtr,Y[train]);error(LMte,Y[test])

## Quadratic regression
k=2
model_LMq=glm(Y[train]^(1/k)~.,data=X[train,])
LMqtr=predict(model_LMq,as.data.frame(X[train,]))^k
LMqte=predict(model_LMq,as.data.frame(X[test,]))^k
error(LMqtr,Y[train]);error(LMqte,Y[test])

## Exponential regression
model_LMe=glm(log(Y[train])~.,data=X[train,])
LMetr=exp(predict(model_LMe,as.data.frame(X[train,])))
LMete=exp(predict(model_LMe,as.data.frame(X[test,])))
error(LMetr,Y[train]);error(LMete,Y[test])

## Support vector machine
algo_SVM=svm(Y[train]~.,data=X[train,])
SVMtr=predict(algo_SVM,X[train,])
SVMte=predict(algo_SVM,X[test,])
error(SVMtr,Y[train]);error(SVMte,Y[test])

## Random forest
n=100
algo_RF=randomForest(Y[train]~.,data=X[train,],ntree=n)
RFtr=predict(algo_RF,X[train,])
RFte=predict(algo_RF,X[test,])
error(RFtr,Y[train]);error(RFte,Y[test])

## K nearest neighbors
n=1
KNNtr=as.numeric(knn(X[train,],X[train,],Y[train],n))
KNNte=as.numeric(knn(X[train,],X[test,],Y[train],n))
error(KNNtr,Y[train]);error(KNNte,Y[test])

## Neural network
h=2
algo_NN=neuralnet(as.formula(paste("Y[train]~",paste(names(X),collapse="+"))),data=X[train,],h
idden=h,
linear.output=T,threshold = .1,stepmax = 1e+07)#,act.fct=function(x)log(1+exp(x*1))/1)
NNtr=round(as.vector(compute(algo_NN,X[train,])$net.result))
NNte=round(as.vector(compute(algo_NN,X[test,])$net.result))
error(NNtr,Y[train]);error(NNte,Y[test])
```

Zu den Abbildungen 38 bis 46 (2/2)

```
## Mittelwert der Algorithmen (multiple learning algorithms)
#MEANtr=apply(cbind(LMtr,LMqtr,LMetr,SVMtr,RFtr,KNNtr,NNtr),1,mean)
#MEANte=apply(cbind(LMte,LMqte,LMete,SVMte,RFte,KNNte,NNte),1,mean)
MEANtr=apply(cbind(SVMtr,RFtr,NNtr),1,mean)
MEANte=apply(cbind(SVMte,RFte,NNte),1,mean)

## Plots der Ergebnisse
par(mfrow=c(4,4))
par(mar=c(4,4,2.5,2))
par(mgp=c(2.5,.75,0))

plot_fct=function(Yb,Y,m){
        plot(Y,Yb,xlim=range(c(Y,Yb)),ylim=range(c(Y,Yb)),col=rgb(0,0,0,.1),xlab="RUL",ylab
="Predicted RUL",main=m,cex.main=1)
        legend("topleft",paste(round(error(Yb,Y),5)),cex=.9,bty='n')
        abline(b=1,a=0,col=2)
        hist(Yb,main="",xlab="Predicted RUL",breaks=10,xlim=range(c(Yb,Y)))}

if(F){
        plot_fct(LMtr,Y[train],"Linear regression")
        plot_fct(LMqtr,Y[train],"Quadratic regression")
        plot_fct(LMetr,Y[train],"Exponential model")
        plot_fct(SVMtr,Y[train],"Support vector machine")
        plot_fct(RFtr,Y[train],"Random forest")
        plot_fct(KNNtr,Y[train],"K nearest neighbors")
        plot_fct(NNtr,Y[train],"Neural network")
        plot_fct(MEANtr,Y[train],"Mean of the algorithms")
        hist(Y[train],xlab="RUL",main="")
        error(mean(Y[train]),Y[train])}else{
        plot_fct(LMte,Y[test],"Linear regression")
        plot_fct(LMqte,Y[test],"Quadratic regression")
        plot_fct(LMete,Y[test],"Exponential model")
        plot_fct(SVMte,Y[test],"Support vector machine")
        plot_fct(RFte,Y[test],"Random forest")
        plot_fct(KNNte,Y[test],"K nearest neighbors")
        plot_fct(NNte,Y[test],"Neural network")
        plot_fct(MEANte,Y[test],"Mean of the algorithms")
        hist(Y[test],xlab="RUL",main="")
        error(mean(Y[test]),Y[test])}
error(MEANtr,Y[train]);error(MEANte,Y[test])

## Testglättung von Inputdaten mit exponentiellem Kernel
#Xs=X;for(k in 1:ncol(X)) X[,k]=es(X[,k],"ANN")$fitted
#par(mfrow=c(2,3)); for(k in 1:ncol(X)){plot(Xs[1:100,k],type='l');lines(X[1:100,k],col=2)}
```