

Herausforderungen und Mehrwerte beim Einsatz von Predictive Maintenance unter Verwendung von Methoden der intelligenten Datenanalyse in der diskreten Fertigung

Dissertation

zur Erlangung des akademischen Grades

Doktoringenieurin

(Dr.-Ing.)

von Dipl.-Ing. (FH) Jessica Fisch

geb. am 23.05.1987 in Berlin

genehmigt durch die Fakultät für Elektrotechnik und Informationstechnik
der Otto-von-Guericke-Universität Magdeburg

Gutachter:

Prof. Dr.-Ing. Christian Diedrich

Prof. Dr.-Ing. Jörg Krüger

Promotionskolloquium am 23.02.2022

*Das Unmögliche zu schaffen, gelingt einem nur,
wenn man es für möglich befindet.*

Zitat aus "Alice im Wunderland"

Kurzfassung

Der Einsatz von voraussagender Instandhaltung (engl. Predictive Maintenance) in der diskreten Fertigung ist, u.a. für die Steigerung des Wettbewerbsvorteils, zielführend. Dadurch kann die Lebensdauer von Komponenten optimal ausgenutzt, die technische Verfügbarkeit gesteigert und Wartungen können geplant durchgeführt werden. Für die Realisierung von voraussagender Instandhaltung sind Vorhersagemodelle notwendig. Diese werden in der Regel mit Hilfe von Daten eines Fertigungssystems und Methoden des maschinellen Lernens trainiert.

Jedoch ist die Umsetzung der voraussagenden Instandhaltung nicht trivial und es sind entsprechend der Literaturrecherche dieser Arbeit bislang keine alles abdeckenden Software-Werkzeuge am Markt verfügbar. Dies lässt sich sicherlich auch darauf zurückführen, dass es laut einer Umfrage fünf Kernherausforderungen bei der Umsetzung von Predictive Maintenance gibt. Die Herausforderung über eine ausreichend große Datenbasis zu verfügen, ist dabei als einzige nicht messbar. Bedingt dadurch sind für die Fertigung bisher keine fundierten Kenntnisse über die Größe der für den Einsatz von voraussagender Instandhaltung notwendigen Datenbasis entwickelt, wissenschaftlich analysiert und in unterschiedlichen Szenarien verglichen worden.

Einleitend wird in dieser Arbeit auf die Herausforderungen in der Fertigung für die Vorhersage von Komponentenausfälle eingegangen. Durch einen Systementwurf wird die Generierung von Vorhersagemodellen aufgezeigt und somit eine Möglichkeit zur Realisierung der vorausschauenden Instandhaltung gegeben. Allerdings wird dabei erneut deutlich, dass ohne ausreichende Datenbasis kein Vorhersagemodell bei einem datengetriebenen Ansatz trainiert werden kann.

Im Hauptteil der vorliegenden Promotionsarbeit wird somit eine Methodik zur Bestimmung der idealisierten Vorhersagegüte entwickelt, um die notwendige Größe einer Datenbasis abschätzen zu können. Anschließend wird die Methode für ein einzelnes Fertigungssystem, nämlich einem Bearbeitungszentrum, unter Betrachtung dreier Szenarien demonstriert. Im ersten Szenario wird davon ausgegangen, dass der Beginn der Datenerfassung sowie das Training und die Nutzung des Vorhersagemodells mit dem Start der Betriebsphase zusammenfällt. Das zweite Szenario hingegen geht von einem späteren Beginn der Datenerfassung und somit auch einem verspäteten Einsatz von voraussagender Instandhaltung aus. Ein weiteres drittes Szenario befasst sich mit der Verbesserung der Vorhersagegüte durch die Nutzung von Daten bzw. Vorhersagemodellen aus baugleichen Systemen. Alle Szenarien basieren dabei auf einem datengetriebenen Ansatz für das Training der Machine-Learning-Modelle.

In einem weiteren Schritt wird die Methodik auf eine Stichprobe bestehend aus 34 Produktionssystemen bzw. 536 Komponenten der Motoren-Komponenten-Produktion angewandt, wodurch sich allgemeingültige Gütwerte für jegliche Art von

Vorhersagemodellen ableiten lassen. Eine Validierung im Anschluss stellt die Güte-
werte direkt auf die Probe. Durch die Werte kann die obere Schranke für das
bestmögliche Vorhersagemodell abgeleitet und somit der maximalmögliche Nutzen
bestimmt werden. Damit kann auch direkt ermittelt werden, ob sich der Einsatz von
Predictive Maintenance bzw. von irgendeinem Vorhersagemodell überhaupt rentie-
ren kann. Außerdem ergibt sich daraus direkt, wie groß die dafür notwendige
Datenbasis sein muss.

Abschließend wird basierend auf den bestimmten Gütewerten für eine Klasse von
Produktionssystemen noch quantitativ der Mehrwert von Predictive Maintenance
in Form von einer gesteigerten technischen Verfügbarkeit bestimmt und mittels ei-
nes Kostenvergleich zwischen reaktiver und voraussagender Instandhaltung
aufgezeigt. Die in dieser Arbeit entwickelte und vorgestellte Methodik ist dement-
sprechend der erste bekannte Ansatz, welcher eine objektive Abschätzung von
voraussagender Instandhaltung in der diskreten Fertigung sowie der für Vorher-
sagemodelle notwendigen Datenbasis ermöglicht.

Abstract

The use of predictive maintenance in discrete manufacturing is useful, among other things, for increasing competitive advantage. This allows the lifetime of components to be optimally utilized, technical availability to be increased, and maintenance to be carried out in a planned manner. Predictive models are necessary for the realization of predictive maintenance. These are usually trained with the help of data from a manufacturing system and machine learning methods.

However, the implementation of predictive maintenance is not trivial and, according to the literature research of this thesis, no software tools covering everything are available on the market so far. This can certainly also be attributed to the fact that, according to a survey, there are five core challenges in the implementation of predictive maintenance. The challenge of having a sufficiently large database is the only one that cannot be quantized. As a result, no well-founded knowledge about the size of the database required for the use of predictive maintenance has yet been developed, scientifically analyzed and compared in different scenarios for manufacturing.

As an introduction, this thesis discusses the challenges in manufacturing for the prediction of component failures. Through a system design the realization of prediction models is described and, thus, a possibility for the realization of predictive maintenance is given. However, it becomes clear again that without sufficient data, no predictive model can be trained in a data-driven approach.

In the main part of this thesis a methodology for the determination of the idealized prediction power is developed in order to estimate the necessary size of the underlying data pool. Subsequently, the methodology is demonstrated for a single manufacturing system, namely a machining center, considering three scenarios. In the first scenario, it is assumed that the start of data collection as well as the training and use of the predictive model coincide with the start of the operational phase. The second scenario, on the other hand, assumes a later start of data collection and, hence, a delayed use of predictive maintenance. A third scenario deals with the improvement of the prediction power by using data or, more precisely, prediction models from identical systems. All scenarios are based on a data-driven approach for training the machine learning models.

In a further step, the methodology is applied to a sample consisting of 34 production systems or 536 components of engine-component production, allowing generally applicable goodness values to be derived for any type of predictive model. A validation afterwards puts these goodness values directly to the test. The values can be used to derive the upper bound for the optimum prediction model and, thus, the maximum possible benefit. This also makes it possible to determine directly, whether the use of predictive maintenance or of any prediction model can be profitable at all. In addition, the size of the necessary data pool can be determined directly.

Finally, based on the determined goodness values for a class of production systems, the added value of predictive maintenance in terms of increased technical availability and cost comparison between reactive and predictive maintenance is determined quantitatively. Accordingly, the methodology developed and presented in this thesis is the first known approach that allows an objective estimation of predictive maintenance in discrete manufacturing as well as of the data pool necessary for predictive models.

Inhaltsverzeichnis

Kurzfassung	III
Abstract	V
Inhaltsverzeichnis	VII
Verzeichnis der Veröffentlichungen des Verfassers	X
Verzeichnis der Abkürzungen	XII
Verzeichnis der Formelzeichen	XIII
Verzeichnis der Abbildungen	XIV
Verzeichnis der Tabellen	XVIII
1 Einleitung	1
1.1 Neue Möglichkeiten zur Steigerung der Wettbewerbsfähigkeit	1
1.2 Zielstellung.....	3
1.3 Struktur der Arbeit	4
2 Grundlagen und Stand der Technik	6
2.1 Produktionssysteme.....	6
2.1.1 Technische Verfügbarkeit	7
2.1.2 Eigenschaften von Ausfällen	8
2.1.3 Lebensdauer eines technischen Systems	9
2.1.4 Einflussgrößen auf die Lebensdauer von Komponenten.....	11
2.1.5 Fertigung in der Motoren-Komponenten-Produktion.....	11
2.2 Instandhaltungsstrategien und -werkzeuge	14
2.2.1 Mehrwerte der verschiedenen Instandhaltungsarten	14
2.2.2 Marktanalyse der Software-Werkzeuge für Predictive Maintenance	18
2.3 Notwendigkeit der Arbeit.....	23
2.4 Fazit	28
3 Systementwurf für die Generierung von Vorhersagemodellen	30
3.1 Datenbasis für Vorhersagemodelle von Ausfällen.....	30
3.1.1 Gewinnung von historischen Daten für das Black-Box- Modell	31
3.1.2 Vorgehensweise einer physikalischen Modellierung für das White-Box-Modell	33

3.1.3	Kennzeichnung von Ausfällen	33
3.2	Vorhersagemodelle durch Klassifikation	36
3.2.1	Vorgehensweise zur Erstellung eines Klassifikationsmodells.....	37
3.2.2	Übertragung eines Klassifikationsmodells	39
3.3	Fazit.....	40
4	Methodik für die Bestimmung der Vorhersagegüte.....	41
4.1	Einführung in die Methodik und Szenarien.....	41
4.1.1	Festlegung eines idealen Klassifikators für die Methodik	42
4.1.2	Formulierung unterschiedlicher Szenarien zur Anwendung der Methodik.....	43
4.2	Herleitung der Methodik zur Bestimmung der Vorhersagegüte.....	44
4.2.1	Auswahl der notwendigen Gütemaße für eine Leistungsbewertung der Datenmenge	46
4.2.2	Bestimmung von relevanten Gütemaßen.....	49
4.2.3	Vorgehensweise zur Verwendung der Gütemaße	51
4.3	Modellierung der Gütemaße.....	52
4.3.1	Vorgehensweise zur Bestimmung einer empirischen Verteilungsfunktion.....	53
4.3.2	Vorgehensweise zur Bestimmung einer Näherungsfunktion	53
4.4	Fazit.....	54
5	Bestimmung der Vorhersagegüte für ein Bearbeitungszentrum	55
5.1	Analyse der Ausgangslage	55
5.2	Durchführung von Szenario 1 für ein Bearbeitungszentrum	58
5.2.1	Bestimmung der erlernbaren Ausfälle	59
5.2.2	Bestimmung der vermeidbaren Ausfälle.....	62
5.2.3	Bestimmung der verbleibenden Ausfälle	64
5.3	Durchführung von Szenario 2 für ein Bearbeitungszentrum	66
5.4	Anwendung der Methode auf ein baugleiches Bearbeitungszentrum	72
5.5	Fazit.....	75
6	Generalisierung und Validierung der Vorhersagegüte.....	78
6.1	Motivation der Generalisierung und Validierung	78
6.2	Stichprobenumfang für eine Klasse von Produktionssystemen.....	79
6.3	Generalisierung der Vorhersagegüte für eine Klasse von Produktionssystemen	80

6.3.1	Vorgehensweise der Generalisierung.....	81
6.3.2	Generalisierung der Vorhersagegüte im Szenario 1	81
6.4	Validierung der Vorhersagegüte für eine Klasse von Produktionssystemen.....	87
6.5	Fazit	91
7	Ableitung der Mehrwerte	92
7.1	Vorgehensweise zur Berechnung der gesteigerten technischen Verfügbarkeit	92
7.2	Technische Verfügbarkeit – Stichprobe 1	94
7.3	Exemplarische Berechnung des Gewinns – Stichprobe 1	96
7.4	Fazit	101
8	Zusammenfassung und Ausblick.....	102
8.1	Zusammenfassung.....	102
8.2	Wissenschaftlicher Beitrag dieser Arbeit	104
8.3	Ausblick.....	106
A	Bestimmung der Vorhersagegüte.....	109
A.1	Mittlere Lebensdauern je Komponente zugeordnet zum jeweiligen System.....	109
A.2	Generalisierung der Vorhersagegüte für Stichprobe 1 im Szenario 2	112
A.3	Validierung der Stichprobe 1 durch BAZ 1 im Szenario 2	115
A.4	Validierung der Stichprobe 2	118
	Literaturangaben	125
	Verzeichnis der betreuten studentischen Arbeiten.....	131
	Ehrenerklärung	132

Verzeichnis der Veröffentlichungen des Verfassers

- Arora, S.-J., Ebbecke, C., Rabe, M. u. Fisch, J.: *Methodology for the assessment of potentials, selection, and design of Predictive Maintenance solutions*. Procedia CIRP 104 (2021), S. 708–713
- Stephan, P., Fisch, J., Can, A., Heimann, O., Thiele, G. u. Krüger, J.: (im Druck) *Data driven automatic parameter inference for robotic assembly programs*. 54th CIRP Conference on Manufacturing Systems (CMS) 2021
- Can, A., Kolesnik, M., Moltchanov, A., Fisch, J. u. Krüger, J.: (im Druck) *Smart Remote Maintenance. Datengetriebene Fernwartung von Produktionssystemen*. atp magazin 8/2021
- Stephan, P., Fisch, J., Can, A., Heimann, O., Thiele, G. u. Krüger, J.: *Generic Architecture for the Automatic Parametrization of Production Machine Assembly Programs*. IOP Conference Series: Materials Science and Engineering 1140 (2021) 1, S. 12025
- Can, A., Fisch, J., Thiele, G. u. Krueger, J.: *Automated continuous learn and improvement process of energy efficiency in manufacturing*. 46th Annual Conference of the IEEE Industrial Electronics Society (IECON) 2020
- Can, A., Thiele, G., Krueger, J., Fisch, J. u. Klemm, C.: *A practical approach to reduce energy consumption in a serial production environment by shutting down subsystems of a machine tool*. 16th Global Conference on Sustainable Manufacturing (GCSM) 2018
- Fisch, J., Can, A., Diedrich, C. u. Krueger, J.: *Transformation der Wissenstreppe zum digitalen Fahrstuhl durch die Verwendung von Big Data-Analysen*. Baden-Baden, 03. und 04. Juni. 19. Leitkongress der Mess- und Automatisierungstechnik. AUTOMATION 2018 (VDI-Berichte 2330)
- Auris, F., Fisch, J., Brandl, M., Suess, S., Soubar, A. u. Diedrich, C.: *Enhancing Data-Driven Models with Knowledge from Engineering Models in Manufacturing*. 14th IEEE International Conference on Automation Science and Engineering (CASE) 2018
- Fisch, J. u. Diedrich, C.: *Methodische Untersuchung des Komplexitätsanstiegs von Produktionssystemen*. at - Automatisierungstechnik 2018, 66(6), S. 449–455
- Fisch, J., Diedrich, C., Rossdeutscher, M. u. Remlein, A.: *Einsatz von Korrelationsanalysen zur Zuordnung von Produktionsdaten zu Funktionsgruppen in einer Werkzeugmaschine*. Baden-Baden, 27. und 28. Juni. 18. Leitkongress der Mess- und Automatisierungstechnik. AUTOMATION 2017 (VDI-Berichte 2293)
- Fisch, J., Haeussler, B., Can, A. u. Diedrich, C.: *Aufbau einer Schnittstelle für die dienstorientierte Automatisierung mittels Analytik-Entscheidungen. Parametrierung von Steuerungen im Cloud-Zeitalter*. atp edition 2018, 4-5, S. 26–29

Fisch, J., Rossdeutscher, M. u. Diedrich, C.: *Anwendung datenbasierter Methoden auf Werkzeugmaschinen Daten zur Abweichungserkennung vom Normalbetrieb*. Le-
onberg bei Stuttgart, 17. und 18. Mai 2017. 28. VDI-Fachtagung Technische
Zuverlässigkeit

Vorträge, Wettbewerbe und Workshops

Fisch, J. u. Harrendorf, M.: *Predictive Maintenance at Mercedes-Benz Cars using ad-
vanced technologies*. Vortrag. Smart Remote Service. Berlin 2019

Fisch, J. u. Ostapenko, I.: *Predictive Maintenance with MATLAB: Pattern Matching for
Time-Series Data*. Vortrag. AUTOMOTIVE CONFERENCE. Stuttgart 2018

Fisch, J. u. Ostapenko, I.: *Predictive Maintenance mit MATLAB: Time Series Production
Data Analysis*. Vortrag. MATLAB EXPO. München 2018

Fisch, J.: *Implementierung von Big-Data-Analysen in die Produktion - Herausforderun-
gen, Mehrwerte und Möglichkeiten*. Workshop. Fabrik des Jahres 2017. Ulm 2018

Fisch, J.: *Fabrik des Jahres 2017. Standortabsicherung durch Digitalisierung*. Projekt-
leiter Mercedes-Benz Werk Berlin. Wettbewerb. [https://www.fabrik-des-
jahres.de](https://www.fabrik-des-jahres.de)

Fisch, J.: *Entwicklung einer MATLAB-Analysesoftware für Big Data in der Motoren-
Komponenten-Produktion*. Vortrag. MATLAB EXPO. München 2016

Fisch, J.: *Wie schafft es die Instandhaltung die Menschen zu faszinieren?* Podiumsdis-
kussion. 37. VDI/VDEh-Forum Instandhaltung, Vernetzte Instandhaltung im
Unternehmen. Berlin 2016

Fisch, J.: *Methode zur Bestimmung des Komplexitätsgrades von Produktionsanlagen*.
Vortrag. 37. VDI/VDEh-Forum Instandhaltung, Vernetzte Instandhaltung im
Unternehmen. Berlin 2016

Verzeichnis der Abkürzungen

A		MTTF	Mean Time To Failure
AI	Artificial Intelligence	MTTFF	Mean Time To First Failure
B		MTTR	Mean Time To Repair
B	Betriebsdauer	N	
BAZ	Bearbeitungszentrum	N	Negatives
BI	Business Intelligence	NC	Numeric Control
C		NPV	Negative-Predictive-Value
CM	Condition Monitoring	O	
CNC	Computerized Numerical Control	OEM	Original Equipment Manufacturer
F		P	
FN	False-Negativ	P	Positives
FNR	False-Negativ-Rate	PaaS	Plattform-as-a-Service
FP	False-Positiv	P-F-I	Intervall zwischen potenziellem und funktionalem Ausfall
FPR	False-Positiv-Rate	PM	Predictive Maintenance
FS	Fertigungssystem	PPV	Positive-Predictive-Value
I		PR	Positive-Rate
I4.0	Industrie 4.0	S	
IaaS	Infrastruktur-as-a-Service	SaaS	Software-as-a-Service
IIoT	Industrial Internet of Things	SPS	Speicherprogrammierbare Steuerung
IoT	Internet of Things	SVM	Support Vector Maschinen
K		T	
K	Komponente	TCO	Total Cost of Ownership
KDD	Knowledge Discovery in Database	TN	True-Negativ
kNN	k-Nearest-Neighbor-Algorithmen	TNR	True-Negativ-Rate
L		TP	True-Positiv
L	Lebensdauer	TPR	True-Positiv-Rate
LP	Learn-Positive	U	
LPR	Learn-Positive-Rate	umati	universal machine tool interface
M		Z	
ML	Machine Learning	ZE	Zeiteinheit
MQTT	Message Queuing Telemetry Transport		
MTBF	Mean Time Between Failure		

Verzeichnis der Formelzeichen

A_T	technische Ausfallrate
A_W	wartungsbedingte Ausfallrate
a	Ausfall
E	Erlös
F	Fixkosten
G	Gewinn
k	variable Kosten
K	Kosten
N_G	Gesamtnutzungsgrad
N_T	technischer Nutzungsgrad
o_K	Offset für die Anzahl der nicht-erfassten Ausfälle
p	Gewinn pro Stück
s_K	Anzahl notwendiger Ausfälle einer Komponente zum Erlernen des Ausfalls
t	Zeitpunkt
T_B	geplante Belegungszeit
T_{FS}	Folgeausfallzeit
T_N	Nutzungszeit
T_O	organisatorische Ausfallzeit
T_T	technische Ausfallzeit
T_W	Wartungszeit
T_x	Taktzeit pro Stück
V_T	technische Verfügbarkeit
$V_{T,PM}$	gesteigerte technische Verfügbarkeit durch PM
x	Stückzahl
$\lambda(t)$	Ausfallrate in Abhängigkeit der Zeit

Verzeichnis der Abbildungen

Bild 1.1: Aufbau von Wissen als Basis der Wettbewerbsfähigkeit.....	1
Bild 1.2: Struktur der Arbeit je Kapitel	5
Bild 2.1: Anteile der Kosten über die verschiedenen Nutzungsphasen	6
Bild 2.2: Schematischer Verlauf eines Driftausfalls bei Abnutzung einer Komponente mit P-F-Intervall.....	8
Bild 2.3: Typische Badewannenkurve der Ausfallrate für Komponenten.....	9
Bild 2.4: Grafische Beschreibung der Begriffe MTTF, MTTFE, MTBF und MTTR.....	10
Bild 2.5: Ursache-Wirkungs-Diagramm der Einflussgrößen auf die Lebensdauer von Komponenten	11
Bild 2.6: Funktionsumfang von Werkzeugmaschinen	12
Bild 2.7: Interne Struktur der Steuerung von Werkzeugmaschinen	13
Bild 2.8: Mehrwerte durch eine voraussagende Instandhaltung.....	16
Bild 2.9: Betriebszustände je Instandhaltungsstrategie	17
Bild 2.10: Umfrage zu Predictive Maintenance.....	18
Bild 2.11: Werkzeuge für Predictive Maintenance am Markt	18
Bild 2.12: Applikationen der DMG Mori AG	19
Bild 2.13: Erfüllungsgrad der Software-Werkzeuge je Instandhaltungsstrategie ...	22
Bild 2.14: Herausforderungen bei der Umsetzung von voraussagender Instandhaltung.....	23
Bild 2.15: Schichtenmodell für Cloud-Services.....	24
Bild 2.16: Zwei Vorgehensweisen für Datenanalysen.....	25
Bild 2.17: Überblick der analytischen Themenfelder in Form eines Venn- Diagrammes	26
Bild 3.1: Ansätze zur Generierung eines Vorhersagemodells.....	31
Bild 3.2: Übersicht über verschiedene Clustering-Verfahren	36
Bild 4.1: Wahl der Delta-Verteilung für die Eintrittswahrscheinlichkeit von Ausfällen	45
Bild 4.2: Schematische Darstellung der berechneten Mengen für ein Beispielsystem bei idealisierter Klassifikation.....	47
Bild 5.1: Struktogramm für die Berechnung der Menge der wahren Ausfälle P entsprechend Szenario 1.....	56
Bild 5.2: Normierte Datenpunkte, empirische und genäherte Verteilung der Ausfallrate PR des Systems BAZ 1 für wahre Ausfälle.....	58
Bild 5.3: Struktogramm für die Berechnung der Menge der erlernbaren Ausfälle LP entsprechend Szenario 1	60

Bild 5.4: Normierte Datenpunkte, empirische und genäherte Verteilung der Lernrate LPR des Systems BAZ 1 für erlernbare Ausfälle entsprechend Szenario 1.....	61
Bild 5.5: Struktogramm für die Berechnung der Menge der vermeidbaren Ausfälle TP entsprechend Szenario 1.....	62
Bild 5.6: Normierte Datenpunkte, empirische und genäherte Verteilung der Trefferrate TPR des Systems BAZ 1 für vermeidbare Ausfälle entsprechend Szenario 1.....	63
Bild 5.7: Struktogramm für die Berechnung der Menge der verbleibenden Ausfälle FN entsprechend Szenario 1	65
Bild 5.8: Normierte Datenpunkte, empirische und genäherte Verteilung der Nicht-Alarmrate FNR des Systems BAZ 1 für verbleibende Ausfälle entsprechend Szenario 1.....	66
Bild 5.9: Struktogramm für die Berechnung eines Offsets bei späterem Beginn der Datenerfassung.....	67
Bild 5.10: Generelles Struktogramm für die Berechnung der Mengen P, LP, TP, FN	68
Bild 5.11: Normierte Datenpunkte, empirische und genäherte Verteilung der Lernrate LPR des Systems BAZ 1 für erlernbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb	69
Bild 5.12: Normierte Datenpunkte, empirische und genäherte Verteilung der Trefferrate TPR des Systems BAZ 1 für vermeidbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb	70
Bild 5.13: Normierte Datenpunkte, empirische und genäherte Verteilung der Nicht-Alarmrate FNR des Systems BAZ 1 für verbleibende Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb	71
Bild 5.14: Normierte Datenpunkte, empirische und genäherte Verteilung der Trefferrate TPR des Systems BAZ 1 für vermeidbare Ausfälle entsprechend Szenario 3 mit zeitlichem Versatz der Systeme	74
Bild 6.1: Erhobene mittlere Lebensdauern der Komponenten je Produktionssystem	80
Bild 6.2: Normierte Datenpunkte aller einzelnen Systeme und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Ausfallrate PR für wahre Ausfälle	82
Bild 6.3: Normierte Datenpunkte aller einzelnen Systeme und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Lernrate LPR für erlernbare Ausfälle entsprechend Szenario 1	84
Bild 6.4: Normierte Datenpunkte aller einzelnen Systeme und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Trefferrate TPR für vermeidbare Ausfälle entsprechend Szenario 1	85

Bild 6.5: Normierte Datenpunkte aller einzelnen Systeme und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Nicht-Alarmrate FNR für verbleibende Ausfälle entsprechend Szenario 1.....	86
Bild 6.6: Normierte Datenpunkte des Systems BAZ 1 und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Ausfallrate PR für wahre Ausfälle.....	87
Bild 6.7: Normierte Datenpunkte des Systems BAZ 1 und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Lernrate LPR für erlernbare Ausfälle entsprechend Szenario 1	88
Bild 6.8: Normierte Datenpunkte des Systems BAZ 1 und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Trefferrate TPR für vermeidbare Ausfälle entsprechend Szenario 1.....	89
Bild 6.9: Normierte Datenpunkte des Systems BAZ 1 und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Nicht-Alarmrate FNR für verbleibende Ausfälle entsprechend Szenario 1..	90
Bild 7.1: Berechnete technische Verfügbarkeit im Falle von voraussagender Instandhaltung für die Stichprobe 1 entsprechend Szenario 1	95
Bild 7.2: Berechnete Erlös-, Kosten- und Gewinn-Verteilungen ohne Fixkosten für die Stichprobe 1 entsprechend Szenario 1	99
Bild 7.3: Berechnete Erlös-, Kosten- und Gewinn-Verteilungen mit Fixkosten für die Stichprobe 1 entsprechend Szenario 1	100
Bild A.1: Normierte Datenpunkte aller einzelnen Systeme und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Lernrate LPR für erlernbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb	112
Bild A.2: Normierte Datenpunkte aller einzelnen Systeme und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Trefferrate TPR für vermeidbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb....	113
Bild A.3: Normierte Datenpunkte aller einzelnen Systeme und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Nicht-Alarmrate FNR für verbleibende Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb....	114
Bild A.4: Normierte Datenpunkte des Systems BAZ 1 und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Lernrate LPR für erlernbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb	115
Bild A.5: Normierte Datenpunkte des Systems BAZ 1 und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Trefferrate TPR für vermeidbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb	116

Bild A.6: Normierte Datenpunkte des Systems BAZ 1 und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Nicht-Alarmrate FNR für verbleibende Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb.....	117
Bild A.7: Normierte Datenpunkte der Testsysteme und genäherte Verteilung der Stichprobe 2 mit statistischer Unsicherheit der Ausfallrate PR für wahre Ausfälle	118
Bild A.8: Normierte Datenpunkte der Testsysteme und genäherte Verteilung der Stichprobe 2 mit statistischer Unsicherheit der Lernrate LPR für erlernbare Ausfälle entsprechend Szenario 1	119
Bild A.9: Normierte Datenpunkte der Testsysteme und genäherte Verteilung der Stichprobe 2 mit statistischer Unsicherheit der Trefferrate TPR für vermeidbare Ausfälle entsprechend Szenario 1	120
Bild A.10: Normierte Datenpunkte der Testsysteme und genäherte Verteilung der Stichprobe 2 mit statistischer Unsicherheit der Nicht-Alarmrate FNR für verbleibende Ausfälle entsprechend Szenario 1.	121
Bild A.11: Normierte Datenpunkte der Testsysteme und genäherte Verteilung der Stichprobe 2 mit statistischer Unsicherheit der Lernrate LPR für erlernbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb	122
Bild A.12: Normierte Datenpunkte der Testsysteme und genäherte Verteilung der Stichprobe 2 mit statistischer Unsicherheit der Trefferrate TPR für vermeidbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb	123
Bild A.13: Normierte Datenpunkte der Testsysteme und genäherte Verteilung der Stichprobe 2 mit statistischer Unsicherheit der Nicht-Alarmrate FNR für verbleibende Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb.....	124

Verzeichnis der Tabellen

Tabelle 3.1: Übersicht über die verschiedenen Datenquellen einer Werkzeugmaschine	32
Tabelle 3.2: Aufbau einer Konfusionsmatrix einer binären Klassifikation.....	38
Tabelle 3.3: Auflistung der Gütemaße für Klassifikationsergebnisse und deren Definition sowie die dafür bekannten Synonyme	38
Tabelle 4.1: Konfusionsmatrix bei idealisierter Klassifikation.....	47
Tabelle 5.1: Informationen über die Komponenten von System BAZ 1.....	55

1 Einleitung

Dieses Kapitel führt zu Beginn der Arbeit in das Themenfeld Digitalisierung und dessen Möglichkeiten ein. Anschließend folgt die Zielsetzung der vorliegenden Arbeit und es wird auf deren Struktur samt den entsprechenden Inhalten eingegangen.

1.1 Neue Möglichkeiten zur Steigerung der Wettbewerbsfähigkeit

Das Ziel eines jeden Unternehmens ist es, im Wettbewerb die Nummer eins zu sein. Diese Wettbewerbsfähigkeit lässt sich in der Produktion anhand der Faktoren Kosten, Qualität und Zeit zur Herstellung eines Produktes messen [1]. Damit ein Unternehmen einen Wettbewerbsvorteil erlangen kann, müssen dort entsprechende Kompetenzen existieren und auch entsprechend eingesetzt werden. Die Stufen zur Erreichung der Wettbewerbsfähigkeit hat bereits North mit der Wissenstreppe (dargestellt in Bild 1.1) beschrieben. [2]

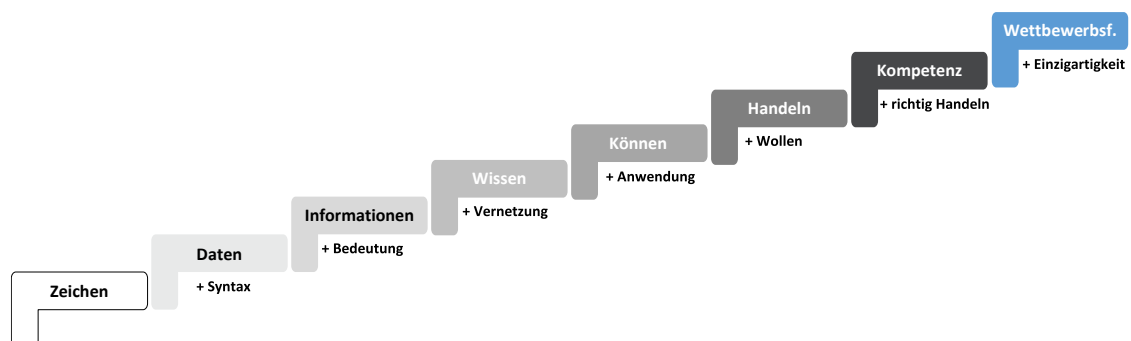


Bild 1.1: Aufbau von Wissen als Basis der Wettbewerbsfähigkeit [2, 3]

Insbesondere durch die digitale Bereitstellung von Informationen im Rahmen einer übergeordneten Digitalisierungsstrategie kann Wissen aufgebaut, ein zielgerichtetes Handeln ermöglicht und ein Wettbewerbsvorteil erzielt werden. Neue Möglichkeiten der Digitalisierung werden heutzutage im Projekt Industrie 4.0 (I4.0), auch vierte Industrie-Revolutionphase genannt, zusammengefasst [4]. Die Basis für die Realisierung der Aktivitäten zu I4.0 bieten die gestiegene Rechenleistung [5] und auch die verbesserten Speichermöglichkeiten von Daten. Ganz nach dem Mooreschen Gesetz verdoppelt sich die Anzahl an Transistoren pro Fläche und damit auch die Rechenleistung statistisch gemittelt alle zwei Jahre, wodurch insbesondere ab der Jahrhundertwende neue Möglichkeiten entstanden sind. [6, 7]

Mit der Verbreitung des Internets und davon abgeleiteter Netzwerk-Technologien ist es außerdem möglich geworden, eine nahezu unüberschaubare Menge an unterschiedlichsten Dingen und alltäglichen Geräten miteinander zu vernetzen. Dies ist besser bekannt als Internet der Dinge oder in Englisch: Internet of Things (IoT). Wenn innerhalb der industriellen Produktion Dinge wie Sensoren, Aktoren, Produkte oder Benutzerschnittstellen miteinander vernetzt werden, wird davon abgeleitet vom Industrial Internet of Things (IIoT) gesprochen. [8 bis 11]

Durch die Möglichkeit der Vernetzung entstehen neue Formen von Daten, welche sich meist durch die Eigenschaften Volume, Velocity und Variety von den bisher bekannten Daten abgrenzen lassen. Volume steht dabei für die immense Menge an Daten, Velocity für die hohe Geschwindigkeit, mit der die Daten anfallen, und Variety für deren große Vielfalt. Diese neue Ära an Daten, gekennzeichnet durch die drei V's, wird häufig mittels des Begriffs – Big Data – beschrieben. Das heißt mit dem Ausdruck Big Data bezeichnet man Daten, welche strukturiert oder unstrukturiert sind, mit einer hohen Abtastrate entstehen sowie in Echtzeit verarbeitet werden wollen und eine große Menge an Speicher belegen, da diese für einen langen Zeitraum gespeichert werden sollen. [12] Es ist also möglich geworden, Daten aus unterschiedlichen Quellen, abgebildet in Variety, zu erhalten und diese in datengetriebenen Analysen zu nutzen und miteinander zu kombinieren. So können zum Beispiel auch die Auswirkungen der Umgebungsbedingungen wie Temperatur und Luftfeuchtigkeit auf einen Fertigungsprozess berücksichtigt werden und in die Optimierung des Fertigungsprozesses als weitere Prozessparameter einfließen oder als Einflussfaktoren für die Vorhersage der Lebensdauer von Komponenten genutzt werden.

Damit künftig Big-Data-Ansätze eingesetzt werden können, mussten neben dem herkömmlichen Data-Warehouse-Modell¹ neue Lösungen entwickelt werden wie die Data-Lake-Technologie [14]. Denn die bislang praktizierte strukturierte Abspeicherung von Daten hat zum Nachteil, dass die Daten bereits vorverarbeitet wurden und somit im Nachgang in der Regel nur noch zuvor festgelegte Fragestellungen analysiert werden können. Dies widerspricht aber dem Big-Data-Gedanken, der vorsieht, dass eine Vielzahl von Daten für einen Wissensgewinn bei Bedarf nutzbringend miteinander verknüpft werden können und durch die erst spät-erfolgende Datenverarbeitung auch jederzeit neue Fragestellungen angegangen werden können. Indem Data Lake's die Daten in der ursprünglichen Form abspeichern, bieten sie folglich die optimale Grundlage für verschiedene, neuere Analytik-Anwendungen [13]. Dies können herkömmliche Business-Intelligence- (BI), aber auch komplexere Advanced-Analytics-Anwendungen sein. Herkömmliche BI-Analysen befassen sich mit der Auswertung und Darstellung von historischen Daten, wohingegen bei Advanced-Analytics-Methoden maschinelles Lernen oder auch andere Arten von künstlicher Intelligenz eingesetzt werden, auf welche in Kapitel 3 näher eingegangen wird. Dadurch ist es möglich, komplexere Analysen durchzuführen und genauere Vorhersagen zu treffen. [13] Um eine solche Vorhersage treffen zu können, ist es dabei notwendig, basierend auf historischen Daten ein Vorhersagemodell zu erlernen und dieses auf aktuellen Daten evaluieren zu können [13, 15]. Diese verbesserten Vorhersagen bieten dann die Grundlage für effizienteres und kompetenteres Handeln und führen damit zu einem Wettbewerbsvorteil, dargestellt in Bild 1.1.

¹ Ein Data Warehouse speichert Daten in einem strukturierten Format ab, sodass diese bereits vorverarbeitet sind [13].

Zusammenfassend ist erkennbar, es ist keine Frage, wieviel mehr Unternehmen durch Digitalisierung verdienen können, sondern wie diese ohne den Einsatz von digitalen Methoden am Markt langfristig weiter bestehen können. Für die optimale Produktfertigung und das Bestehen im Wettbewerb ist es folglich notwendig, weiter zu digitalisieren. [16]

1.2 Zielstellung

Durch die Digitalisierung können neue Technologien auch in produzierenden Unternehmen eingesetzt werden. Im Bereich der Instandhaltung von technischen Anlagen ist dies die voraussagende Wartung (Predictive Maintenance) mittels datengetriebener Methoden. Für deren Einsatz und zum Training von Vorhersagemodellen werden historische Daten benötigt [17].

Die Vorhersagemodelle beruhen dabei auf den bereits aufgetretenen und daraus erlernbaren Ausfällen. Hierfür werden die Ausfalldaten der einzelnen Komponenten eines technischen Systems erfasst. Die erfassten Daten bilden dabei einen Betrachtungszeitraum ab. Dieser Zeitraum hat einen maßgeblichen Einfluss auf die Güte der Vorhersagemodelle. Ein Ausfall, der zum Beispiel im Betrachtungszeitraum nicht erfasst wurde und damit nicht aufgetreten ist, kann nicht im Vorhersagemodell berücksichtigt bzw. trainiert werden und folglich auch nicht vorhergesagt werden. Dementsprechend ist es für die Erfassung eines vollständigen Zustandsraums eines Produktionssystems notwendig, über einen längeren Zeitraum die Daten zu erfassen und zu analysieren. Die notwendige Erfassungsdauer bis zum Auftreten nahezu aller Ausfälle ist systemspezifisch und kann gegebenenfalls sehr lang sein.

Mit Methoden des maschinellen Lernens versuchen Experten wie Data Scientists aus den Daten von Produktionssystemen eine korrekte Klassifikation zu erzielen. Oftmals wird erst dann deutlich, dass die Datenbasis nicht ausreichend ist. Demnach wird dann die Menge der historischen Daten vergrößert und erneut analysiert. Bislang ist meist unklar, welche Datenmenge bzw. damit eng verknüpft welcher Betrachtungszeitraum notwendig ist, wodurch zeitaufwendige Tests ohne Vergleichbarkeit der möglichen Vorhersagegüte durchgeführt werden müssen.

In der bisherigen Literatur finden sich keine Angaben über den notwendigen Zeitrahmen der Datenerfassung und einer sich daraus ergebenden Vorhersagegüte. Deswegen wird bislang lediglich eine subjektive Expertenschätzung der nötigen Datenmenge getroffen. Resultierende Gütewerte und Mehrwerte werden beim Einsatz von datengetriebenen Methoden für die voraussagende Instandhaltung meist gar nicht eingesetzt oder auch nur subjektiv durch Experten abgeschätzt.

Die Zielsetzung der vorliegenden Arbeit ist es, eine datengetriebene Methodik für die Bestimmung der theoretischen Güte von Vorhersagemodellen für Ausfälle von Komponenten eines Produktionssystems auszuarbeiten, zu untersuchen und festzulegen. Mittels der Methodik soll es möglich sein, verschiedene Szenarien beim Einsatz von Predictive Maintenance unter Verwendung von Machine Learning zu

bewerten und zu vergleichen. Außerdem soll die notwendige Größe der Datenbasis bzw. die notwendige Länge des Betrachtungszeitraums mithilfe der Methodik abgeschätzt werden können.

Längerfristig könnten damit subjektive Methoden zur Bestimmung der idealisierten Vorhersagegüte und des Mehrwerts von voraussagender Instandhaltung durch wissenschaftlich fundierte und datengetriebene Methoden vollständig abgelöst werden.

1.3 Struktur der Arbeit

Die Arbeit ist entsprechend Bild 1.2 wie folgt aufgebaut.

Nach der bereits erfolgten Einleitung der Arbeit mit dem Fokus auf der Steigerung der Wettbewerbsfähigkeit durch voraussagende Instandhaltung (engl.: Predictive Maintenance, kurz: PM), werden im zweiten Kapitel die Grundlagen und der Stand der Technik zu Produktionssystemen und zur Instandhaltung in der Fertigung erläutert. Abschließend wird aus den ersten zwei Kapiteln die Notwendigkeit der Arbeit abgeleitet.

Bedingt durch die Herausforderungen von Predictive Maintenance wird im dritten Kapitel ein Systementwurf für die Generierung von Vorhersagemodellen ausgearbeitet. Dieser betrachtet zum einen die Gewinnung von Daten und zum anderen die Datenanalyse zur Generierung von Vorhersagemodellen.

Im vierten Kapitel wird dann die Methodik für die Bestimmung der Vorhersagegüte erarbeitet. Basierend auf drei Szenarien, einem idealen Klassifikator und der Lebensdauer von Komponenten werden entsprechende Berechnungsvorschriften definiert.

Das fünfte Kapitel beinhaltet die Ergebnisse bei Anwendung der Methodik auf ein einzelnes System – genauer ein Bearbeitungszentrum. Es werden drei verschiedene Szenarien durchgeführt, die den Einsatz von Datenanalyse untersuchen. Mit Hilfe von Struktogrammen wird die Logik der Methodik verdeutlicht.

Ein sechstes Kapitel betrachtet die Ergebnisse bei Anwendung der Methodik auf insgesamt 34 Produktionssystemen, wodurch sowohl eine Generalisierung der Vorhersagegüte als auch eine Validierung der Methode ermöglicht wird.

Weiterführend werden im siebten Kapitel die Mehrwerte, die sich durch den Einsatz von Vorhersagemodellen in Form einer Steigerung der technischen Verfügbarkeit ergeben, abgeleitet. Außerdem wird zur Verdeutlichung der Mehrwerte ein fiktives Produktionssystem gewählt und der zusätzlich resultierende Gewinn berechnet.

Eine Zusammenfassung und ein Ausblick schließen die Arbeit im achten Kapitel ab.

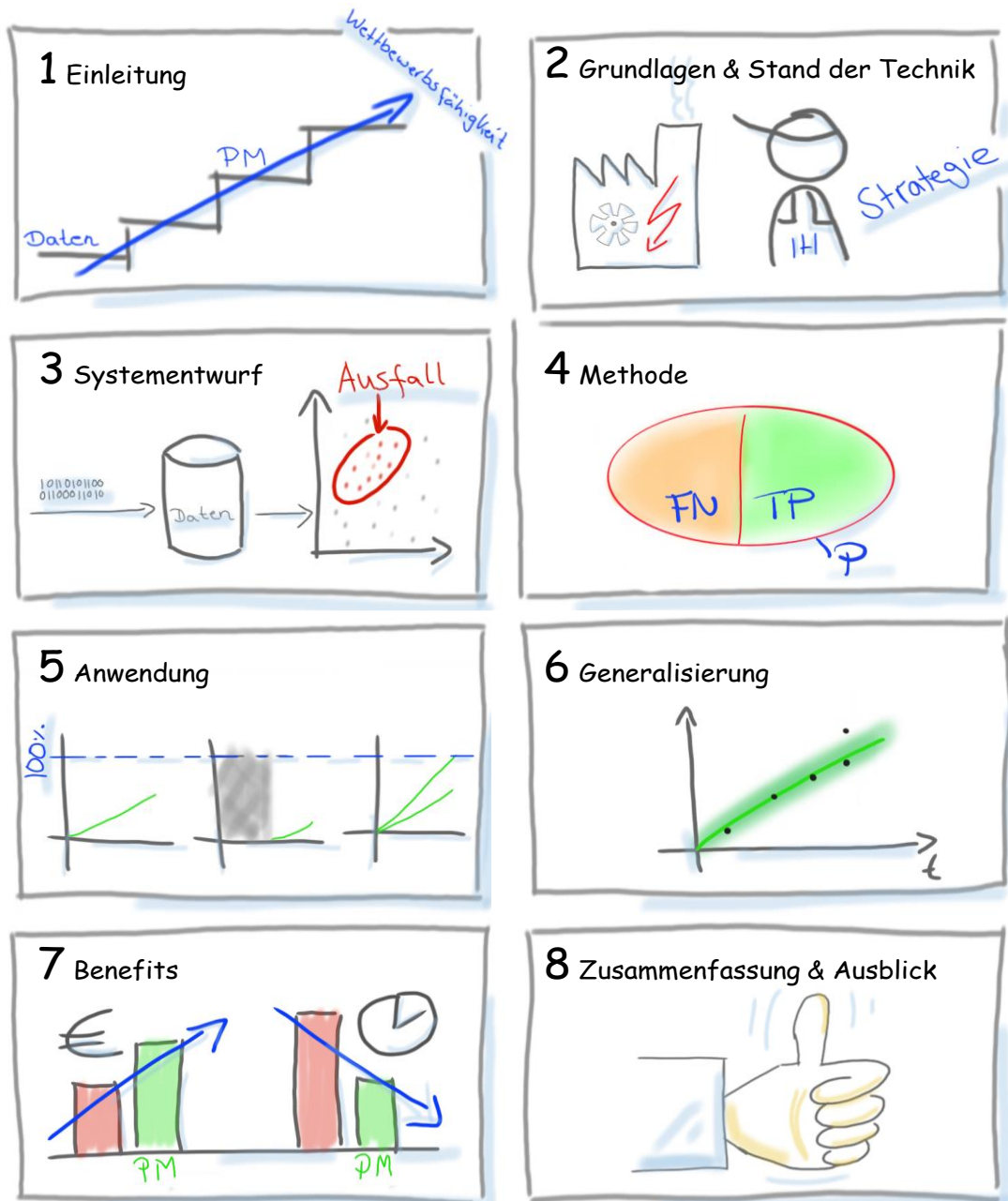


Bild 1.2: Struktur der Arbeit je Kapitel: Skizziert ist der Inhalt jedes Kapitels in einem einzelnen Fenster. Die gewählte Darstellungsform stammt aus der Vorgehensweise zur Skizzierung von Inhalten in Präsentationen (häufig wird diese für Flipcharts angewendet).

2 Grundlagen und Stand der Technik

Das folgende Kapitel gibt Informationen über den Stand der Technik in der Instandhaltung der Motoren-Komponenten-Produktion. Dazu wird als Erstes auf Produktionssysteme mit Schwerpunkt der Fertigung eingegangen und anschließend die Strategien der Instandhaltung mit dem Fokus auf voraussagende Instandhaltung analysiert. Abgeschlossen wird das Kapitel mit einer Begründung der Notwendigkeit der Arbeit.

2.1 Produktionssysteme

Ein Produktionssystem ist ein technisches System, welches nach VDI 3633 als eine „abgegrenzte Anordnung von Elementen, die miteinander in Beziehung stehen“ beschrieben wird [18]. Ein solches System lässt sich in Teilsysteme und wiederum in Komponenten unterteilen, wobei jede Komponente auch ein oder mehrere Funktionselemente besitzen kann [19]. Die Gesamtheit aller Schritte in einem System, welche benötigt werden, um ein Produkt teilweise oder komplett zu fertigen, wird dabei als Prozess bezeichnet [18].

In der vorliegenden Arbeit wird ein technisches System als ein Produktionssystem zur Fertigung, Montage oder Handhabung von Produkten in der Automobilindustrie verstanden. Eine Komponente eines Systems wird dabei durch Sensoren oder Aktoren wie Industrieroboter, Pumpen, Antriebe, Werkzeuge oder Messsysteme aus der Feldebene einer Automatisierungspyramide abgebildet.

Ein technisches System hat einen Lebenszyklus, welcher mit der Planung beginnt und über die Phasen Entwicklung, Herstellung, Inbetriebnahme, Betrieb, Optimierung verläuft und schließlich bei der Entsorgung endet. In dieser Arbeit wird die längste Nutzungsphase, die Betriebsphase, betrachtet. Die generelle Nutzungszeit T_N von Produktionssystemen in der Automobilbranche liegt bei bis zu 15 Jahren. [20]

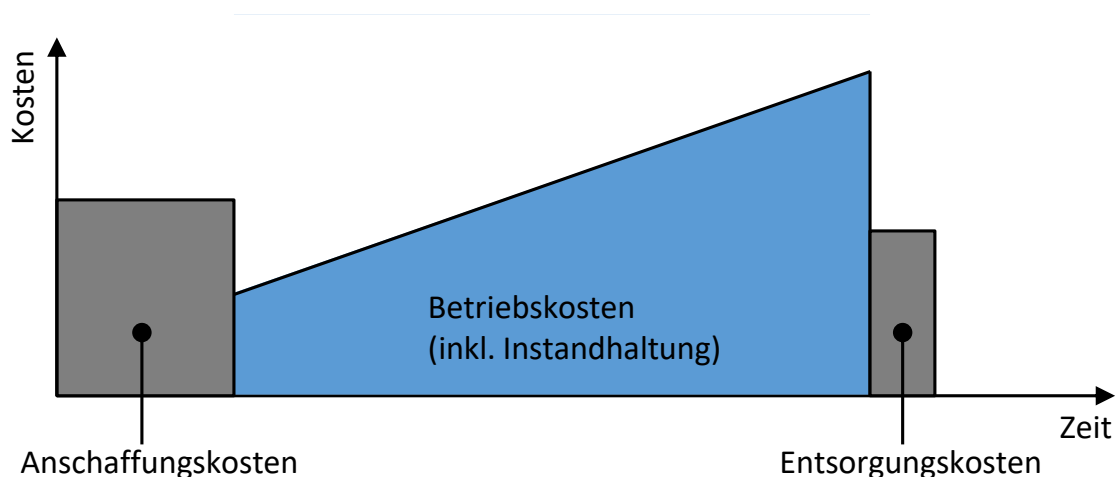


Bild 2.1: Anteile der Kosten über die verschiedenen Nutzungsphasen [20]

Die Kosten über den Lebenszyklus in Anbetracht der unterschiedlichen Verursacher zeigt das Bild 2.1. Aus diesem ist ersichtlich, dass die Betriebskosten und somit gerade auch die Instandhaltungskosten in der Regel den größten Teil der Kosten ausmachen.

2.1.1 Technische Verfügbarkeit

Jede Produktion arbeitet naturgemäß auf eine technische Verfügbarkeit V_T von 100 % hin, welche nur durch die Reduzierung von Ausfällen erreicht werden kann. Dieses hat bereits auch Linß erkannt und wie folgt beschrieben: „Durch eine gezielte und kontrollierte Überwachung der verschleißenden Bauteile lassen sich Fehler, Störungen, Ausfälle und Schäden vermeiden und damit die technische Zuverlässigkeit und technische Verfügbarkeit von Produktionssystemen erhöhen.“ [21]. Hieraus begründet sich dann auch die Notwendigkeit der Umsetzung einer geplanten Instandhaltung in der Produktion.

Die technische Verfügbarkeit V_T , welche mit dem technischen Nutzungsgrad N_T gleichzusetzen ist, beschreibt die Zeit ohne Ausfälle als Prozentangabe. Diese lässt sich aus der geplanten Belegungszeit T_B ² und der technischen Ausfallzeit T_T bestimmen. [22]

$$V_T = N_T = \left(1 - \frac{T_T}{T_B}\right) * 100\% \quad (2.1)$$

Das Maß der technischen Verfügbarkeit V_T nach Formel 2.1 ist sehr gut für die Angabe des Verhältnisses zwischen der Belegungszeit T_B und der technischen Ausfallzeit T_T geeignet. Allerdings werden organisatorische Ausfallzeiten T_O und Wartungszeiten T_W nicht berücksichtigt. Hierfür eignet sich die Kennzahl des Gesamtnutzungsgrads N_G nach Formel 2.2.

$$N_G = \frac{T_N}{T_B} = \frac{T_B - T_O - T_T - T_W}{T_N + T_O + T_T + T_W} * 100 \% \quad (2.2)$$

Es wird dabei davon ausgegangen, dass Wartungen und damit Wartungszeiten T_W in „nicht belegten Zeiten“ erfolgen. Organisatorische Ausfallzeiten T_O sind beispielsweise bestimmt durch Fehlen von Energie, Bauteilen, Materialien oder auch mangelnder Qualifikation. Diese Zeiten können nicht im Rahmen von geplanter Instandhaltung reduziert werden und werden der Einfachheit halber in dieser Arbeit vernachlässigt.

In der Arbeit wird im Folgenden ausschließlich die technische Verfügbarkeit V_T verwendet. Diese ist eine maßgebliche Kennzahl zur Bewertung des Mehrwerts einer voraussagenden Instandhaltung.

² $T_B = T_N + T_O + T_T + T_W$, T_B die Summe der Nutzungszeit T_N , der organisatorischen Ausfallzeit T_O , der technischen Ausfallzeit T_T und der Wartungszeit T_W [22].

2.1.2 Eigenschaften von Ausfällen

Nach DIN 13306 wird die „Beendigung der Fähigkeit einer Einheit, eine geforderte Funktion zu erfüllen“ als Ausfall bezeichnet. Somit steht ein Ausfall eines Produktionssystems an, sobald ein Funktionselement seine elementare Funktion nicht mehr erfüllt. [23]

Ein Ausfall wird nach folgenden Aspekten bewertet [24]:

Beeinträchtigungsumfang Je nachdem welcher Ausfall vorliegt, kann dieser das komplette System oder nur einen Teil betreffen, wodurch ein weiterer Betrieb des Systems möglich ist. Es wird daher zwischen einem Voll- und Teilausfall unterschieden.

Änderungsgeschwindigkeit Eine weitere Charakterisierung eines Ausfalls ist die Eintrittsgeschwindigkeit. Ist ein Ausfall im Vorfeld erkennbar, handelt es sich um einen Driftausfall, welcher bei Abnutzung entsteht. Wenn es jedoch zu einem Ausfall durch Bedienungsfehler oder Ermüdung von Materialien kommt, handelt es sich um einen Sprungausfall.

Bei Driftausfällen wird der Abstand zwischen dem Zeitpunkt eines potenziellen Ausfalls (P) bis zum funktionalen Ausfall (F) P-F-Intervall genannt. Dieses Intervall ist in Bild 2.2 dargestellt und ist maßgeblich dafür, dass ein funktionaler Ausfall einer Komponente vorhergesagt werden kann. Ohne P-F-Intervall ist ein Komponentenausfall nicht vorhersehbar, dies kann z.B. bei elektrischen Komponenten der Fall sein. Wenn das P-F-Intervall zu klein ist, reicht die Reaktionszeit für die Instandhaltung ebenfalls gegebenenfalls nicht aus. [25]

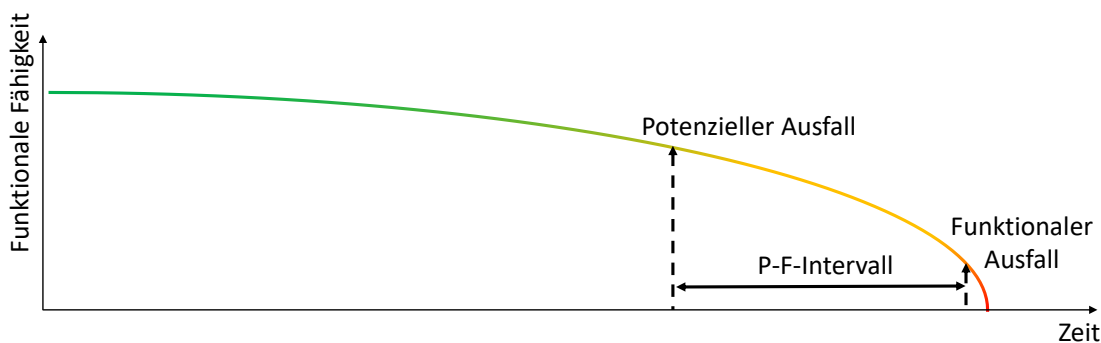


Bild 2.2: Schematischer Verlauf eines Driftausfalls bei Abnutzung einer Komponente mit P-F-Intervall (vgl. [25]): Dargestellt ist ein Driftausfall mit einem exemplarischen P-F-Intervall. Als P-F-Intervall wird die Dauer zwischen dem potenziellen und funktionalen Ausfall definiert.

Ausfallursache Es führen verschiedene Ursachen zu Ausfällen. Die Ursache kann in die vier Kategorien Entwurf, Fertigung, Abnutzung und intermittierend eingeteilt werden.

- Ein entwurfsbedingter Ausfall ist bereits durch die Auslegung vorherbestimmt, wenn zum Beispiel bei der Konstruktion von anderen Belastungswerten ausgegangen wird als auftreten werden, wodurch die

Komponenten falsch ausgelegt werden und damit irgendwann versagen werden.

- Ein fertigungsbedingter Ausfall entsteht in der Fertigung durch die Verwendung von falschen Materialien oder dem nicht-korrekten Einbau.
- Der Abnutzungsausfall liegt begründet in der Beanspruchung³ von Komponenten. Durch Aktion einer Komponente entsteht zum Beispiel Reibung, wodurch das zugrunde liegende Material nach einer bestimmten Zeit nicht mehr funktionstüchtig ist und bricht.
- Unter intermittierendem Ausfall wird ein Ausfall bedingt durch eine temporäre Änderung der Funktionalität verstanden. Dieser kann beispielsweise durch den Test eines neuen Fertigungsprozesses auftreten.

2.1.3 Lebensdauer eines technischen Systems

Ein Produktionssystem besteht aus mehreren Komponenten, welche sich untereinander in ihrer Lebensdauer beeinflussen und somit zum frühzeitigen Ausfall führen können. In dieser Arbeit wird eine vereinfachte Herangehensweise gewählt und nur die mittlere Lebensdauer einer Komponente ohne weitere Einflüsse betrachtet.

Die Ausfallrate in Abhängigkeit der Zeit $\lambda(t)$ wird mittels der sogenannten Badewannenkurve beschrieben, siehe Bild 2.3. Dabei lassen sich drei Phasen unterscheiden, die je nach Komponente unterschiedlich durchlaufen werden. Die erste Phase ist durch Frühausfälle, die zweite Phase durch Zufallsausfälle und die dritte Phase durch Spätausfälle geprägt. [26]

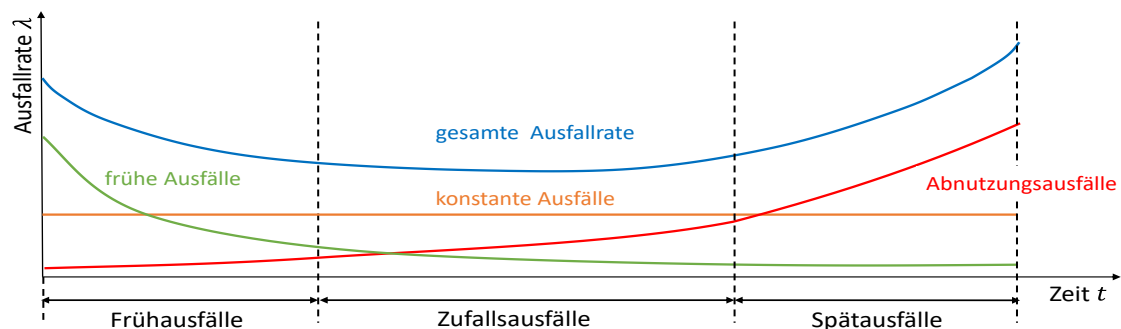


Bild 2.3: Typische Badewannenkurve der Ausfallrate für Komponenten: Dargestellt ist die gesamte Ausfallrate (in Blau) für die Komponenten eines Produktionssystems. Die Ausfallrate ergibt sich aus der Summe von frühen Ausfällen (in Grün), konstanten Ausfällen (in Orange) und Abnutzungsausfällen (in Rot).

In der ersten Phase ist für einen Ausfall in der Regel noch nicht der Verschleiß verantwortlich. Dabei handelt es sich um Konstruktions-, Werkstoff- oder Herstellungsfehler. In der zweiten Phase sind Ausfälle häufig durch Umwelteinflüsse, Überlasten, Bedienungsfehler oder fehlerhafte Wartung ausgelöst. In der

³ „Gesamtheit oder Teilgesamtheit der Einwirkungen, denen die Einheit ausgesetzt ist, wird oder ausgesetzt sein kann.“ [24]

dritten Phase wirken zusätzlich noch Verschleiß oder Ermüdung auf das System und bewirken die Ausfälle. [20]

Die entstehenden Frühausfälle werden innerhalb dieser Arbeit nicht weiter betrachtet, da von einem bereits erprobten und fehlerfreien System zu Beginn der Betriebsphase auszugehen ist. Diese Vorgabe ist z.B. bei der Daimler AG im Lastenheft festgehalten [27]. In der Produktion wird in der Regel dafür mittels einer Abnahme die Zuverlässigkeit⁴ des Produktionssystems im Vorfeld getestet.

Zufallsausfälle sind meist Spontanausfälle und besitzen daher ein geringes P-F-Intervall, es ist daher meist nur möglich diese Ausfälle im Nachgang zu diagnostizieren. Durch die korrekte Nutzung der Komponenten werden diese bestmöglich vermieden.

In der dritten Phase wird die Ausfallrate durch den Verschleiß bestimmt. Dieser erfüllt die Eigenschaften eines hinreichend großen P-F-Intervalls und ist bei der Erfassung der relevanten Merkmale im Vorfeld erkennbar. Somit konzentriert sich eine mögliche Vorhersage auf Verschleißausfälle.

In dieser Arbeit wird der Fokus somit auf die Spätausfälle gelegt, welche durch Verschleiß oder Ermüdung von Komponenten auftreten. Die mittlere Lebensdauer von technischen Systemen, Teilsystemen und Komponenten wird mittels der Mean-Time-Between-Failure (MTBF) beschrieben, wobei der erste Einsatz bis zum funktionalen Ausfall bei reparierbaren Systemen als Mean-Time-To-First-Failure (MTTFF) und bei nicht reparablen Systemen als Mean-Time-To-Failure (MTTF) bezeichnet wird, siehe Bild 2.4. Die Reparaturdauer wird im Englischen als Mean-Time-To-Repair (MTTR) bezeichnet. [21]



Bild 2.4: Grafische Beschreibung der Begriffe MTTF, MTTFF, MTBF und MTTR (vgl. [21]): Abgebildet ist links die nicht reparierbare Einheit mit ausschließlich einer Mean-Time-To-Failure (MTTF). Rechts ist die reparierbare Einheit dargestellt, welche durch eine Mean-Time-To-First-Failure (MTTFF), Mean-Time-To-Repair (MTTR) und künftig durch eine Mean-Time-Between-Failure (MTBF) charakterisiert wird.

Übergreifend kann zwischen der Mean-Time-Between-Failure (MTBF), der Mean-Time-To-Repair (MTTR), der geplanten Belegungszeit T_B und der technischen Ausfallzeit T_T folgender Zusammenhang hergeleitet werden.

$$V_T = \left(1 - \frac{T_T}{T_B}\right) * 100\% \Leftrightarrow \left(1 - \frac{\sum MTTR}{\sum(MTBF + MTTR)}\right) * 100\% \quad (2.3)$$

⁴ Unter Zuverlässigkeit wird verstanden, dass das System eine vorgegebene Eignung erfüllt. Diese Eignung kann im Rahmen eines Tests über eine definierte Zeitspanne erfolgen. [24]

2.1.4 Einflussgrößen auf die Lebensdauer von Komponenten

Die Einflüsse auf die Lebensdauer von Komponenten eines Produktionssystems sind vielseitig und wurden bereits nach Linß [21] als Ausfallursachen definiert. In das im Bild 2.5 dargestellte Ursache-Wirkungs-Diagramm wurden diese sinngemäß überführt und mittels Expertenwissens ergänzt. Das Ursache-Wirkungs-Diagramm oder auch Ishikawa-Diagramm und Fischgräten-Diagramm genannte Diagramm stellt die Einflussfaktoren auf die Lebensdauer von Komponenten dar. Mittels der herkömmlichen „5M“-Methode Maschine, Mensch, Material, Methode und Mitwelt wurden die Einflüsse klassifiziert. Die Angabe der mittleren Lebensdauer von Komponenten erfolgt nach einer Analyse dieser Faktoren. Bei Veränderung der Einflussgrößen hat dies eine Auswirkung auf die Lebensdauer. Denn nicht nur die Betriebsdauer ist relevant, auch störende Einflüsse wie Temperaturveränderung, Dimensionierungsfehler oder Fremdeinwirkungen können zur Reduzierung der Lebensdauer und einem Ausfall führen.

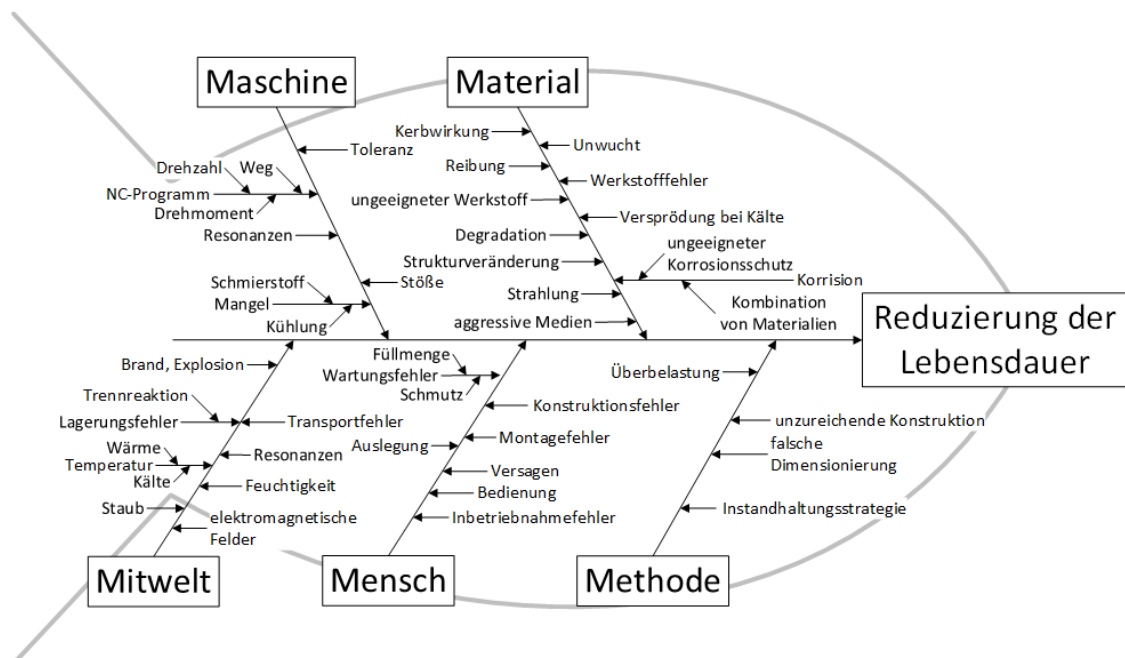


Bild 2.5: Ursache-Wirkungs-Diagramm der Einflussgrößen auf die Lebensdauer von Komponenten

In dieser Arbeit wird die Angabe der Lebensdauer MTBF als ideal angesehen. Das bedeutet, weitere Einflussgrößen auf die Lebensdauer von Komponenten werden nicht berücksichtigt, dennoch können durch eine Modifikation der Lebensdauer MTBF die meisten der Einflussgrößen bei Bedarf ausreichend berücksichtigt werden.

2.1.5 Fertigung in der Motoren-Komponenten-Produktion

Für die Fertigung von Motoren-Komponenten kommen größtenteils Standardmaschinen, welche Bohr-, Fräs-, Dreh- oder Schleifverfahren verwenden, zum Einsatz. Diese können aufgrund der automatisierten Werkzeugwechsel als Bearbeitungszentren (BAZ) eingeordnet werden. Die Einordnung erfolgt nach dem

Automatisierungsgrad und da herkömmliche Werkzeugmaschinen erst einmal keine programmierten, wiederholenden Abläufe absolvieren können, ist der oftmals umgangssprachliche Gebrauch des Begriffes Werkzeugmaschinen nicht korrekt. [28]

In Rahmen dieser Arbeit werden Fertigungszellen betrachtet. Diese beinhalten die Funktionen von CNC⁵-Maschinen und auch der von Bearbeitungszentren, welche durch programmierte und wiederholende Abläufe sowie automatisierte Werkzeugwechsel charakterisiert sind. Zusätzlich sind Fertigungszellen durch einen automatisierten Be- und Entladevorgang von Werkstücken und einer Fertigung nach Losgröße 1⁶ gekennzeichnet. Der Funktionsumfang von Werkzeugmaschinen ist in Bild 2.6 dargestellt.

+ mehrere sich ergänzende und ersetzende Maschinen + automatisierter WSt-Fluss + automatisierte Transportüberwachung und Fertigungssteuerung über Leitrechner + ergänzende Einrichtung wie Mess- und Reinigungsmaschinen			FZ	FS (TS)
+ automatisierte WSt-Be- und Entladeeinrichtung aus eigenem WSt-Speicher + automatisierte Fertigungssteuerung + flexible Werkstückfolge + integrierte Überwachungs- und Messeinrichtung		BAZ		
+ automatisierter Werkzeugwechsel aus eigenem WZ-Speicher + automatisierter WSt-Palettenwechsel und WSt-Überwachung + Mehrseiten-/Komplettbearbeitung (4. und 5. NC-Achse) + automatisierter Programmaufruf				
+ automatisierter Arbeitszyklus für gleiche Werkstücke (WSt) + Ablaufsteuerung der Maschinenfunktionen (+ Werkzeugrevolver)			CNC- Maschine	
Erzeugung der Schnitt- und Vorschubbewegungen sowie der Prozesskräfte	konv. CNC	Fertigungszelle		
			Fertigungssystem (Transferstraße)	

Bild 2.6: Funktionsumfang von Werkzeugmaschinen (Ausschnitt vgl. [28])

Diese Arbeit wird sich weiterführend mit einem Fertigungssystem (FS) befassen, da die Einflüsse bei den Fertigungszellen aufgrund deren niedrigere Komplexität⁷ als geringer angenommen werden können, siehe Bild 2.5 im vorherigen Abschnitt. Der Aufbau von Fertigungszellen im Vergleich zu Fertigungssystemen ist in der Regel sehr ähnlich. Die Hauptbaugruppen wie Gestellbauteile, welche auch die Struktur der Maschine vorgeben, bestehen aus Werkzeug- und Werkstückaufnahme, Spindel, Antrieben und in der Regel Kugelgewindetriebe⁸ und Führungen zur Übertragung der Bewegungen auf die einzelnen Achsen. Neben dieser sind Versorgungs-, Entsorgungs-, Automatisierungs-, Sicherheits- und Bedieneinheiten

⁵ engl.: Computerized Numerical Control, deutsch: rechnergestützte numerische Steuerung

⁶ Losgröße 1 bedeutet das jedes Bauteil individuell und flexibel automatisiert hergestellt werden kann [29].

⁷ Methode zur Untersuchung der Komplexität von Produktionsanlagen, siehe [30].

⁸ Der Kugelgewindetrieb ist ein Spindel-Mutter-Getriebe mit Wälzelementen aus Kugeln [28].

notwendig. Zusätzlich sorgt bei Fertigungssystemen der automatisierte Werkstückfluss zu Einflüssen von anderen Systemen auf das einzelne System.

Die Steuerung von Abläufen und Antrieben erfolgt über die speicherprogrammierbare Steuerung (SPS) und numerische Steuerung⁹ (NC) in der Maschine, die interne Struktur der Steuerung für CNC-Maschinen. Gemeint sind damit allgemein alle Werkzeugmaschinen (Bild 2.6) mit einer Steuerung wie in Bild 2.7 dargestellt. Das technische System, also die Werkzeugmaschine, mit den Sensoren und Aktoren befindet sich in der sogenannten Feldebene. Übergeordnet ist die Steuerungsebene zuständig dafür, Daten von Sensoren, heutzutage meist über ProfiBus oder gängiger ProfiNet, zu empfangen, zu verarbeiten und dann an die Aktoren zu senden. Die SPS und NC arbeiten bei Werkzeugmaschinen zusammen, allerdings hat jede seinen eigenen Aufgabenbereich: So wird die SPS für einfachere Steuerungsaufgaben verwendet wie die Ansteuerung von Ventilen. Auch die Auswertung von Druckschaltern wird über eine SPS realisiert. Die NC hingegen hat eine Echtzeitfähigkeit und feste Zykluszeiten der Programme sind gewährleistet, wodurch die Positionierung von Achsen und rechtzeitige Auswertung der Wegmesssystemsignale möglich ist.

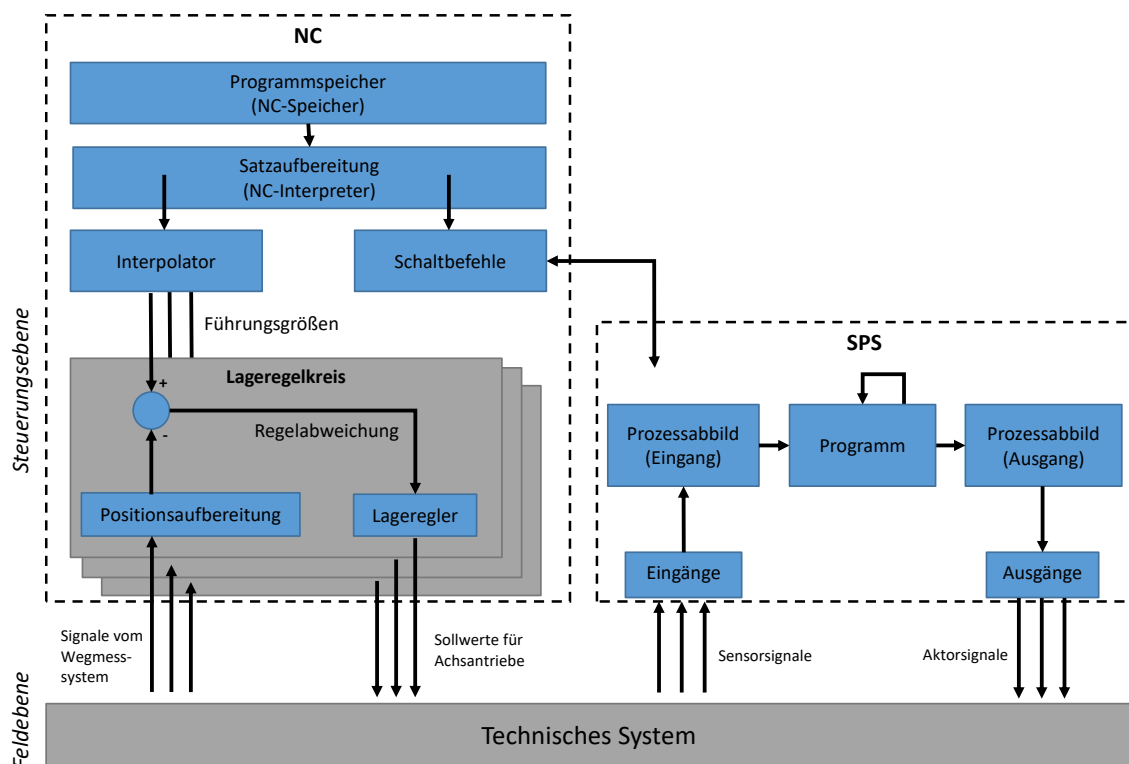


Bild 2.7: Interne Struktur der Steuerung von Werkzeugmaschinen (in Anlehnung [31 bis 34]): Dargestellt ist der Informationsfluss in der NC und SPS und die Schnittstelle zu einem technischen System.

⁹ engl.: Numeric Control

2.2 Instandhaltungsstrategien und -werkzeuge

In diesem Abschnitt werden die Mehrwerte der unterschiedlichen Instandhaltungsstrategien erläutert. Anschließend wird eine Marktanalyse zu den Werkzeugen für die Instandhaltung mit Fokus auf der voraussagenden Instandhaltung durchgeführt, da dieses Themengebiet das künftige Potenzial zur Steigerung der Wettbewerbsfähigkeit beinhaltet. Nachdem die Werkzeuge bekannt sind, wird auf die Herausforderungen beim Einsatz von Predictive Maintenance (PM) eingegangen und damit die Fragestellung der Arbeit näher erläutert.

2.2.1 Mehrwerte der verschiedenen Instandhaltungsarten

Während des Lebenszyklus einer Einheit (Produktionssystem oder Komponente) ist es notwendig, diese zu warten, zu inspizieren, instand zu setzen oder zu optimieren. Diese Grundmaßnahmen übernimmt die Instandhaltung, wodurch ein funktionsfähiger Zustand aufrechterhalten wird. [23, 35]

Jedoch ist die Instandhaltung, je nach Art, zeitaufwändig und/oder kostenintensiv. Deswegen müssen die Vor- und Nachteile der jeweiligen Strategie immer sorgfältig gegeneinander abgewogen werden.

Die Instandhaltung kann in verschiedene Arten unterteilt werden und wird nach DIN EN 13306 [23] wie folgt definiert:

- reaktive Instandhaltung
- präventive Instandhaltung

vorausbestimmte Instandhaltung

zustandsorientierte Instandhaltung

voraussagende Instandhaltung

Die **reaktive Instandhaltung**, auch bekannt als Feuerwehrstrategie, löscht in Ausfallsituationen, welche zum Stillstand der technischen Anlage führen, sprichwörtlich das Feuer. Das heißt die Instandhaltung reagiert nur im Fall eines Produktionsausfalles und muss dann sofort die ausgefallene Einheit instand setzen, sodass diese wieder in einem funktionsfähigen Zustand ist. [23]

Die Nachteile sind, dass bei mehreren gleichzeitigen Ausfällen unter Umständen keine Instandhalter verfügbar sind oder ggf. unnötige Personalkosten durch Leerlaufzeiten entstehen. Des Weiteren müssen viele Ersatzteile¹⁰ bevorratet werden, wodurch hohe Lagerkosten und ein zeitlich vorgeordneter Invest entsteht. Auch der Produktionsausfall führt zu Kosten für Bedienpersonal und der Ausblieb von produzierten Produkten. Ein weiteres Risiko der reaktiven Instandhaltung ist das durch

¹⁰ „Einheit zum Ersatz einer entsprechenden Einheit, um die ursprünglich geforderte Funktion der Einheit zu erhalten“ [23].

den Primärausfall¹¹ ein Sekundärausfall¹², auch Folgeausfall genannt, entsteht, wodurch weitere Instandsetzungen notwendig werden und Folgekosten auftreten.

Jedoch gibt es auch Vorteile durch die ausfallbasierte Instandhaltung wie die Nutzung des kompletten Abnutzungsvorrates einer Komponente. Dieser bezeichnet die Dauer, bis eine Einheit ihre Funktion nicht mehr erfüllen kann. [35] Dadurch wird die Einheit so lange wie möglich genutzt und die Kosten für den frühzeitigen Austausch vermieden. Daher kann die reaktive Instandhaltung gerade auch hinsichtlich Ressourcenschonung und Nachhaltigkeit eine gewünschte und vorteilhafte Strategie sein.

Präventive Instandhaltungsarten wie die folgenden genannten haben hingegen den Vorteil Produktionsausfälle in produktiven Zeiten zu vermeiden.

Die **vorausbestimmte Instandhaltung** wird geplant und nach einer festen Anzahl oder Zeit durchgeführt [23]. Der Zeitpunkt kann von der Betriebsdauer der Einheit, einer produzierten Stückzahl oder einer festen Anzahl an Schaltzyklen wie dem Öffnen eines Ventils abhängen und wird im Vorfeld festgelegt.

Ein wesentlicher Nachteil der vorausbestimmten Instandhaltung ist es, dass der Zustand der Einheit irrelevant für die Instandhaltung ist. Dadurch werden die Einheiten oftmals zu früh getauscht, wodurch hohe Aufwände in der Instandhaltung und somit eigentlich unnötige Ersatzteil- und Personalkosten entstehen. Auch die Produktion ist davon betroffen, da die Nutzungszeit T_N der Anlage geringer eingeplant werden kann als theoretisch möglich wäre.

Ein Vorteil dieser Instandhaltungsstrategie ist dennoch, dass Ausfälle in der Belegungszeit T_B verhindert und ein ausfallfreier Betrieb weitestgehend gewährleistet werden kann. Auch die Notwendigkeit von Ersatzteilen kann geplant werden, wodurch die Lagerkosten reduziert werden können.

Bei der **zustandsorientierten Instandhaltung** wird zunächst eine Zustandsüberwachung oder Inspektion durchgeführt. Die Inspektion, auch Konformitätsprüfung genannt, untersucht den Zustand der Einheit durch Messungen, Beobachtungen oder Tests zu einem bestimmten Zeitpunkt. [23] Zusätzlich oder stattdessen wird vermehrt auf eine Zustandsüberwachung der Einheiten, bekannt als Condition Monitoring (CM), gesetzt. Hierbei werden Merkmale der Einheit kontinuierlich gemessen und bewertet. Dies erfolgt teils automatisch, aber auch manuell durch die Vorgabe von festen Grenzwerten.

Die Nachteile dabei sind der teils hohe Aufwand zur Durchführung einer Inspektion oder auch die korrekte Parametrisierung der Überwachungsgrenzen der Zustandsüberwachung. Ein Nachteil von Inspektionen kann auch sein, dass die Abnutzung

¹¹ „Ausfall einer Einheit, der weder direkt noch indirekt durch einen Ausfall oder einen Fehler einer anderen Einheit verursacht wurde“ [23].

¹² „Ausfall einer Einheit, der entweder direkt oder indirekt durch einen Ausfall oder einen Fehler einer anderen Einheit verursacht wurde“ [23].

von außen nicht erkennbar ist und teilweise, z.B. mangels der Erreichbarkeiten von Komponenten im Maschinenraum, nicht möglich ist.

Die Vorteile der zustandsorientierten Instandhaltung sind, dass die Einheit bis zur Abnutzungsgrenze betrieben werden kann und somit relativ gut genutzt wird. Dieses Vorgehen spart Ersatzteilkosten ein. Außerdem werden oftmals Ausfälle in der Produktionszeit verhindert, wodurch eine gute technische Verfügbarkeit erzielt werden kann.

Die **voraussagende Instandhaltung** basiert darauf den Zustand einer Einheit zu analysieren und basierend auf Vergangenheitsinformationen eine Vorhersage abzuleiten [36]. In diesem Fall würde der Ausfall einer Einheit vorhergesagt werden.

Zu den Vorteilen dieser Vorgehensweise zählen die optimale Belegungszeit T_B der Anlage als auch die geplante Instandhaltung, welche eine verbesserte, maximale Nutzung der Einheit ermöglicht. Neben der in Bild 2.8 dargestellten Mehrwerte ergeben sich weitere Vorteile durch einen lückenlosen Informationsfluss, ein effizientes Betriebsmittelmanagement und eine verbesserte Wettbewerbsfähigkeit. Letztere wird unter anderem durch die erhöhte Anlagenverfügbarkeit, also der gesteigerten technischen Verfügbarkeit, erzielt.

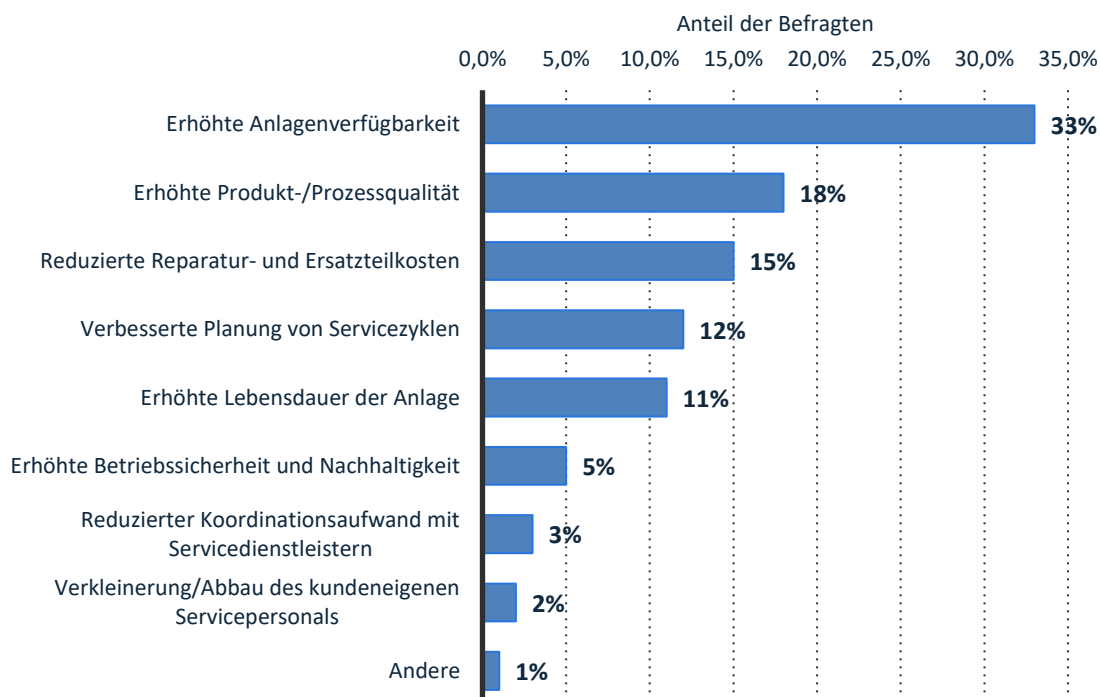


Bild 2.8: Mehrwerte durch eine voraussagende Instandhaltung [37]: Dargestellt ist eine Umfrage des Unternehmens Roland Berger zur Fragestellung: „Worin liegt Ihrer Einschätzung nach der Kundennutzen von Predictive Maintenance?“. Befragt wurden 153 Personen von Unternehmen aus den Branchensegmenten der Antriebs-/Fluidtechnik, der elektrischen Automation/Robotik, der diskreten Fertigungstechnik sowie der Software- und Digitalisierungstechnik.

Basierend auf der voraussagenden Instandhaltung ist es möglich, weitere Informationen bereitzustellen. Durch die Entscheidungsunterstützung kann der Instandhalter schneller reagieren und Prozesse wie die Ersatzbeschaffung oder die

Personalplanung können teil- oder vollautomatisiert erfolgen. Die höheren Implementierungskosten sollten jedoch immer mit den Einsparungen hinsichtlich einer gesteigerten technischen Verfügbarkeit abgeglichen werden. Möglicherweise ist in Einzelfällen die Strategie einer reaktiven, vorausbestimmten oder zustandsorientierten Instandhaltung wirtschaftlich besser geeignet. Dies wird die Arbeit im siebten Kapitel nochmal genauer untersuchen.

Veranschaulicht zeigt Bild 2.9 die Betriebszustände je Instandhaltungsstrategie auf. Dabei werden die Vorteile einer voraussagenden Instandhaltungsstrategie wie eine kurze Wartezeit, keine weitere Ausfallzeit und eine optimale Nutzung des Abnutzungsvorrats der Komponenten deutlich. Es ist somit festzuhalten, dass bei einer Bewertung der technischen Verfügbarkeit die voraussagende Instandhaltung die besten Ergebnisse erzielt. Dennoch ist nicht jede Komponente für diese Art der Instandhaltung geeignet, was in Abschnitt 2.1.2 entsprechend diskutiert wurde.

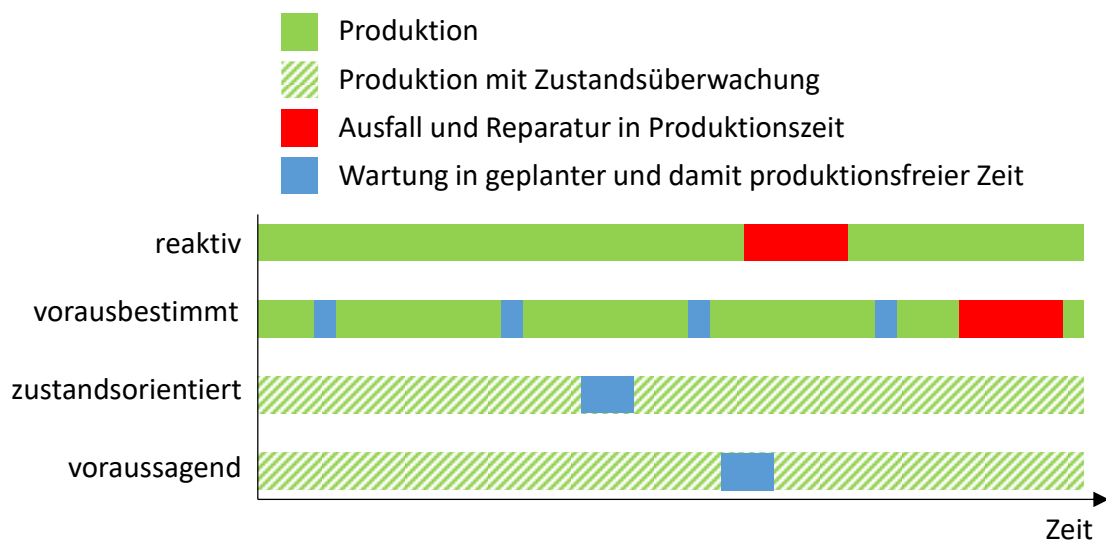


Bild 2.9: Betriebszustände je Instandhaltungsstrategie (in Anlehnung [38]): Dargestellt sind für die reaktive, vorausbestimmte, zustandsorientierte und voraussagende Instandhaltungsstrategie die Betriebszustände und damit auch Ausfall- und Wartezeiten zu den Produktionszeiten.

Zusammenfassend kann man sagen, dass alle Instandhaltungsstrategien Vorteile und auch Nachteile besitzen. Die Wahl der Strategie sollte daher immer von den folgenden Faktoren abhängig sein: Belegungszeit, Ersatzteilkosten, Personalverfügbarkeit, Einheiten der technischen Anlage und die Vorhersagbarkeit von deren Ausfällen sowie Umsetzungskosten für die jeweilige Instandhaltungsstrategie. Allerdings sagen in einer Umfrage im Jahr 2017, dargestellt in Bild 2.10, bereits 77 % der Befragten, dass eine voraussagende Instandhaltung (Predictive Maintenance) notwendig sei und dass diese eine Voraussetzung sei, um künftig auch wettbewerbsfähig zu bleiben.

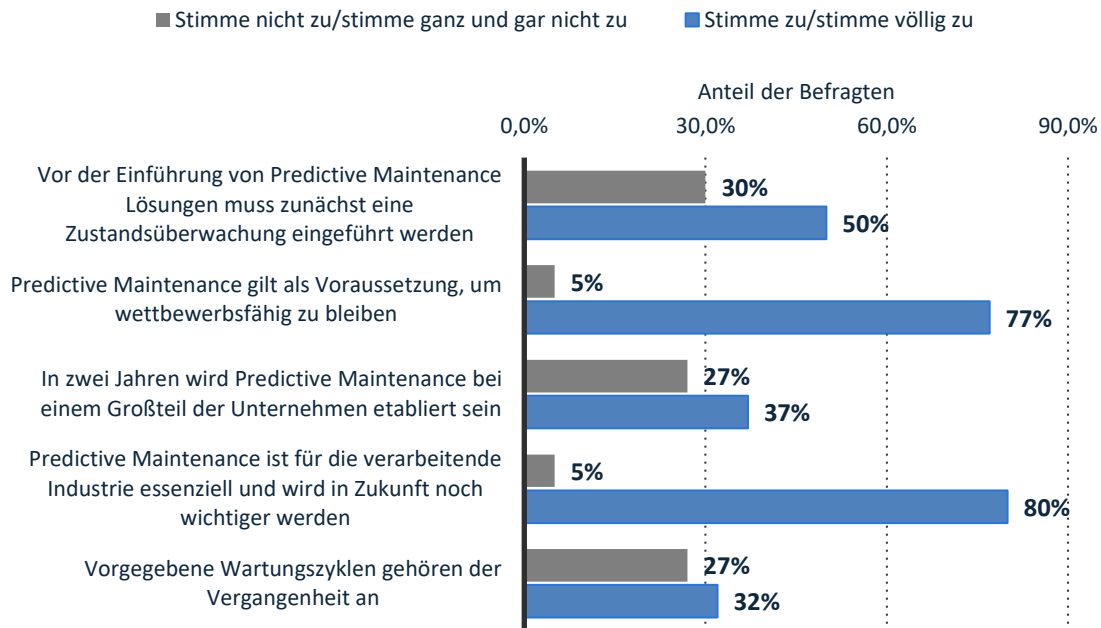


Bild 2.10: Umfrage zu Predictive Maintenance [39]: Dargestellt ist eine Umfrage zur Fragestellung: „Stimmen Sie den nachfolgend genannten Aussagen zu oder nicht zu?“. Befragt wurden ca. 290 Personen aus der verarbeitenden Industrie.

2.2.2 Marktanalyse der Software-Werkzeuge für Predictive Maintenance

Laut einer Umfrage, zusammengestellt in Bild 2.11, sind 11% der Unternehmen der Meinung, es gäbe ein vollumfängliches Produkt-/Leistungsangebot an Werkzeugen für die Realisierung von Predictive Maintenance. 30% der Unternehmen sind der Meinung, es gäbe ein Basisangebot am Markt. [40]

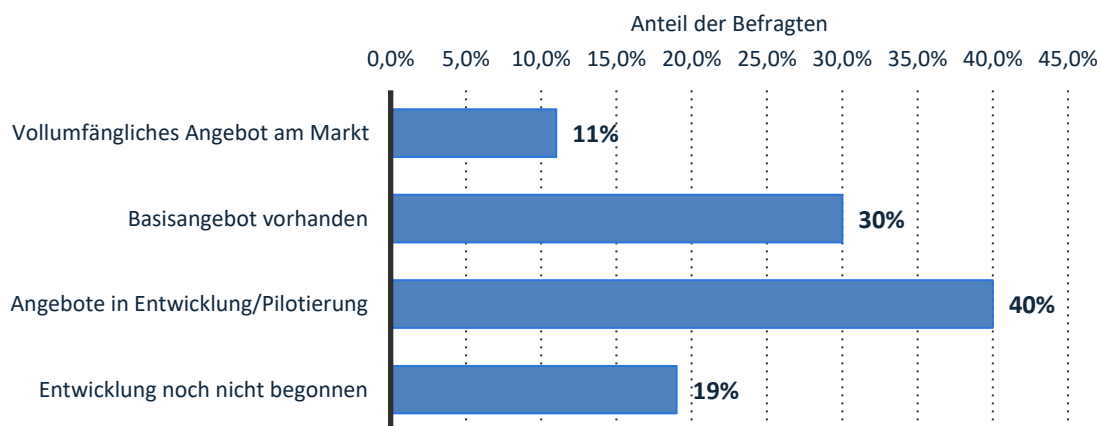


Bild 2.11: Werkzeuge für Predictive Maintenance am Markt [40]: Dargestellt ist eine Umfrage des Unternehmens Roland Berger zur Fragestellung: „Inwiefern besteht in Ihrem Unternehmen bereits ein Produkt-/Leistungsangebot im Bereich Predictive Maintenance?“. Befragt wurden 153 Personen von Unternehmen aus den Branchensegmenten der Antriebs-/Fluidtechnik, der elektrischen Automation/Robotik, der diskreten Fertigungstechnik sowie der Software- und Digitalisierungstechnik.

Dieser Abschnitt zeigt verfügbare Produkte in der Motoren-Komponenten-Produktion in der Automobilbranche und deren jeweiligen (zum Teil beschränkten)

Funktionsumfang auf, wodurch der noch notwendige Handlungsbedarf in Bezug auf Predictive Maintenance deutlich werden sollte. Bei der Analyse wurden unter anderem sowohl die fünfzehn umsatzstärksten Werkzeugmaschinenhersteller in Europa betrachtet als auch bekannte Hersteller aus der Automobilbranche [41].

An erster Stelle der umsatzstärksten Unternehmen steht die **DMG Mori AG**. Bei diesem Unternehmen ist eine Vielzahl an Applikationen für digitale Services verfügbar, welche in Bild 2.12 aufgeführt sind. Für die Vorhersage von Komponentenausfälle eignet sich die Condition Analyser APP, welche mittels eines Sensors und Instandhaltungspakets für Industrie 4.0 ein optimales Ergebnis liefern soll. Die vorbeugende Instandhaltung in Form von Wartung basiert auf der Analyse von Kraft, Vibration, Temperatur und Schmierung.

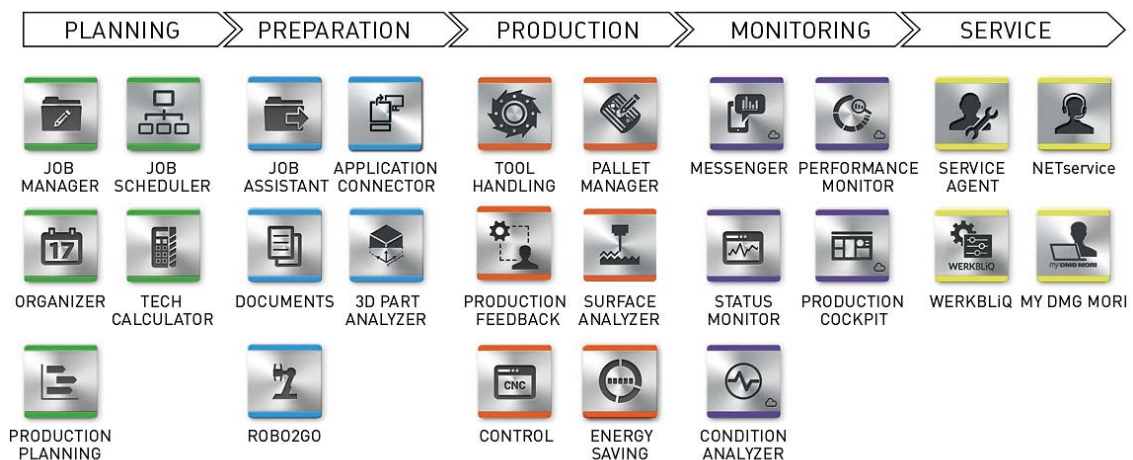


Bild 2.12: Applikationen der DMG Mori AG [42]

Auf dem zweiten Platz der umsatzstärksten Unternehmen folgt der Maschinenbauer **GROB**, welcher auch auf digitale Services setzt und mit GROB⁴Track eine Zustandsüberwachung für die Kugelgewindetriebe anbietet. Hierbei wird der Instandhalter mittels einer Nachricht über eine Zustandsänderung informiert. [43]

Die **Schweizer United Grinding Group** hat vier Bausteine für eine Umsetzung von Industrie 4.0: Production Monitor, Remote Service, Service Monitor und die IT-Sicherheit. Der erste Baustein – Production Monitor – gibt einen Überblick über die Maschine und ermöglicht es unabhängig vom Ort die Informationen zu Betriebszuständen und Kennzahlen anzuzeigen. Der Remote Service ermöglicht den Fernzugriff durch einen Spezialisten. Der Service Monitor unterstützt die Instandhaltung bei einer vorausbestimmten und intervallbasierten Instandhaltung, welche auf den Angaben des Betriebshandbuchs beruht. Der Baustein der IT-Sicherheit unterstützt den Kunden bei Themen wie Firmware- und Software-Aktualisierungen. [44]

Die **EMAG Gruppe** bietet einen Fingerprint für ihre Maschinen an, welcher den Status der Achsen von Werkzeugmaschinen anzeigt. Außerdem arbeitet die EMAG Gruppe aktuell an einem Produkt zu Predictive Maintenance, wie aktuellen Berichten von der Messe EMO 2019 zu entnehmen ist. [45 bis 47]

Die **Index Gruppe** vertreibt bereits in der Serie ihre Maschinen mit dem sogenannten iXpanel, welches unter anderem Wartungsintervalle, Stückzahlen und Produktionszeiten darstellt. [48]

Das Unternehmen **HELLER** stellt größtenteils Bearbeitungszentren her. Aufgrund der Möglichkeiten von Industrie 4.0 und damit der möglichen Digitalisierung der Prozesse bietet das Unternehmen HELLER4Industry an, welches auf der MindSphere¹³ betrieben wird. Das Produkt besteht aus drei Bausteinen: HELLER4Operation, HELLER4Services und HELLER4Performance, wobei in Bezug auf die Instandhaltung lediglich der zweite Baustein relevant ist. Das Produkt hat aktuell die Funktionalität, Zustände und Werte zu visualisieren. Die Berichterstattung lässt darüber hinaus erahnen, dass weitere Ideen und Konzepte vorhanden sind, jedoch wurde beim Vergleich der Messeauftritte von 2017 und 2019 keine technische Weiterentwicklung festgestellt. [50]

Mit CHIRON DATAline bringt die **Chiron Gruppe** eine Plattform für die Zustandsanzeige mit, welche einen Überblick über den aktuellen und historischen Status der Maschine, den Änderungen im Bearbeitungsprogramm, der Stückzahl, der OEE¹⁴ und den Wartungszählern gibt. [51]

Bei der **Hermle Gruppe** wird ein zentrales Überwachungstool im Bereich der digitalen Welt beschrieben, dieses kann den aktuellen Betriebszustand und auch den historischen Zustand sowie eine Meldungshistorie anzeigen. Weitere Funktionen sind der Export und die Benachrichtigungsfunktion per E-Mail. [52]

Zum Unternehmen **Niles Simmons Hegenscheidt** sind noch keine Veröffentlichung mit Bezug zu Predictive Maintenance zu finden.

Die **Schwäbische Werkzeugmaschinen GmbH** (SW) bietet ein breites Portfolio an Angeboten an, welche Schulung bis hin zu Zugriff auf die jeweilige Maschine umfassen. Insbesondere das Vorhaben den Kunden mit seinen Daten zu unterstützen, wird unter dem Produkt „life data“ verstanden. Hierbei geht es bisher jedoch noch um eine systematische Analyse der Daten und die manuelle Ableitung von Vorhersagen. [53]

Auf der Internetseite und auch bei der Suche mittels Suchmaschinen konnten keine Produkte im Bereich Industrie 4.0 für die **Starrag Gruppe** gefunden werden. Jedoch scheint das Unternehmen auch aktiv in dem Bereich zu sein, wie eine aktuelle Stellenausschreibung zeigt. [54]

Der Konzern **GF Milling** gehört zur Division GF Machining Solutions. Eine Internetrecherche hat gezeigt, dass die Division im Bereich des Einkaufs auf digitale Services setzt, jedoch im Bereich der Instandhaltung noch auf bewährte Technologien wie Inspektionen vor Ort und regelmäßigen Wartungen vertraut. Mit dem Produkt

¹³ „MindSphere ist das cloudbasierte, offene IoT-Betriebssystem von Siemens, das Ihre Produkte, Anlagen, Systeme und Maschinen verbindet und es Ihnen ermöglicht, die Fülle von Daten aus dem Internet der Dinge (IoT) mit umfangreichen Analysen zu nutzen.“ [49].

¹⁴ Overall Equipment effectiveness = Gesamtanlageneffektivität (GAE)

rConnect Dashboard ist 2019 bekannt geworden, dass das Unternehmen auch Industrie 4.0-Lösungen anbietet. Das Produkt bietet die Möglichkeit Maschinenkennzahlen (Status, Verfügbarkeit, Effizienz und Produktivität) aus realen Daten zu visualisieren. [55]

Bei der **Mikron Group** wird bereits von Mikron 4.0 gesprochen, allerdings bezieht sich dies auf eine stärkere Vernetzung von Maschinen und einen Wissensaustausch. Auch eine Echtzeit-Anzeige des Maschinenstatus ist möglich. [56]

Die **DVS Technology Group** lässt wenig von ihren zukünftigen Absichten erahnen. In ihren bisherigen Veröffentlichungen werden lediglich die Nutzung von Produktionsinfrastrukturen in Anlehnung an die Industrie 4.0-Prinzipien genannt. [57]

Die **Danobat Group** ermöglicht Ihren Kunden eine Darstellung des Status und Ablauf des Fertigungsprozesses, auch können bereits Trends erkannt und manuelle Handlungen abgeleitet werden. [58]

Zusätzlich zu den umsatzstärksten Unternehmen wurde noch eine Auswahl von weiteren Unternehmen, die vorrangig als Zulieferer der Automobilbranche tätig sind, betrachtet und deren Produktportfolio analysiert.

Unter ARTIS hat sich die Firma **MARPOSS Monitoring Solutions GmbH** einen Namen im Bereich der Werkzeug- und Prozessüberwachung, der adaptiven Regelung und dem Condition-Monitoring gemacht. Insbesondere die Reaktion in Echtzeit bei Fehlern als auch die effiziente und zustandsabhängige Nutzung von Werkzeugen sind Stärken des Unternehmens. [59]

Mittels des Produkts von GfM (**Gesellschaft für Maschinendiagnose mbH**) ist es möglich die Spindeln einer Werkzeugmaschine zu überwachen. Hierzu bietet die Firma den Peakanalyser an. Dieser vergleicht die aktuellen Daten mit einer vorher festgelegten Referenzkurve. Zusätzlich bieten sie an, unter Verwendung von Referenzspektren die Daten zu untersuchen. [60]

Das Unternehmen **Voith** hat bereits eine komplette Strategie „Value-driven Intelligence“ aufgebaut, um Unternehmen aus verschiedenen Branchen bei der Digitalisierung zu unterstützen. Für die Instandhaltung wird unter dem Begriff „On-Care.Asset“ eine erhöhte Wartungseffizienz fokussiert. [61]

Bewertet man die Unternehmen hinsichtlich der Produkte für die verschiedenen Instandhaltungsarten entsteht Bild 2.13. Oftmals verstehen die Unternehmen bereits ihr Angebot als Predictive Maintenance und damit als voraussagende Instandhaltung. Dies ist jedoch nicht korrekt. Das Beratungsunternehmen Roland Berger beschrieb bereits, dass der Unterschied zwischen Predictive Maintenance und Condition-Monitoring darin liegt, dass erst bei PM von einer Vorhersage auszugehen ist. Die herkömmlichen CM-Systeme ermöglichen bereits in Echtzeit den Zustand von Komponenten anzuzeigen, jedoch werden meist noch einfachere statistische oder

manuelle Verfahren verwendet, um die Daten zu analysieren. [62] Dies bedingt unter anderem die in Bild 2.13 durchgeführte Analyse der Werkzeuge je Instandhaltungsstrategie.

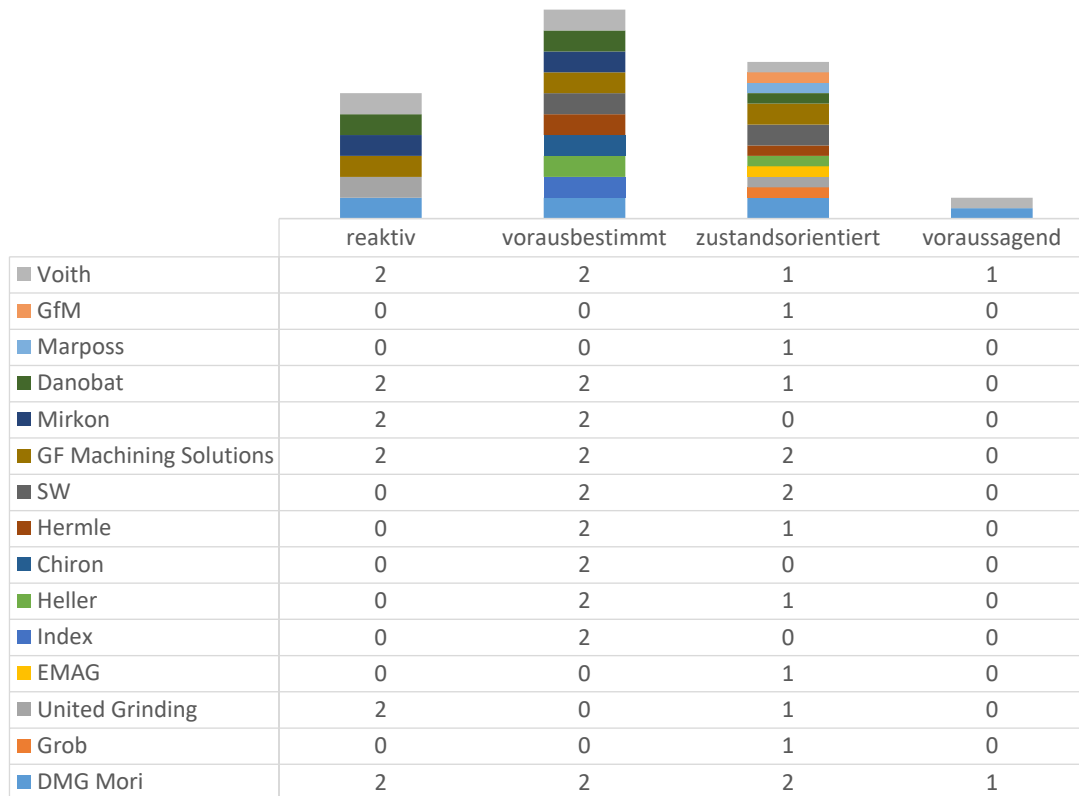


Bild 2.13: Erfüllungsgrad der Software-Werkzeuge je Instandhaltungsstrategie: Dargestellt sind zu den fünfzehn umsatzstärksten Werkzeugmaschinenhersteller der Erfüllungsgrad je Instandhaltungsstrategie. Ein Erfüllungsgrad von „0“ bedeutet das entweder kein Produkt vorhanden ist oder keine Informationen hierzu vorlagen. Bei „1“ bietet das Unternehmen teilweise ein Produkt für die entsprechende Instandhaltungsstrategie an und bei „2“ gibt es ein vollumfängliches Produkt zur Strategie.

Fazit Kein Unternehmen hat bisher ein vollumfängliches Produkt zu Predictive Maintenance, also der vorbeugenden Instandhaltung, und nur wenige Unternehmen können heute schon Teilumfänge bezüglich Predictive Maintenance abdecken. So bietet das Unternehmen DMG Mori zum Beispiel bereits eine Vorhersage mittels zusätzlicher Sensoren für einzelne Messwerte an. Im Gegensatz dazu setzt Voith auf einen ganzheitlichen Ansatz, welcher jedoch noch kein vollständiges Produkt darstellt.

2.3 Notwendigkeit der Arbeit

Aus dem Fazit des vorherigen Kapitels 2.2.2 ist zu entnehmen, dass viele Unternehmen noch kein komplettes Angebot an Produkten für die voraussagende Instandhaltung haben. Eine Umfrage, abgebildet in Bild 2.14, zeigt die Herausforderungen bei der Umsetzung von Predictive Maintenance. Daraus sind fünf Herausforderungen ersichtlich, welche mindestens 43 % der Befragten als (sehr) große Herausforderungen ansehen. Im Folgenden werden daher diese fünf Punkte IT-Sicherheit, Auswahl und Verfügbarkeit der Daten, IT-Infrastruktur, Anwendung statistischer Methoden und ausreichend große Datenbasis genauer analysiert, wodurch der Mangel an passenden Produkten für die vorausschauende Instandhaltung verdeutlicht werden soll.

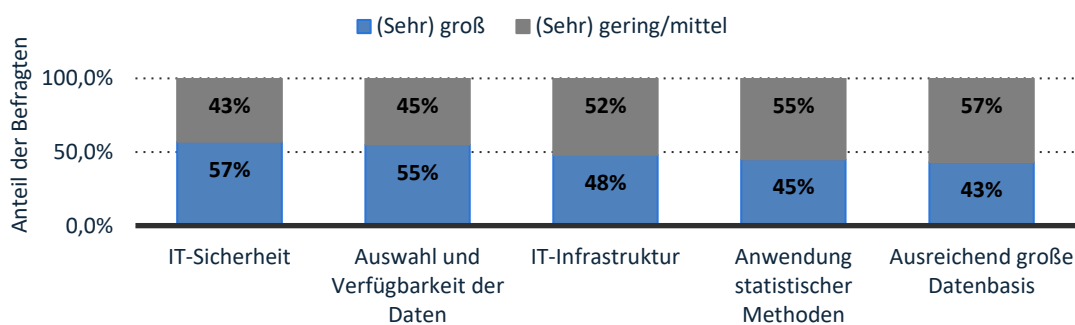


Bild 2.14: Herausforderungen bei der Umsetzung von voraussagender Instandhaltung [63]: Dargestellt ist eine Umfrage des Unternehmens BearingPoint zur Fragestellung: „Wie groß schätzen Sie folgende technische Herausforderungen für die erfolgreiche Umsetzung von Predictive Maintenance in Ihrem Unternehmen ein?“. Befragt wurden 74 Fachexperten aus dem DACH-Gebiet.

Eine generelle Voraussetzung für die Datenanalyse und damit auch für die voraussagende Instandhaltung sind die Themen IT-Sicherheit und -Infrastruktur. Diese beiden Themen haben insgesamt eine große Bedeutung, müssen aber in dieser Arbeit nicht weitergehend betrachtet, wie nachfolgende Ausführungen zeigen werden.

Die Kategorie **IT-Sicherheit** im Zusammenhang mit digitalen Services ist ein sehr relevantes Thema bei der immer stärker werdenden Vernetzung. Dabei spielen Punkte wie Datenschutz und Datenübertragung eine relevante Rolle. In entsprechenden Veröffentlichungen werden diese bereits umfassend thematisiert [29, 64]. Angriffe auf Produktionssysteme und auf dessen Steuerungen wie WannaCry [65] treten nur noch selten auf und wenn werden diese meist schnell kontrolliert und beseitigt. Dies zeigt im Kontext der Automobilindustrie, dass diese Herausforderung in großen Unternehmen durch entsprechende Prozesse bereits heute ausreichend gemeistert werden kann.

Bereits in der Einleitung wurden heutige Möglichkeiten mit performanten Speicher- und Verarbeitungslösungen vorgestellt. Sicherlich ist der Aufbau einer **IT-Infrastruktur** und die Vernetzung von Maschinen trotzdem eine Herausforderung, jedoch gibt es konkrete Vorgehensweisen zur Realisierung einer voraussagenden Instandhaltung. Hierzu haben Unternehmen bereits unter anderem Cloud-Lösungen entwickelt, die von Infrastruktur-as-a-Service (IaaS) über Plattform-as-a-Service

(PaaS) bis hin zu Software-as-a-Service (SaaS) reichen. Auch der Betrieb einer On-Premise Plattform ist eine Möglichkeit. Bild 2.15 stellt die Unterschiede der Cloud-Services in einem Schichtenmodell dar.

OEM's (Original Equipment Manufacturer) wie das Unternehmen Mercedes-Benz AG setzen auf Cloud-Lösungen und -Plattformen wie Microsoft Azure. [66]

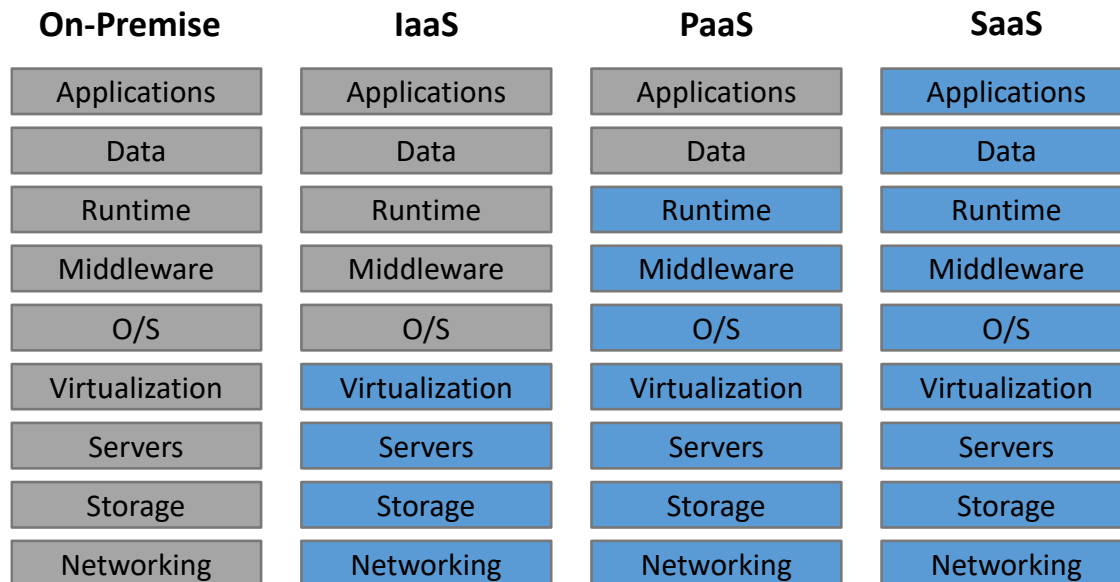


Bild 2.15: Schichtenmodell für Cloud-Services (vgl. [67]): Dargestellt sind die verschiedenen Modelle für eine Nutzung von Cloud-Services. Unterschieden wird in On-Premise, IaaS, PaaS und SaaS. Der bereitgestellte Service durch den Cloud-Betreiber ist in Blau markiert. In Grau dargestellte Felder müssen durch den Nutzer selbst organisiert werden.

Neben den IT-seitigen Voraussetzungen benötigt eine herkömmliche Datenanalyse wie die voraussagende Instandhaltung immer Daten, welche zunächst ausgewählt und verfügbar gemacht werden müssen. Mittels analytischer (u.a. statistischer) Methoden können aus den Daten neue Informationen und Erkenntnisse, bspw. die Wahrscheinlichkeit für einen bevorstehenden Ausfall, herausgefunden werden. Jedoch kann dies bei einer unzureichenden Datenbasis zu ungenügenden Ergebnissen führen. Eine nachfolgende Beschreibung der drei Herausforderungen bezüglich Daten zeigt die Notwendigkeit einer weiteren Analyse aller drei Punkte in verschiedenen Detailierungsgraden.

Die **Auswahl und Verfügbarkeit der Daten** ist sowohl eine technische als auch eine organisatorische Herausforderung. Zum einen ist dabei das Thema der Datenqualität, welche mit Aktualität, Eindeutigkeit, Einheitlichkeit, Genauigkeit, Konsistenz, Korrektheit, Redundanzfreiheit, Relevanz, Verständlichkeit, Vollständigkeit und Zuverlässigkeit beschrieben wird, relevant [68]. Zum anderen ist die Auswahl der Daten und Datenquellen eine individuelle Herausforderung, welche je nach Zielsetzung unterschiedlich angegangen werden muss. Doch auch die technischen Lösungen zur Erfassung und somit auch zur Bereitstellung von Daten sind gerade in den letzten Jahren immer einfacher geworden [29, 69]. Mittlerweile können Daten im Millisekunden Abständen erfasst und gespeichert werden, sodass sie

in (Nahe-)Echtzeit verarbeitet werden können [70]. Oftmals ist die Rede von sogenannten IoT-Gateways, welche die Erfassung von Daten aus proprietären oder nicht-performanten Quellen ermöglichen [71]. Eine aktuelle Herausforderung hierbei ist die Standardisierung der Daten bei Anbindungen von heterogenen Quellen. In Projekten wie umati (universal machine tool interface) wird deswegen versucht einen Standard zu schaffen [72].

Die Vorgehensweise zur Auswahl der korrekten Daten leitet sich von der jeweiligen Fragestellung ab. Dies entspricht einem Top-Down-Ansatz, welcher in Bild 2.16 dargestellt ist. Für die voraussagende Instandhaltung kann die Fragestellung „Wann wird meine Maschine ausfallen?“ abgeleitet werden. Hierzu muss die Datenanalyse Wissen zur Vorhersage von Ausfällen aufbauen. Dieses Wissen kann aus Informationen wie die Ähnlichkeiten von Daten einer Maschine erlangt werden. Die Grundlage für die entsprechende Analyse sind Daten wie Zeitreihen aus Sensoren und Aktoren. Mit Hilfe des Top-Down-Ansatzes könnte nun die Auswahl der Daten und der Vorgehensweise zur Beantwortung der Fragestellung geklärt werden.

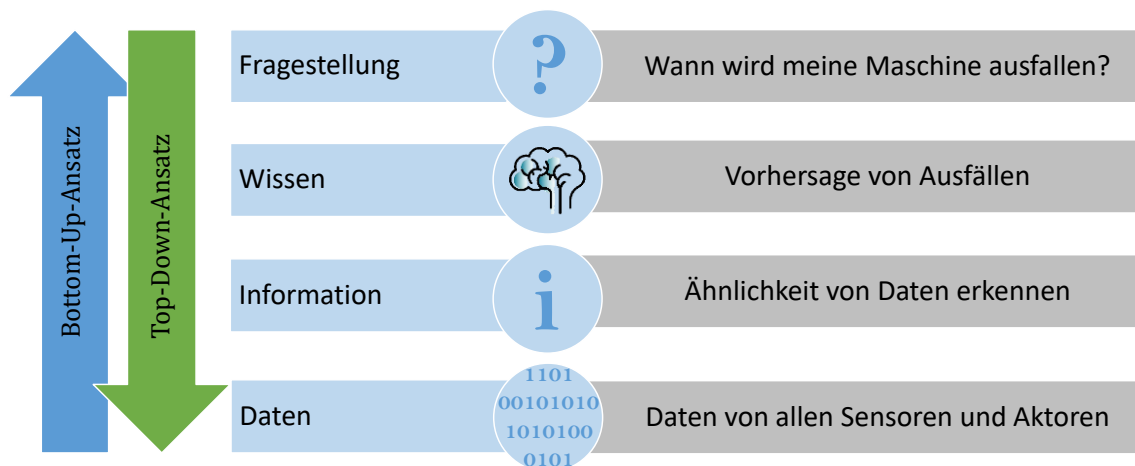


Bild 2.16: Zwei Vorgehensweisen für Datenanalysen: Dargestellt ist der Bottom-Up- (blauer Pfeil) und Top-Down-Ansatz (grüner Pfeil) für die Nutzung von Daten bzw. Beantwortung einer Fragestellung. Zusätzlich sind die Schritte Fragestellung, Wissen, Information und Daten für die voraussagende Instandhaltung skizziert.

Neben der vorhergehenden Herausforderung ist eine weitere die **Anwendung der statistischen Methoden** auf verfügbare Daten. Es gibt eine Vielzahl an Begrifflichkeiten zur Einordnung der möglichen Algorithmen. Dabei liegt der Kern und Anfang im Data Mining, welches weiterentwickelt wurde zum sogenannten Knowledge Discovery in Database (Abk.: KDD, deutsch: Wissensentdeckung in Datenbanken). Beide nutzen statistische Methoden zur Analyse von Daten. Auch die Mustererkennung (engl.: Pattern recognition) nutzt zur Analyse statistische Methoden. Weitere Methoden sind das Lernen mittels neuronaler Netze und die Themenfelder Artificial Intelligence¹⁵ (AI), Machine Learning¹⁶ (ML) und Deep Learning. Für die Verständlichkeit der verschiedenen Begriffe werden in Bild 2.17 die verschiedenen Analytik-

¹⁵ deutsch: Künstliche Intelligenz (KI)

¹⁶ deutsch: Maschinelles Lernen (ML)

Methoden geordnet und nach deren Aufwand bei der Umsetzung positioniert. Der Aufwand ist abhängig von der Menge der Trainingsdaten, der Komplexität und Nachvollziehbarkeit der Algorithmen und der notwendigen Rechenleistung [38].

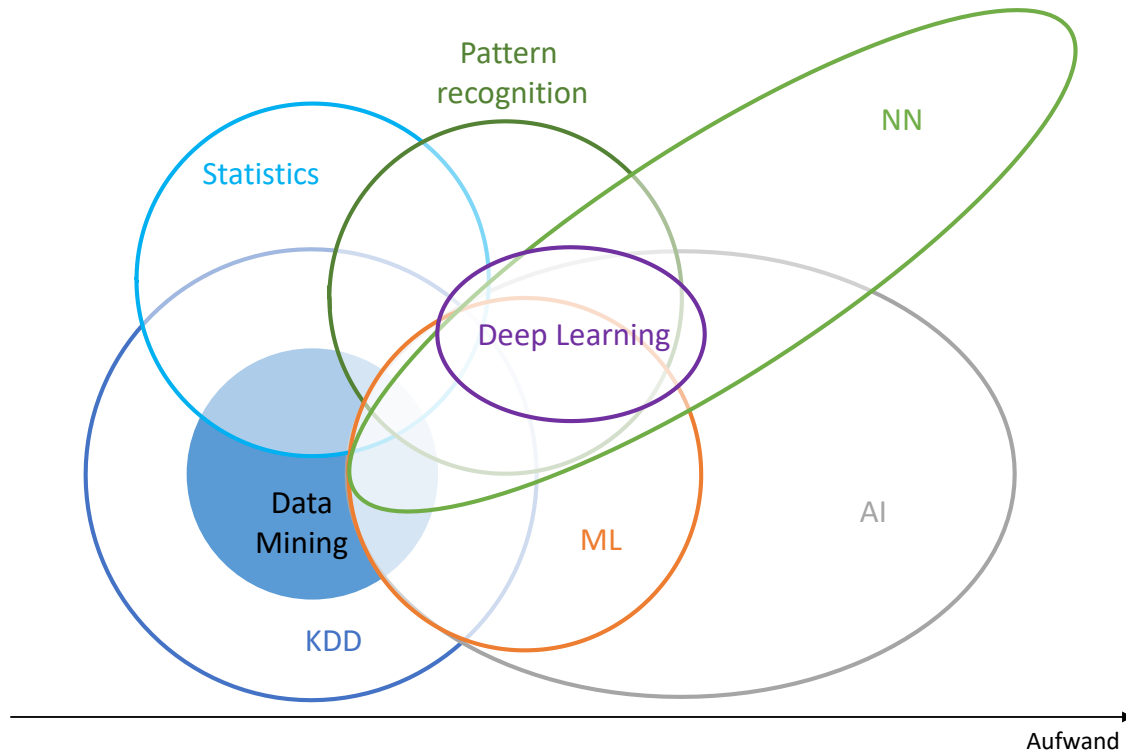


Bild 2.17: Überblick der analytischen Themenfelder in Form eines Venn-Diagrammes [73, 74]

Die Herausforderung beim Einsatz von statistischen Methoden ist daher zunächst einmal die richtige Auswahl der Methode und der Algorithmen, da eine Vielfalt an Algorithmen besteht. Doch auch die Art der Datenerhebung ist zu analysieren. Daher wird im folgenden Kapitel 3 ein Systementwurf für die voraussagende Instandhaltung erarbeitet, welcher die Auswahl und Verfügbarkeit der Daten und Anwendung von statistischen Methoden beinhaltet.

In vielen Veröffentlichungen wird immer wieder von einer **ausreichenden Datenbasis** gesprochen und auch bei der oben gezeigten Umfrage taucht diese Herausforderung auf. Die Eigenschaften einer ausreichenden Datenbasis werden jedoch nicht definiert. Eine weitergehende Recherche zu dem Thema ergibt eine mögliche Gruppierung des Themas in die Schwerpunkte Quantität, Qualität und Optimierung der Datenbasis.

In der einschlägigen Literatur zur voraussagenden Instandhaltung ist oftmals die Sprache von einem Zeitraum, in dem die Daten erfasst wurden oder von einer vorliegenden Datenmenge. Beide Beschreibungen zielen jedoch auf die Quantität, in dem Daten für eine voraussagende Instandhaltung vorliegen sollen, ab. Laut [29, 69, 70, 75 bis 77] ist es notwendig über einen längeren Zeitraum bereits Daten zu erfassen, jedoch wird die Angabe nicht weiter konkretisiert und bleibt in allen Veröffentlichungen als weiche Aussage stehen. Hierzu ein Beispiel: „Im Training verbessert sich das Modell stetig mit der Menge der zur Verfügung gestellten Daten.“ [75] Weiterhin geben die veröffentlichten Angaben zur Datenmenge in [38,

69, 71, 78 bis 80] keinen faktischen Aufschluss, welches eine ausreichende Datenmenge ist. Hierzu findet sich beispielhaft folgendes Zitat „Voraussetzung für [den Einsatz von Machine Learning-Verfahren] ist die Verfügbarkeit einer ausreichenden Datenbasis [...]“ [78]

Ein Zitat von [80] deutet bereits auf die notwendige Qualität der Daten und somit deren Informationsgehalt hin: „Voraussetzung für die Anwendung der verschiedenen Methoden des maschinellen Lernens ist das Vorhandensein von Daten in ausreichend großer Menge und in erforderlicher Güte.“ Die Qualität der Daten ist nach [70, 75, 81 bis 83] ebenso relevant, laut derer sollten in den Daten sowohl Gut- als auch Schlecht-Daten beinhalten sein, wobei Schlecht-Daten Belastungen und Ausfälle eines Produktionssystems bzw. dessen Komponente abbilden. Zusätzlich sollten laut [81] diese Schlecht-Daten mehrfach vorhanden sein, wobei dies nicht weiter präzisiert wird. Weiterhin weist [83] daraufhin, dass auch eine umfangreiche Datenbasis durch Verfälschungen belastet sein könnte. Das heißt, Komponenten könnten durch vorausbestimmte intervallbasierte Instandhaltung in ihrer Lebensdauer verlängert oder durch zustandsbasierte Instandhaltung bereits zu früh getauscht werden.

Neben der Quantität und Qualität tritt vermehrt die Optimierung der Datenbasis für eine vorausschauende Instandhaltung durch Nutzung von Daten aus vergleichbaren Systemen in den Fokus. [26, 29, 75]

Insgesamt lässt sich aus obigen, zusammengestellten Aussagen folgern, dass also sowohl in der Literatur wie auch im industriellen Umfeld noch Unklarheit besteht, wann eine ausreichend große Datenbasis aufgebaut ist. Hierzu werden herkömmlicherweise Daten mittels Versuchen oder Testdaten gekennzeichnet und die Ergebnisse validiert. Oftmals zeigt erst die Analyse der Daten ein unzureichende Datenbasis auf. Eine Verbesserung dieser wird meist mit einer weiteren, zeitaufwendigen Datenerfassung und anschließenden Datenanalyse versucht zu beheben. Ein Maß für die Vorhersagegüte in Abhängigkeit der Datenbasis ist bisher nicht gegeben.

Fazit zur Notwendigkeit der Arbeit In diesem Abschnitt wurden die Herausforderungen für die Umsetzung einer voraussagenden Instandhaltung dargestellt. Dabei konnte gezeigt werden, dass für die ersten drei Kategorien IT-Sicherheit, Datenzugriff und IT-Infrastruktur bereits Konzepte und zum Teil entsprechende Lösungen vorhanden sind. Daher wird in dieser Arbeit darauf nicht weiter eingegangen. Vielmehr liegt der Handlungsbedarf darin, eine ausreichend große Datenbasis und eine Auswahl von geeigneten statistischen Methoden, welche für das Training von Vorhersagemodellen notwendig sind, zu definieren.

2.4 Fazit

In diesem Kapitel der Arbeit wurden die grundlegenden Eigenschaften von Produktionssystemen zusammengefasst, die Vor- und Nachteile der verschiedenen gebräuchlichen Instandhaltungsstrategien analysiert und die verfügbaren Software-Werkzeuge der bekanntesten Werkzeugmaschinenhersteller und -zulieferer zusammengetragen und bewertet.

Die Zusammenfassung der Eigenschaften von Produktionssystemen wurde durch eine Definition des Produktionssystems eingeleitet. Eine Vertiefung der Kenntnisse über solche Systeme erfolgte durch die Definition der technischen Verfügbarkeit, wozu die geplante Belegungszeit T_B und die technische Ausfallzeit T_T eingeführt wurden. Im Weiteren wurden die Unterschiede von Ausfällen und die Bedeutung der Lebensdauer und deren Einflussgrößen in Produktionssystemen übersichtlich zusammengefasst. Zur Einordnung der Produktionssysteme in dieser Arbeit wurde der Bezug zur Motoren-Komponenten-Produktion mit deren Werkzeugmaschinen hergestellt. Zum weiteren Verständnis der Arbeit wurde der Aufbau einer Werkzeugmaschine erörtert und damit bereits ein erster Bezug zu Daten in einem System geschaffen.

Darüber hinaus wurde untersucht, inwiefern Ausfälle in Systemen vermieden bzw. abgestellt werden können, dazu wurden die vier Instandhaltungsstrategien – reaktiv, vorausbestimmt, zustandsorientiert und voraussagend – eingeführt. Die Unterschiede dieser vier Strategien sind: Die reaktive Instandhaltung agiert erst bei einem Ausfall, dadurch entsteht der Nachteil, dass das System auch in produktiven Zeiten ausfällt. Daher wird alternativ eine vorausbestimmte Instandhaltungsstrategie eingesetzt. Bei dieser wird eine Einheit unabhängig vom Zustand getauscht, dadurch besteht jedoch das Risiko eine Einheit zu früh zu tauschen. Dem entgegenwirken kann die Verwendung einer zustandsorientierten Instandhaltung, bei welcher der Zustand einer Einheit zum aktuellen Zeitpunkt bewertet wird. Allerdings erfordert eine Einschätzung des Zustandes und damit verbunden einer Abschätzung der Restlebensdauer sehr viel Expertise. Falls basierend auf dem Zustand eine Vorhersage durchgeführt wird, so handelt es sich dann um eine voraussagende Instandhaltungsstrategie, welche eine Abschätzung der restlichen Lebensdauer einer Einheit durchführt – umsetzbar ist dies meist nur durch ein Vorhersagemodell, welches mittels historischer Daten und Machine-Learning-Algorithmen erstellt wird. Die Vorteile der voraussagenden Instandhaltung (dargestellt in Bild 2.8) sind unter anderem die optimale Ausnutzung des Vorrates einer Einheit, die Planung von Instandhaltungstätigkeiten und die Steigerung der technischen Verfügbarkeit V_T durch Reduktion der technischen Ausfallzeit T_T .

Eine weiterführende, eigens aufgestellte Marktanalyse von Software-Werkzeugen brachte die Erkenntnis, dass es noch keine schlüsselfertigen Lösungen für Predictive Maintenance auf dem Markt gibt. Darauf aufbauend erfolgte deshalb eine Analyse der Herausforderungen beim Einsatz von Predictive Maintenance. Resultierend aus

dieser Analyse wurde der Handlungsbedarf – im Bereich der Anwendung statistischer Methoden und der Definition einer ausreichenden Datenbasis – und damit die Notwendigkeit der Arbeit aufgezeigt.

Folglich wurde durch dieses Kapitel die Arbeit in das Themengebiet der Instandhaltung mit dem Einsatz von Machine-Learning für die voraussagende Instandhaltung in der Fertigung eingeordnet.

3 Systementwurf für die Generierung von Vorhersagemodellen

Im folgenden Kapitel wird eine ausführliche Beschreibung der Datenbasis von Werkzeugmaschinen vorgenommen. Es werden Möglichkeiten aufgezeigt, wie die entsprechenden Daten mittels verschiedener Methoden des Machine Learning ausgewertet werden können und darauf basierende Vorhersagemodelle berechnet werden können.

3.1 Datenbasis für Vorhersagemodelle von Ausfällen

Bereits Abschnitt 0 hat gezeigt, dass eine der Herausforderungen von Predictive Maintenance eine ausreichend große Datenbasis ist. Bevor jedoch der notwendige Umfang an Daten diskutiert werden kann, muss erst einmal eine Definition von Daten aus einer Fertigung erfolgen. Bereits im vorherigen Abschnitt wurde in Bild 2.16 vorweggenommen, dass es sinnvoll ist, Daten aus der Feldebene und damit die Werte der Sensoren und Aktoren einer Werkzeugmaschine zu verwenden, um ein Abbild derselben zu erhalten und damit ein Vorhersagemodell berechnen zu können.

Die Daten für die Erzeugung eines solchen Vorhersagemodells lassen sich jedoch auf zwei unterschiedlichen Wegen erfassen. Herkömmlicherweise erfolgt eine Erhebung der Daten direkt an dem realen Produktionssystem. Dieser datengetriebene Ansatz zum Training eines Vorhersagemodells mit historischen Daten ist meist einfacher umzusetzen, da die erfassten Daten dem realen System entsprechen. Jedoch ist der datengetriebene Ansatz auch sehr langwierig, denn es müssen erst eine Vielzahl an unterschiedlichen Datenpunkten, einschließlich Datenpunkte von zugehörigen Ausfällen, aufgezeichnet werden, bevor das Vorhersagemodell unterschiedliche Kennzeichnungen von Ausfällen erlernen kann.

Eine weitere, in der Motoren-Komponenten-Produktion bislang aber eher unübliche Möglichkeit ist es, ein Simulationsmodell vom System zu erstellen. Dies kann zum Beispiel durch eine nahezu vollständige Modellierung des Produktionssystems unter Berücksichtigung der physikalischen Gesetze und der funktionalen Abläufe im System erfolgen. Man spricht dann auch vom Digitalen Zwilling [84]. Auf Basis eines solchen Simulationsmodell können bzw. sollten alle Arten von Ausfällen simuliert werden, sodass entsprechend geeignete Daten für das Training eines Vorhersagemodells zur Verfügung stehen.

Für die Bestimmung eines Vorhersagemodells gibt es somit zwei unterschiedliche Ansätze, dargestellt in Bild 3.1. Der eine ist der datengetriebene Ansatz – auch Black-Box-Verfahren genannt. Hierbei wird mittels aufgezeichneter, historischer Daten das Systemverhalten identifiziert. Der andere Ansatz ist modellgetrieben, auch bekannt als White-Box-Verfahren, bei dem ein physikalisches und funktionales Modell unter Verwendung von Simulationen generiert wird. [38]

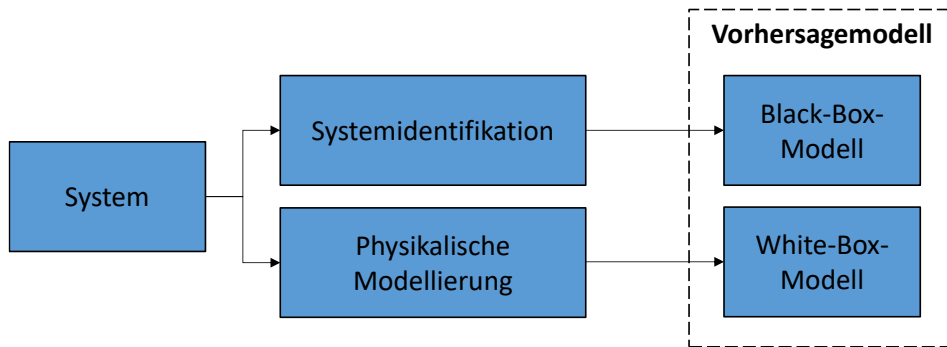


Bild 3.1: Ansätze zur Generierung eines Vorhersagemodells: Dargestellt ist zum einen die Systemidentifikation und zum anderen die physikalische Modellierung. Im ersteren Fall würde das Vorhersagemodell einem Black-Box-Modell entsprechen und im zweiten Fall würde es ein White-Box-Modell darstellen.

Die nachfolgenden zwei Abschnitte zeigen zum einen die Möglichkeiten zur Gewinnung von Daten aus realen Produktionssystemen und zum anderen die Vorgehensweise zur Erstellung einer physikalischen Modellierung. Der dritte Abschnitt zeigt die Möglichkeiten zur Kennzeichnung von Daten als potenzieller oder funktionaler Ausfall.

3.1.1 Gewinnung von historischen Daten für das Black-Box-Modell

Der Ansatz der Black-Box-Modellierung wird oftmals gewählt, wenn die konkrete Beschreibung einer Einheit oder eines Systems sehr komplex, zu zeitaufwendig oder häufig gar nicht möglich ist. Die Vorteile einer Black-Box-Modellierung sind ein geringer Aufwand bei der Systembeschreibung und ein leichtes Erlernen von Korrelationen und Abläufen in der Einheit. [85] Allerdings sind hierfür auch viele Daten notwendig und es ergeben sich aufgrund des gewählten Black-Box-Ansatzes erst einmal keine tiefergehenden Informationen über die Beschaffenheit der zugrunde liegenden physikalischen Einheit. Dadurch sind vorausgehende Verfahren der Regressionsanalyse oder der Mustererkennung für die Erstellung von Vorhersagemodellen zwingend notwendig.

Ein technisches System in der Fertigung wie in Abschnitt 2.1.5 vorgestellt hat eine speicherprogrammierbare Steuerung (SPS), welche die Steuerung der Komponenten im System ermöglicht. Die relevanten Daten liefern die Eingänge und Ausgänge bzw. Sensoren und Aktoren, welche physikalisch im System verbaut sind, sowie die zur Bereitstellung zusätzlicher Informationen über das Produktionssystem vorgesehenen Merker. [34] Eine automatisierte Werkzeugmaschine ist zusätzlich noch mit einer numerischen Steuerung¹⁷ (NC) ausgestattet, welche die Bewegungsteuerung der einzelnen Achsen übernimmt. Die relevanten Daten der numerischen Steuerung sind somit die Signale des Wegmesssystems und die Vorgabewerte der Achsantriebe. [31] Die Datenquellen SPS und NC generieren grundsätzlich Daten vom Typ sensorgenerierte Daten. [86] Basierend auf dem internen Aufbau einer

¹⁷ engl.: Numerical Control

Werkzeugmaschine werden die in Tabelle 3.1 aufgelisteten, unterschiedlichen Datenquellen erhoben. Alle Daten liegen als Zeitreihe vor. Der Zyklus der Steuerungen ist unterschiedlich. Bei einer SPS liegt der Abarbeitungszyklus beispielsweise meist zwischen 10 und 20 ms, wobei auch nicht äquidistant abgetastet werden kann bzw. Daten erfasst werden können. Bei einer NC hingegen ist ein Echtzeit-Verhalten notwendig, wodurch der Lageregelungstakt immer gleichbleibend, meist von einer Dauer von 2 ms, ist. Die erfassten Daten in der SPS sind in der Regel binäre Werte (bool) mit den Werten 0 und 1, aber auch ganzzahlige Zahlen (int) oder Gleitkommazahlen (float) wie bei einer NC sind möglich.

Tabelle 3.1: Übersicht über die verschiedenen Datenquellen einer Werkzeugmaschine: Aufgelistet sind die Eigenschaften der Datenquellen unterteilt nach Kategorie, Zyklus, Abtastung und Datentyp.

Datenquelle	Zyklus	Abtastung	Datentyp
Sensoren SPS	Abarbeitungszyklus	10-20 ms	bool, int, float
Aktoren SPS	Abarbeitungszyklus	10-20 ms	bool, int, float
Wegmesssysteme	Lageregelungstakt	2ms	float
Achsantriebe	Lageregelungstakt	2ms	float

Für die Erfassung der Steuerungsdaten sind in den letzten Jahren immer mehr Lösungen entworfen worden. Insbesondere sogenannte IoT¹⁸-Gateways haben sich dabei immer mehr auf dem Markt etabliert. IoT-Gateways ermöglichen „Dingen“ wie Steuerungen Daten zu erfassen, zwischenspeichern und an zentrale Datenbroker zu versenden. Oftmals sind ältere Steuerungen in Bestandsmaschinen verbaut, an denen entweder keine Anpassung mehr vorgenommen werden soll, um z.B. den Serienbetrieb nicht zu unterbrechen, oder an denen steuerungstechnisch keine Möglichkeit zur Erfassung der Daten besteht. In diesem Fall eignen sich zusätzliche IoT-Gateways zur Aufrüstung der bestehenden Steuerung. Einer der bekanntesten Steuerungshersteller ist die Firma Siemens. Diese bietet sogenannte Simatic bzw. Sinumerik Edge Geräte an, welche neben der Datenerfassung, -speicherung und -weitergabe auch noch die Daten direkt verarbeiten können. Allerdings ist nicht immer ein zusätzliches IoT-Gateway notwendig: Eine einfachere Variante zur Erfassung von SPS-Daten ist die Nutzung von Funktionsbausteinen, welche die Daten z.B. per MQTT¹⁹ direkt aus der SPS ohne zusätzliches Gerät (IoT-Gateway) versenden können [88].

¹⁸ engl.:Internet of Things; deutsch: Internet der Dinge

¹⁹ MQTT steht für Message Queuing Telemetry Transport und ist ein Nachrichtenprotokoll zum Versenden und Empfangen von Nachrichten [11, 87].

3.1.2 Vorgehensweise einer physikalischen Modellierung für das White-Box-Modell

Der zweite Ansatz, die sogenannte White-Box-Modellierung, nutzt mathematische Modelle, in der Regel Differentialgleichungen, mit denen versucht wird, das reale Verhalten eines physikalischen Systems zu beschreiben. Ein solches Modell ermöglicht es verschiedene Situationen zu simulieren und das Systemverhalten zu berechnen, welches dann für die Erstellung der Vorhersagemodelle genutzt werden kann [79]. Jedoch ist der Aufwand bei der Erstellung von White-Box-Modellen sehr hoch. Zum einen wird eine hohe Expertise benötigt und zum anderen ist die Erstellung sehr zeitaufwendig oder unmöglich. [85]

Teilweise werden FMU-Modelle²⁰ zu Komponenten mitgeliefert, allerdings müssen diese dann für ein System zusammengebaut werden. Dabei werden oftmals nicht alle Einflussfaktoren berücksichtigt, da diese sehr vielfältig und nur schwer vollständig zu erfassen sind, siehe Kapitel 2.1.4. Außerdem bringt das Zusammenspiel von unterschiedlichen Einheiten ein neues Systemverhalten mit sich, wodurch der Erstellungs- und Abstimmungsaufwand für ein solches Modell sehr hoch sein kann. [89]

Eine Variante zur Berücksichtigung von Einflussfaktoren kann die sogenannte Gray-Box-Modellierung sein. Bei dieser wird vorausgesetzt, dass bereits ein zumindest grobes Simulationsmodell vorliegt. Mit Hilfe von realen Daten aus dem tatsächlichen System werden dann die Modellparameter abgeschätzt. Das Verfahren der Parameterschätzung beruht hierbei auf einem iterativen Prozess, bei dem im Vorfeld ein geeignetes Gütekriterium z.B. die Summe der Fehlerquadrate²¹ und ein akzeptierter Modellfehler festgelegt werden müssen.

3.1.3 Kennzeichnung von Ausfällen

Die Ansätze der vorherigen zwei Abschnitte haben Optionen zur Erzeugung, Erfassung und Speicherung von Daten für ein Vorhersagemodell betrachtet. Für die Erstellung eines Vorhersagemodells sind zum einen Daten, welche das Systemverhalten beschreiben, notwendig. Zum anderen ist eine Kennzeichnung der Daten erforderlich, denn nur so kann ein Ausfall überhaupt erst erkannt und das Vorhersagemodell entsprechend trainiert werden. Dies ist also eine notwendige Voraussetzung, damit das Vorhersagemodell einen erneuten Ausfall vorhersagen kann.

In dieser Arbeit wird zur Unterscheidung der Kennzeichnung von Ausfällen das in Abschnitt 2.1.2 bereits eingeführte P-F-Intervall verwendet. Das heißt, ein potenzieller Ausfall (P), welcher noch keinen sichtbaren Ausfall hervorruft, muss bereits

²⁰ Functional Mock-up Unit beschreibt die Funktion einer Komponente mittels einer definierten Schnittstelle. Eine solche modellhafte Beschreibung lässt sich in verschiedene Simulationssoftware implementieren.

²¹ engl. least-squares method

erkannt werden, bevor ein funktionaler Ausfall (F) eintritt. Nur wenn eine Kennzeichnung der potenziellen Ausfälle erfolgt, ist eine Vorhersage der funktionalen Ausfälle möglich. Umso kleiner das P-F-Intervall ist, umso kürzer ist die Reaktionszeit für den Instandhalter. Daher ist es umso wichtiger, den potenziellen Ausfall so früh wie möglich zu erkennen und darauf reagieren zu können. Es gibt Komponenten, bei denen der Zeitraum des potenziellen Ausfalls sehr gering oder nicht vorhanden ist. Dies sind vor allem elektrische Einheiten.

Ein ausreichendes P-F-Intervall definiert sich abhängig vom Ausfalltyp einer Komponente. Folgt aus dem Primärausfall ein Sekundärausfall mit hohen Kosten und/oder einer erhöhten Ausfallzeit, ist ein P-F-Intervall von bereits wenigen Minuten relevant. Steht eine zeitlich flexible Instandhaltung zur Verfügung, ist auch bei einem Primärausfall ohne Sekundärausfall eine Vorhersage mit Vorlauf von einigen Stunden hilfreich, da so eine Reparatur innerhalb einer nicht-produktiven Zeit eingeplant werden kann. In der Regel erfolgt eine Instandhaltungsplanung allerdings mindestens täglich, sodass der Vorhersagehorizont wenigstens 24 h betragen sollte. Wenn der Einsatz von Vorhersagemodellen auch zu einer Reduktion der Ersatzteilbevorratung führen soll, müssen bei der Betrachtung des anzustrebenden Vorhersagezeitraums noch Lieferzeiten entsprechend berücksichtigt werden.

Die **Kennzeichnung von potenziellen Ausfällen** ist komplex und zum Teil nicht immer möglich. Die Kennzeichnung der potenziellen Ausfälle kann dabei sowohl manuell als auch automatisch erfolgen und ist abhängig vom Anwendungsfall.

Im Falle der in dieser Arbeit betrachteten Predictive Maintenance ist die manuelle Kennzeichnung von potenziellen Ausfällen sehr aufwendig, da meistens potenzielle Ausfälle in einem System von außen nicht einfach erkennbar sind. Eine Prüfung der einzelnen Komponenten im Rahmen einer Inspektion zeigt nur physische Veränderungen zu einem Zeitpunkt. Sollten sich beispielsweise die Temperatur eines Antriebs im Betrieb verändert haben, ist dies während der Inspektion nicht unbedingt erkennbar. Im Vergleich hierzu kann bei Predictive Quality²² recht einfach eine manuelle Kennzeichnung von Daten durchgeführt werden, da dort quantitative Maße der Produkte (i.O. oder n.i.O.) zur Beurteilung der Klasse des Produkts herangezogen werden können.

Eine komplett-automatisierte Kennzeichnung von potenziellen Ausfällen in historischen Daten ist jedoch sehr schwierig, da es in der Regel keine eindeutigen und allgemeingültigen Indizien für einen potenziellen Ausfall gibt. Lediglich die Abweichung von Zeitreihen oder ein Trend in den Daten können Hinweise auf einen potenziellen Ausfall geben. Allerdings ist es sehr häufig schwierig bis unmöglich aus Indizien ohne Expertenwissen eine Ursache abzuleiten und eine fehlerhafte bzw. nichtzutreffende Kennzeichnung von potenziellen Ausfällen vollständig auszuschließen. Daher ist es sinnvoll, einen kombinierten Ansatz aus automatischer und manueller Kennzeichnung von Daten zu wählen. Das heißt, die erhobenen Daten aus

²² Predictive Quality ist die Vorhersage der Qualität durch Beobachtung der Prozesse eines Produktionssystems.

einem realen Produktionssystemen werden automatisch analysiert. Aufgezeigte Abweichungen oder Trends sollten anschließend durch Experten untersucht und zu einem vorliegenden potenziellen Ausfall zugeordnet werden.

Bei Daten aus einer Simulation ist eine Kennzeichnung ebenso aufwendig, denn es müssten alle Anzeichen für einen potenziellen Ausfall im Simulationsmodell abgebildet werden. Dies ist aufgrund der Komplexität eines Systems meist nur mit hohem Aufwand möglich.

Diese Arbeit konzentriert sich deshalb weiterführend auf den datenbasierten Ansatz, da eine Modellierung aller Systeme und deren untergeordneter Einheiten zu aufwendig wäre. Außerdem entspricht der datengetriebene Ansatz dem Stand der Technik und wird gängigerweise in der Motoren-Komponenten-Produktion eingesetzt.

Aufgrund der Fokussierung dieser Arbeit auf den datengetriebenen Ansatz – der Black-Box-Modellierung – wird nachfolgend die Vorgehensweise zur Erkennung von Abweichungen in Daten beschrieben. Durch den Einsatz von statistischen Methoden können Abweichungen in den Daten analysiert werden. Da die Daten in diesem Fall zunächst keine Kennzeichnungen (Labels) besitzen, wird ein unüberwachtes Lernen²³ gewählt. Einer der bekanntesten Methoden dieser Kategorie ist das Clustering.

Das Clustering unterscheidet sich je nach Algorithmus in seiner Vorgehensweise. Die bekanntesten Clustering-Verfahren (siehe Bild 3.2) sind das partitionierende, das hierarchische und das dichtebasierte Clustering. Ersteres gruppiert ähnliche Datenpunkte zusammen, wobei eine feste Anzahl k an Gruppen von Datenpunkten im Vorfeld definiert werden muss. Ein Vertreter dieser Clustering-Verfahren ist das k-Means-Clustering. Das zweite Clustering-Verfahren hingegen basiert auf der Distanz der Punkte im n -dimensionalen Raum wie z.B. der euklidischen Distanz. Dabei unterscheidet das sogenannte hierarchische Clustering zwei Sub-Verfahren, zum einen das divisive und zum anderen das agglomerative Clustering. Das divisive Clustering-Verfahren geht von einer großen Gruppe von allen Datenpunkten aus und spaltet diese dann immer weiter auf. Das agglomerative Clustering-Verfahren hingegen sieht jeden Datenpunkt zunächst als einzelne Gruppe an und fügt diese, abhängig von der Distanz der Punkte bzw. der Clusterzentren, dann zusammen. Unabhängig von dem Verfahren (divisiv oder agglomerativ) muss beim hierarchischen Clustering eine Entscheidung getroffen werden, wann eine Gruppe gebildet werden soll. Hierzu werden meist Dendrogramme verwendet, da sich damit sehr gut die Distanzen zwischen den Gruppen darstellen lassen. Neben dem partitionierenden und hierarchischen Clustering gibt es noch das dichtebasierte Clustering. Bei diesem Verfahren wird z.B. vom letzten Datenpunkt der nächste Nachbar gesucht, wodurch Ketten gebildet werden können. [7, 90]

²³ engl.: unsupervised Learning

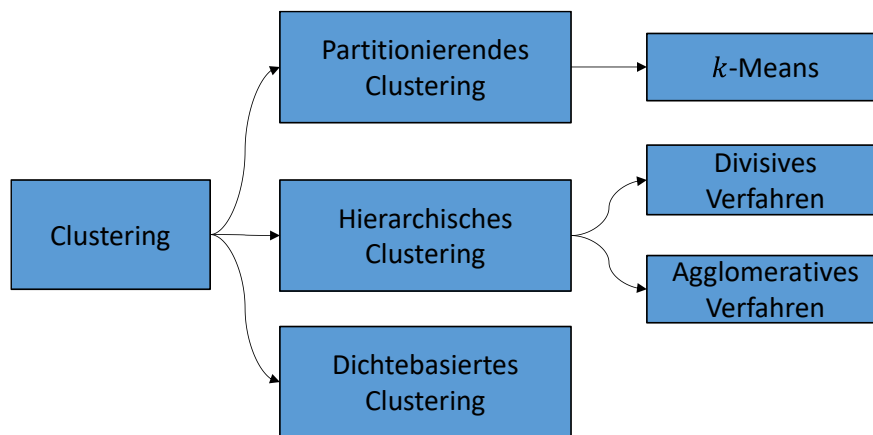


Bild 3.2: Übersicht über verschiedene Clustering-Verfahren: Dargestellt sind das partitionierende, hierarchische und dichtebasierte Clustering mit den wesentlichen Vertretern.

Zusammenfassend kann festgehalten werden, dass mit Hilfe des Clusterings ähnliche Gruppen festgestellt und damit Abweichungen erkannt werden können. Dadurch können häufig dann auch potenzielle Ausfälle identifiziert und durch Experten eindeutig bewertet werden.

Im Unterschied zu den potenziellen Ausfällen ist die **Kennzeichnung von funktionalen Ausfällen** in der Regel leicht umsetzbar, da bei diesen das System ausfällt und somit dieser Ausfall nicht mehr übersehbar ist. Es kann daher eine manuelle, automatisierte oder regelbasierte Kennzeichnung der Daten gut umgesetzt werden. Sollten die Daten aus einer Simulation stammen, müssen zusätzlich feste Grenzwerte vorgegeben werden, bei denen das System ausfällt.

3.2 Vorhersagemodelle durch Klassifikation

Die Erstellung von Vorhersagemodellen ist basierend auf gekennzeichneten Daten relativ einfach umsetzbar: Mittels den im vorhergehenden Abschnitt beschriebenen Methoden – Clustering und Expertenwissen – ist es möglich für potenzielle Ausfälle gekennzeichnete Daten zu erhalten. Durch Algorithmen des überwachten Lernens²⁴ kann basierend auf den gekennzeichneten Daten ein Klassifikationsmodell, welches Voraussetzung für eine Vorhersage ist, erlernt werden.

Im nachfolgenden Abschnitt werden zuerst Algorithmen zur Klassifikation eingeführt und anschließend wird eine Methode zur Validierung des Klassifikationsmodells vorgestellt. Im daran anschließenden Abschnitt wird analysiert, inwiefern Daten aus anderen Systemen zum Training von Modellen eingesetzt werden können.

²⁴ engl.: supervised Learning

3.2.1 Vorgehensweise zur Erstellung eines Klassifikationsmodells

Eine Klassifikation beschreibt die Zuordnung von Datenpunkten zu Klassen. Im Falle dieser Arbeit sind die Klassen – kein Ausfall oder potenzieller bzw. funktionaler Ausfall. Es gibt eine Vielzahl an Algorithmen, um die Datenpunkte einer Klasse zuzuordnen. Abhängig von der Beschaffenheit der Daten und der Tauglichkeit des gewählten Algorithmus können die Datenpunkte besser oder schlechter der korrekten Klasse zugeordnet werden.

Zur Erstellung des Klassifikationsmodells wird zunächst aus einer Menge an Daten eine Teilmenge, häufig 80 % der Daten, zum Training des Klassifikationsmodells gebildet [91]. Die bekanntesten Algorithmen für die Klassifikation sind Diskriminanzanalysen²⁵, Support Vector Maschinen (SVM), Naive Bayes-Klassifikatoren, k-Nearest-Neighbor-Algorithmen (kNN), Entscheidungsbäume²⁶ und neuronale Netze. [92]

Anschließend wird mit den restlichen Daten (häufig 20 %) eine Validierung des Klassifikationsmodells durchgeführt. Dazu werden die Datenpunkte, für die die Kennzeichnung bekannt ist, mittels des trainierten Klassifikators erneut gekennzeichnet (die Klasse wird vorhergesagt). Ein Vergleich zwischen der wahr richtigen Kennzeichnung und der vorhergesagten Kennzeichnung ermöglicht dann eine Bewertung des Klassifikators in Form einer Konfusionsmatrix – auch Wahrheitsmatrix genannt. In Tabelle 3.2 wird der Aufbau einer Konfusionsmatrix für eine binäre Klassifikation (Zuordnung in zwei Klassen) demonstriert. Wenn ein Datum richtig zu einer positiven Klasse zugeordnet wurde, wird dies als True-Positiv (TP) bezeichnet. Wenn ein positives Datum fälschlicher Weise zu einer negativen Klasse zugeordnet wurde, wird dies als False-Negativ (FN) gezählt. Für Daten, welche in Wahrheit negativ sind, gilt: Sollte das Datum als positiv vorhergesagt worden sein, so wird es als False-Positiv (FP) angesehen und sollte es korrekterweise als negativ vorhergesagt worden sein, so würde es einem True-Negativ (TN) entsprechen. Zusätzliche Informationen liefern die Summen der wahren positiven Daten $P = TP + FN$, der wahren negativen Daten $N = FP + TN$, der vorhergesagten positiven Daten $P^* = TP + FP$ und die der vorhergesagten negativen Daten $N^* = FN + TN$.

In dieser Arbeit geht es, wie bereits gesagt, um eine binäre Klassifikation, welche durch die Klasse „Ausfall“ und „kein Ausfall“ beschrieben wird. Übertragen für diese Arbeit bedeutet dies, dass True-Positiv TP die Gesamtanzahl der richtig vorhergesagten Ausfälle beschreibt. True-Negativ TN hingegen beschreibt die Gesamtanzahl der richtig vorhergesagten Nicht-Ausfälle.

²⁵ engl.: Discriminant Analysis

²⁶ engl.: Decision Tree

Tabelle 3.2: Aufbau einer Konfusionsmatrix einer binären Klassifikation

	Als Positiv (+) vorhergesagt	Als Negativ (-) vorhergesagt	Summe
In Wahrheit positiv +	True-Positiv (<i>TP</i>)	False-Negativ (<i>FN</i>)	<i>P</i>
In Wahrheit negativ -	False-Positiv (<i>FP</i>)	True-Negativ (<i>TN</i>)	<i>N</i>
Summe	<i>P</i> *	<i>N</i> *	<i>N + P</i>

Basierend auf den absoluten Gütemaßen können die in Tabelle 3.3 aufgelisteten relativen Gütemaße zur Beurteilung der Leistung eines Klassifikators berechnet werden [93]. Die Güte von Klassifikationsmodellen ist zum einen abhängig von der Vollständigkeit, Genauigkeit und Auswahl der erfassten Daten eines Systems [94], zum anderen entscheidet aber auch der gewählte Algorithmus in Zusammenspiel mit den Eigenschaften der Daten über die Güte der Klassifikationsergebnisse.

Eine Problematik besteht jedoch weiterhin darin, dass der Klassifikator gegebenenfalls einen zukünftigen Datenpunkt bzw. ein Datum nicht korrekt klassifizieren kann, sofern dieser ein vom bisherigen Vorhersagemodell stark abweichenden Wertezustand aufweist. Eine Konfusionsmatrix kann immer nur die Güte der bekannten Datenmenge bewerten. Die nachfolgend entwickelte Methode wird dazu dienen, die theoretisch bestmöglichen Gütemaße zu berechnen und damit einen Zielwert für die Leistungsbeurteilung aufzustellen.

Tabelle 3.3: Auflistung der Gütemaße für Klassifikationsergebnisse und deren Definition sowie die dafür bekannten Synonyme

Name	Definition	Synonyme
<i>TPR</i>	$= TP/P$	Richtig-Positiv-Rate (engl. True-Positive-Rate), Power, Sensitivität, 1-Betafehler, Recall, Trefferquote, Empfindlichkeit
<i>FNR</i>	$= FN/P$	Falsch-Negativ-Rate (engl. False-Negative-Rate), Fehlender Alarm, Betafehler
<i>FPR</i>	$= FP/N$	Falsch-Positiv-Rate (engl. False-Positiv-Rate), Alphafehler, Typ-1-Fehler, 1-Spezifität, Fehlalarm
<i>TNR</i>	$= TN/N$	Richtig-Negativ-Rate (engl. True-Negative-Rate), Spezifität, 1-Alphafehler
Richtigkeit	$= TP + TN/(N + P)$	Korrektklassifikationsrate, Gesamtgenauigkeit, Vertrauenswahrscheinlichkeit
Falschheit	$= FP + FN/(N + P)$	Falschklassifikationsrate
<i>PPV</i>	$= TP/P^*$	Positiver Vorhersagewert (engl. Positive-Predictive-Value), Präzision, Relevanz, Wirksamkeit, Genauigkeit
<i>NPV</i>	$= TN/N^*$	Negativer Vorhersagewert (engl. Negative-Predictive-Value), Segreganz, Trennfähigkeit

3.2.2 Übertragung eines Klassifikationsmodells

Nicht immer ist es sinnvoll, ein Klassifikationsmodell selbst zu trainieren. Insbesondere Lieferanten von Komponenten oder Produktionssystemen haben oftmals die Möglichkeit auf eine große Anzahl von Daten aus dem Feld zuzugreifen. Diese können dann die deutlich größere Menge an Daten nutzen, um ein wesentlich besseres Black-Box-Modell zu trainieren als durch eine Teilmenge an Daten jemals zu erreichen wäre. Denn durch die größere Menge an und Varianz der Daten, welche naturgemäß bei einer größeren Anzahl an betrachteten Systemen auch mit einer wesentlichen höheren Geschwindigkeit entstehen, kann eine deutlich größere Vielfalt an Zuständen aus den Daten erlernt werden. Diese wiederum können in einem entsprechenden Klassifikationsmodell abgebildet werden. Jedoch sind nicht alle Systeme oder Komponenten identisch, äußere Einflussfaktoren oder veränderte Betriebsanforderungen führen dazu, dass sich ein Klassifikationsmodell unter Umständen nicht eins zu eins übertragen lässt. Es sollte nicht unerwähnt bleiben, dass die Varianz zwischen den Systemen oder Komponenten und damit innerhalb der Daten dabei so weit gehen kann, dass eine Übertragung eines Klassifikationsmodells zwischen zwei ähnlichen oder gar baugleichen Systemen nicht oder nur mit großem Aufwand möglich ist. Im später betrachteten Szenario 3 wird aber vom Regelfall des Transfer Lernens ausgegangen, dass die Übertragung des entsprechenden Klassifikationsmodells zwischen zwei baugleichen Systemen ohne weiteres möglich ist.

Dieses sogenannte Transfer Lernen ermöglicht es, bereits vortrainierte Modelle auf ähnliche Daten anzuwenden. Im Vorfeld ist es erforderlich, mit einer großen Menge an klassifizierten Daten einen geeigneten Klassifikator – meistens wird ein neuronales Netz verwendet – zu trainieren. Die gelernten Muster können anschließend auf neue Daten übertragen werden. Im Rahmen des Transfers kann das Modell dann auch für die neuen Daten geeignet trainiert werden, wobei das Erfordernis eines solchen Nachtrainings²⁷ von der Ähnlichkeit und Größe der zwei Datensätze abhängt. Der Vorteil von Transfer Lernen ist das Modelle übernommen werden können. Der Nachteil liegt hierbei jedoch im Training, denn zu Beginn werden viele klassifizierte Daten benötigt. [95, 96]

²⁷ Als Nachtrainieren wird die Anpassung des Klassifikators auf den vorliegenden Arbeitspunkt gemeint. So bleiben die grundsätzlichen Zusammenhänge gleich, allerdings verändert sich der Wertebereich.

3.3 Fazit

In diesem Kapitel wurden zunächst zwei Ansätze zur Erstellung eines Vorhersagemodells – Black-Box-Modell und White-Box-Modell – vorgestellt. Dabei stand im Fokus die Vorgehensweise zur Gewinnung von Daten aus Werkzeugmaschinen. Im Weiteren wurde demonstriert, wie eine Kennzeichnung der für die voraussagende Instandhaltung genutzten Daten erfolgen kann.

Basierend auf den gekennzeichneten Daten wurden dann unterschiedliche, gängige Algorithmen zur Modellierung eines Vorhersagemodells für Produktionssysteme kurz erläutert. Der Fokus wurde dabei auf die Leistungsbewertung des Klassifikators gelegt. Die eingeführten Gütemaße dienen der weiteren Arbeit als Grundlage.

Für die Herstellung eines Bezugs zur Praxis wurde zusätzlich noch das Transfer-Learning eingeführt. Insbesondere in der Zusammenarbeit zwischen Lieferanten und Kunden wird dieser Anwendungsfall oft diskutiert, aber auch bei baugleichen Systemen in der eigenen Fertigung kann dieser Ansatz Anwendung finden. Aus diesem Grund wird dieser Ansatz im späteren Kapitel 5.4 auch nochmals näher untersucht.

4 Methodik für die Bestimmung der Vorhersagegüte

Dieses Kapitel beschreibt die Methodik zur Bestimmung der Vorhersagegüte für verschiedene Szenarien²⁸. Zunächst werden die Beweggründe für die Entwicklung der Methode dargelegt. Danach wird die Methode zur Bestimmung der Vorhersagegüte zusammen mit den entsprechenden Gütemaßen hergeleitet. Abschließend werden Möglichkeiten zur Modellierung der Gütemaße erarbeitet und beschrieben.

4.1 Einführung in die Methodik und Szenarien

Die vorhergehenden Kapitel haben insgesamt sehr gut veranschaulicht, dass die in Abschnitt 0 aufgezeigten Herausforderungen – IT-Sicherheit, Auswahl und Verfügbarkeit der Daten, IT-Infrastruktur und Anwendung statistischer Methoden – beherrscht sind. Doch trotzdem ist das Angebot an marktreifen Lösungen für eine voraussagende Instandhaltung dürftig. Dies liegt sicherlich oftmals an den ungenügend erforschten und bisher nur subjektiv abgeschätzten Potenzialen, aber auch an der Unkenntnis über die ausreichend große Datenbasis, welche zum Erlernen von Ausfällen notwendig ist.

Eine umfassende Recherche des bekannten Wissens über die notwendige Datenbasis für die voraussagende Instandhaltung in Abschnitt 0 konnte keine konkrete Antwort geben. Lediglich wurde deutlich, dass das heutige Vorgehen darauf beruht, Daten zu erheben und hierauf verschiedene Algorithmen (siehe Bild 2.17) anzuwenden. Mittels eines Kreuzvalidierungsverfahrens wird dann die vorliegende Datenmenge aufgeteilt und die Leistung des Vorhersagemodells überprüft. Das heißt, erst einmal müssen mit mehr oder weniger großen Aufwand Daten erhoben werden und hinterher kann bewertet werden, ob ein Produktionssystem bzw. dessen Daten für die voraussagende Instandhaltung ausreichend sind. Sollte dies nicht der Fall sein, so werden weitere Daten erhoben und der Prozess wiederholt.

In diesem Kapitel wird eine Methode entwickelt, welche die Güte von Vorhersagemodellen vorhersagt, ohne Zeitreihen-Daten aus einem Produktionssystem (beschrieben in Abschnitt 3.1.1) zu nutzen. Die Methode kann somit keine Vorhersage der real auftretenden Ausfälle für ein Produktionssystem machen, aber sie kann eine obere Schranke für die erreichbare Vorhersagegenauigkeit bestimmen. Die Methode bietet also die Möglichkeit bereits vor Beginn der Datenerfassung den Informationsgehalt von Daten und daraus resultierend die optimale Leistung des Vorhersagemodells zu bestimmen.

Darauf basierend hat der Data Scientist dann auch nach der Entwicklung eines Vorhersagemodells die Möglichkeit die Leistungsfähigkeit seines Klassifikators zu überprüfen. Doch wesentlich entscheidender ist es, dass der/die Ingenieur(in), Data Engineer, Data Scientist etc. nun im Voraus ermitteln kann, welche Datenmenge zu

²⁸ Ein Szenario beschreibt den hypothetischen Verlauf von Ereignissen [18].

welchem bestmöglichen Klassifikationsergebnis führen kann. Dadurch ist es möglich objektiv eine Dauer für die Datenerfassung zu benennen und damit zu entscheiden, wann eine Datenbasis ausreichend ist. Obendrein kann dem/der Anwender(in) mitgeteilt werden, von welcher maximal erreichbaren Leistung eines Vorhersagemodells er/sie auszugehen hat²⁹.

4.1.1 Festlegung eines idealen Klassifikators für die Methodik

Die selbstentwickelte und nachfolgend vorgestellte Methode zur Bestimmung der Vorhersagegüte funktioniert unabhängig vom angewandten Klassifikations-Algorithmus. Hierdurch kann die Fehleranfälligkeit durch die Wahl eines spezifischen Klassifikations-Algorithmus vernachlässigt werden.

Prinzipiell wäre es möglich die Einflüsse, welche durch die Wahl des Klassifikations-Algorithmus entstehen, miteinzubeziehen. Allerdings soll mittels der Methodik das bestmögliche Ergebnis für die Güte eines Vorhersagemodells unabhängig vom jeweils gewählten Algorithmus bestimmt werden. Würde der jeweilige Algorithmus hingegen miteinbezogen werden, so wäre dies nicht möglich, da einfachere Algorithmen in der Regel auch im besten Fall nicht die gleiche Vorhersagekraft wie komplexere Algorithmen erreichen können.

In dieser Arbeit wird deshalb von einem idealen Klassifikator ausgegangen. Dieser ist durch folgende Annahmen charakterisiert:

- Für das Training eines Klassifikators muss ein Ausfall eine gewisse Anzahl aufgetreten sein. Sobald ein ausreichender Ausfall-Datensatz vorliegt, wird der Klassifikator neutrainiert. Damit kann der Zeitpunkt des Trainings in dieser Arbeit vernachlässigt werden und der Klassifikator ist fortlaufend aktuell.
- Jedes Datum ist eindeutig identifizierbar und kann somit zweifellos einer Klasse zugeordnet werden. Somit ist eine korrekte Gruppierung der Daten in die Klassen True-Positiv TP , False-Negativ FN , False-Positiv FP und True-Negativ TN gegeben.

Aus den beiden getroffenen Annahmen kann abgeleitet werden, dass kein Fehlalarm FP bei einem idealen Klassifikator auftreten wird. Denn entweder entspricht ein Datum einem Ausfall und gehört zur Klasse True-Positiv TP (auch Treffer genannt) oder es ist kein Ausfall und es kann somit zur Klasse True-Negativ TN zugeordnet werden. Diese Annahme ist berechtigt, da die theoretisch maximal erreichbare Vorhersagegüte bestimmt werden soll. Allerdings kann ein Ausfall auch

²⁹ Persönliche Erfahrungen haben gezeigt, dass die Erwartungen an die voraussagende Instandhaltung sehr hoch sind. Sobald ein Ausfall nicht vorhergesagt wird und ein ungeplanter Produktionsausfall entsteht, wird das System weniger akzeptiert bzw. zurück zur intervallbasierten Instandhaltung gewechselt. Interviews haben gezeigt, dass die Erwartungshaltung relevant für die Akzeptanz des Systems ist. D.h. wenn im Vorfeld deutlich wird, dass nur einige Ausfälle vorhergesagt werden können und weiterhin Produktionsausfälle durch fehlende Alarme entstehen, dann kann dies entsprechend eingeplant werden.

beim erstmaligen Eintreffen in die Klasse False-Negativ FN und damit als sogenannter Nicht-Alarm eingeordnet werden.

Eine weitere Annahme ist die Diagnostizierbarkeit eines Ausfalls. Bereits in Kapitel 2.1.2 wurde der potenzielle Ausfall eingeführt. Dieser beschreibt, dass die Komponente bereits vor dem funktionalen Ausfall Symptome aufweist, welche einen möglichen Ausfall andeuten. Sicherlich gibt es unterschiedliche Anzeichen für einen potenziellen Ausfall und damit auch Ursachen für einen funktionalen Ausfall. Jedoch ist dieses gesondert zu untersuchen bzw. kann entsprechend dadurch berücksichtigt werden, dass für eine Komponente mehrere potenzielle Ausfälle mit unterschiedlichen Symptomen und Lebensdauern vorgesehen werden. In dieser Arbeit wird jedenfalls davon ausgegangen, dass das Symptom für einen bevorstehenden funktionalen Ausfall immer gleich ist. Somit wird von einem einfachen Ausfall bzw. einem Ausfall mit immer gleichen Anzeichen ausgegangen, welcher eindeutig klassifizierbar ist³⁰.

Die nachstehende, entwickelte Methode zeigt also die bestmöglich erreichbare Vorhersagegenauigkeit unter Annahme eines idealisierten³¹ Klassifikators bei der Verwendung von historischen vorhersagbaren Daten auf. Entgegen bekannten Veröffentlichungen zur Durchführung oder Optimierung von Klassifikations-Algorithmien stellt die vorliegende Arbeit nicht die Güte eines bestimmten Klassifikators dar, sondern beschreibt allgemeingültig und unabhängig vom gewählten Klassifikator die bestmöglichen Gütemaße für die Vorhersage von Ausfällen.

4.1.2 Formulierung unterschiedlicher Szenarien zur Anwendung der Methodik

In der Motoren-Komponenten-Produktion, aber auch in diversen weiteren Expertengesprächen wurde deutlich, dass die voraussagende Instandhaltung in unterschiedlichen Konstellationen eingesetzt werden soll. Dieser Abschnitt stellt die berücksichtigten Szenarien für die Entwicklung der Methode zunächst vor. Im späteren Verlauf der Arbeit wird die Vorhersagegüte für diese bestimmt.

Experteninterviews haben weiterhin gezeigt, dass der Einsatz von voraussagender Instandhaltung zu verschiedenen Zeitpunkten der Nutzungsphase erfolgt. Beginnt die Datenerfassung mit Beginn der Nutzungsphase des Betriebs, wird künftig vom

³⁰ Neben diesen Annahmen, welche für die entwickelte Methode relevant sind, ist festzuhalten, dass das P-F-Intervall, also der Zeitraum zwischen potenziellen und funktionalen Ausfall, immer ausreichend groß sein sollte. Denn wenn sich ein Ausfall ankündigt, muss die Instandhaltung noch rechtzeitig reagieren können. Das bedeutet, Ersatzteile müssen beschafft, Instandhalter verfügbar sein und das System auch noch rechtzeitig erreicht werden. Eine Vorhersage fünf Minuten vor einem Ausfall wäre beispielsweise nur noch für das Abschalten der Anlage zur Vermeidung von Folgeausfällen hilfreich.

³¹ Die Verwendung eines nicht-idealisierten Klassifikators führt zu einer Abnahme der im Bestfall erreichbaren Vorhersagegenauigkeit, da es sich hierbei um die Bestimmung einer oberen Schranke für die Güte von Vorhersagmodellen handelt und nur im Idealfall die maximal mögliche obere Schranke erreicht wird. Dementsprechend führt die Verwendung eines idealisierten Klassifikator zu Ergebnissen, welche einen realen Klassifikatoren inkludieren.

ersten Szenario gesprochen. Dies ist ein grundlegendes Szenario, welches den klassischen Fall der Anwendung von voraussagender Instandhaltung in der Produktion beschreibt. Doch dieser Fall tritt nicht immer auf, denn oftmals sollen auch noch bereits bestehende Produktionssysteme mit voraussagender Instandhaltung ausgestattet werden. Die Auswirkungen eines zeitlich später stattfindender Einsatz von voraussagender Instandhaltung auf die Vorhersagegüte sind bisher unerforscht und sollen im Rahmen dieser Arbeit in einem zweiten Szenario untersucht werden. Dazu wird davon ausgegangen, dass die notwendige Datenerfassung erst eine gewisse Zeitspanne nach Betriebsbeginn installiert wird.

Eine weitere Möglichkeit zur Verbesserung der Vorhersagegüte durch übertragendes Lernen wurde bereits in Abschnitt 3.2.2 thematisiert. Hierbei werden Daten aus einem System genutzt und auf weitere baugleiche Systeme übertragen. Wichtig ist, dass die Anzahl und Art der Features³² in all diesen Systemen gleich sein müssen. In einem dritten Szenario werden somit die Vorteile durch die Nutzung von Daten aus baugleichen Systemen beleuchtet.

In dieser Arbeit sollen dementsprechend die drei verbreitetsten Szenarien untersucht und verglichen werden.

4.2 Herleitung der Methodik zur Bestimmung der Vorhersagegüte

Im Folgenden soll eine Methodik für die Bestimmung der (best-)möglichen Vorhersagegüte unter der Annahme einer idealen Klassifikation entwickelt werden, welche auf den Ausfällen von Einheiten in einem technischen System basiert.

Zum Erhalt der Ausfallzeitpunkte einer Einheit ist es zunächst notwendig, deren Lebensdauer zu kennen oder andernfalls zu bestimmen. Die entsprechende Einheit wird im Folgenden als Komponente K bezeichnet und sie hat eine bestimmte Lebensdauer L_K . Diese wird über die mittlere Lebensdauer ($MTTF$ oder $MTBF$) beschrieben, welche bereits in Kapitel 2.1.3 ausführlich erklärt wurde. Die Angabe der Lebensdauer erfolgt immer in Stunden. Eine Verallgemeinerung dieser Angabe auf die Einheit „beliebige Zeiteinheiten (ZE)“ ist weder sinnvoll noch notwendig, da nach VG-Norm 95287 die generelle Einheit von Komponenten-Lebensdauern Stunden ist [97].

In dieser Arbeit werden nur Komponenten K betrachtet, deren Lebensdauer L kleiner einer Betriebsdauer B sind. Die Betriebsdauer B beschreibt die Dauer des technischen Systems in der Nutzungsphase des Betriebs, welches eine rein zeitliche Betrachtung ist.

³² Als Feature wird ein Merkmal eines Merkmalsträgers, in diesem Fall eines Systems, im maschinellen Lernen beschrieben. Die Art eines Merkmals wird zum einen durch die Ausprägungen und zum anderen durch das Skalenniveau (nominal, ordinal, metrisch) definiert.

$$(MTTFF = MTBF = L_K) < B \quad (4.1)$$

Aus der Lebensdauer einer Komponente L_K lassen sich die Ausfallzeitpunkte der Komponente a_K über den Betriebszeitraum t_B bestimmen. n drückt dabei die Anzahl des Ausfalls seit Betriebsbeginn aus.

$$a_K = \{a_{K1}, \dots, a_{Kn}, a_{Kn+1}\}, \text{ wobei } a_{Kn} = L_K * n \quad (4.2)$$

Die Herangehensweise zur Berechnung der Ausfälle a ist nur korrekt, wenn die Eintrittswahrscheinlichkeit eines Ausfalls mit einer einfachen Delta-Funktion beschrieben wird. Normalerweise folgt die Wahrscheinlichkeit für das Eintreten eines Ausfalls einer bestimmten Verteilung. Allerdings hat jede Komponente ihre eigene Verteilung und auch wenn diese meist Weibull-verteilt sind, so unterscheiden sie sich in Lage und Streuung [98]. Es ist festzuhalten, dass die individuelle Verteilung zu jeder Lebensdauer einer Komponente weitgehend unbekannt ist. In dieser Arbeit wird daher auf eine komplexe Verteilung für die Ausfallwahrscheinlichkeit verzichtet und eine Delta-Verteilung gewählt. Diese ist auch als Dirac-, Impuls-, Stoß-, Nadel- oder Einheitsimpuls-Verteilung bekannt. Die Übertragungsfunktion ist sehr simpel und in Formel 4.3 beschrieben. [99]

Interpretiert man die Formel 4.3 bedeutet dies, dass sobald ein Ausfall eintreten sollte, tritt dieser auch ein. In Bild 4.1 ist der Verlauf der Sprungfunktion und damit die Verteilung der Ausfallwahrscheinlichkeit F über der Zeit t dargestellt. Deutlich erkennbar ist, dass die Eintrittswahrscheinlichkeit für einen Ausfall erst einmal bei null liegt, sobald jedoch die mittlere Lebensdauer $MTBF$ bzw. $MTTFF$ erreicht wird, fällt die Komponente sofort aus und die Ausfallwahrscheinlichkeit liegt bei eins bzw. 100 %.

$$G(s) = s \quad (4.3)$$

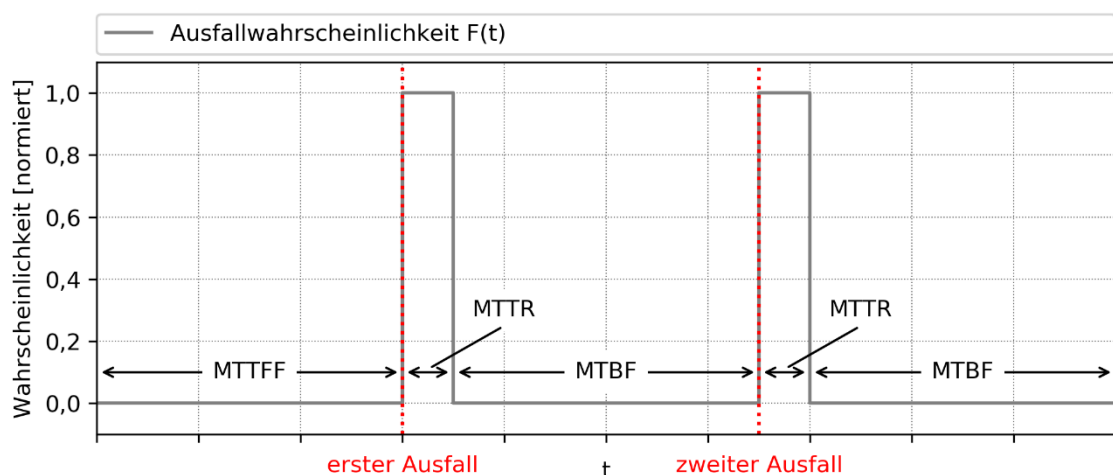


Bild 4.1: Wahl der Delta-Verteilung für die Eintrittswahrscheinlichkeit von Ausfällen: Dargestellt ist der zeitliche Verlauf der Ausfallwahrscheinlichkeit $F(t)$, wie er in dieser Arbeit durch Annahme einer Delta-Verteilung für die Ausfallwahrscheinlichkeit angenommen wird. Dargestellt werden zwei Ausfälle mit entsprechender Reparaturzeit $MTTR$, welche entsprechend der gewählten Delta-Verteilung exakt nach einer mittleren Lebensdauer ($MTTFF$ vor erstem Ausfall und $MTBF$ vor weiterem Ausfall) eintreten.

Die Lebensdauer einer Komponente L_K spiegelt somit die funktionsfähige Zeit einer Einheit wider. Nachdem diese Zeit abgelaufen ist, kommt es bei einer idealisierten Betrachtung mittels einer Delta-Funktion zu einem Ausfall zu einem exakten Zeitpunkt t . Damit erfüllt dann die übergeordnete Einheit, also das technische System, nicht mehr seine Funktion und muss instandgesetzt werden. Es kommt zu einer technischen Ausfallzeit T_T mit möglichen Folgeausfallzeiten T_{FS} ³³ und damit einer Reparaturzeit von $MTTR = T_T + T_{FS}$.

Die Dauer des Sprungs $s=1$ ist abhängig von der sich an den Ausfall anschließenden Reparaturzeit $MTTR$. Diese kann oftmals variieren durch die Komplexität der Reparatur, der Ausführlichkeit der Maschinendokumentation, dem Expertenwissen des Instandhalters und beispielsweise dem Instandhaltungsmanagement. Es ist daher nicht sinnvoll, im Weiteren individuelle Reparaturzeiten zu betrachten. Das vorliegende Beispiel in Bild 4.1 stellt zwei Ausfälle mit einer Zeitdauer von 10 % der Betriebsdauer B für ein System dar. Die Lebensdauer einer Komponente L_K deckt in Summe 90 % ab. Dies entspricht damit einer technischen Verfügbarkeit $V_T = 90 \%$.

Weiterführend wird davon ausgegangen, dass keine unbekanntem Einflüsse auf die einzelnen Komponenten und somit auf das System wirken sowie keine Korrelationen zwischen Ausfällen verschiedener Komponenten bestehen. Daher ergibt sich auch keine Streuung der Ausfallzeitpunkte und damit auch keine Veränderung der mittleren Lebensdauern von einzelnen Komponenten. Exemplarisch bedeutet dies, dass wenn Komponente A ausfällt, hat dies keinen weiteren Einfluss auf die zusammenhängende Komponente B.

Die Grundlage für die Herleitung der weiteren Methode zur Vorhersage der Vorhersagegüte sind demnach die Ausfälle a , welche aus den Lebensdauern der Komponenten L_K unter Annahme einer Delta-Funktion bestimmt werden.

4.2.1 Auswahl der notwendigen Gütemaße für eine Leistungsbewertung der Datenmenge

Mit Hilfe eines idealisierten Klassifikators lassen sich Ausfälle a für eine Leistungsbewertung in vorhergesagte und nicht-vorhergesagte Ausfälle aufteilen. Die generelle Vorgehensweise zur Leistungsbewertung eines Klassifikators wurde bereits in Abschnitt 3.2.1 ausgeführt. In diesem Abschnitt werden nun die zu prognostizierenden Gütemaße konzipiert und definiert.

Bild 4.2 stellt die Aufteilung der Gesamtmenge an Ausfällen³⁴ dar, dabei wird noch einmal deutlich, dass nur Elemente der Menge der wahren Ausfälle P ("in Wahrheit negativ") aufgrund des idealen Klassifikators auftreten. Folglich werden nur die Mengen der Nicht-Alarme FN , der Treffer TP und die Gesamtmenge P sowie deren Gütemaße nachfolgend berücksichtigt und berechnet.

³³ Folgeauszeiten setzen sich aus organisatorischen, technischen und wartungsbedingten Folgeausfallzeiten zusammen. $T_{FS} = \sum(T_{FO} + T_{FT} + T_{FW})$

³⁴ P = Anzahl der (Ausfälle = True); konträr: N = Anzahl der (Ausfälle = False)

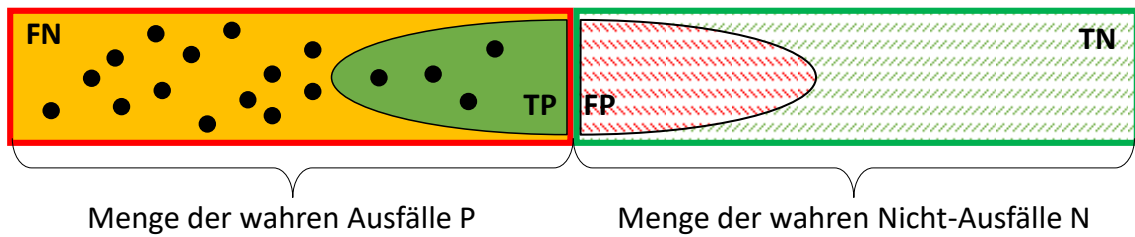


Bild 4.2: Schematische Darstellung der berechneten Mengen für ein Beispielsystem bei idealisierter Klassifikation: Dargestellt ist die Menge der Treffer TP (in Grün), die Menge der Nicht-Alarme FN (in Orange) und deren Vereinigung als Menge der wahren Ausfälle P (Rot umrandet). Zusätzlich ist die Menge der Fehlalarme FP (in Rot gestrichelt), die Menge der True-Negatives TN (in Grün gestrichelt) und deren Vereinigung als Menge der wahren Nicht-Ausfälle N (Grün umrandet) dargestellt.

Die Menge³⁵ der Daten(punkte) beschreibt somit nur die positive Klasse P („In Wahrheit Ausfall“). Somit ergibt sich die in Tabelle 4.1 dargestellte Konfusionsmatrix. Die Anzahl an Elementen in einer Menge wird durch die Mächtigkeit beschrieben³⁶.

Tabelle 4.1: Konfusionsmatrix bei idealisierter Klassifikation: Aufgelistet ist die Anzahl der Treffer $|TP|$, die Anzahl der Nicht-Alarme $|FN|$ und resultierend die Anzahl der wahren Ausfälle $|P|$, als Element der natürlichen Zahl \mathbb{N} . Zusätzlich ist die Anzahl der Fehlalarme $|FP|$, die Anzahl der True-Negatives $|TN|$ und deren resultierende Anzahl der wahren Nicht-Ausfälle $|N|$ dargestellt, welche immer null sind.

	Als Ausfall vorhergesagt	Als Ausfall nicht vorhergesagt	Summe
In Wahrheit Ausfall	$ TP \in \mathbb{N}$	$ FN \in \mathbb{N}$	$ P = TP + FN \in \mathbb{N}$
In Wahrheit kein Ausfall	$ FP = 0$	$ TN = 0$	$ N = 0$
Summe	$ P^* = TP \in \mathbb{N}$	$ N^* = FN \in \mathbb{N}$	$ P + N \in \mathbb{N}$

Die bereits eingeführten Gütemaße aus Abschnitt 3.2.1 werden nachstehend nun für die ausgearbeitete Methode festgelegt. Somit ergibt sich für die Periodenprävalenz³⁷ immer 1, denn der Anteil an wahren Ausfällen $|P|$ zur Gesamtzahl der Ausfälle $(|P| + |N|)$ ist aufgrund des idealen Klassifikators immer gleich. Daraus folgt, dass der komplette Datensatz nur aus Ausfällen besteht. Die Anzahl der wahren Nicht-Ausfälle $|N|$ ist dementsprechend null. Folglich können für die bereits eingeführten Gütemaße aus Tabelle 3.3 die Maße in der entwickelten Methode berechnet werden:

³⁵ In der Literatur wird häufig vereinfachend der Begriff Summe statt des mathematisch korrekten Begriffs Menge verwendet. Für die korrekte Beschreibung der Methodik und Anwendung dieser im folgenden Kapitel wurde die Benennung jedoch an den wissenschaftlich korrekten Sprachgebrauch angepasst.

³⁶ In der Literatur wird oftmals die Summe und Menge nicht unterschieden. Mathematisch korrekter ausgedrückt sind $P, P^*, TP, FP, N, N^*, FN, TN$ Mengen von Daten und $|P|, |P^*|, |TP|, |FP|, |N|, |N^*|, |FN|, |TN|$ die Anzahl.

³⁷ Die Prävalenz beschreibt den Anteil der Daten, die einen Ausfall beschreiben, in Bezug auf die Summe aller Daten. Die gewählte Periodenprävalenz beschreibt dieses Verhältnis für einen bestimmten Zeitraum. Im Gegensatz dazu steht die Punktprävalenz, welche das Verhältnis nur zu einem bestimmten Zeitpunkt angibt.

- Die True-Positiv-Rate $TPR = |TP|/|P|$ beschreibt die Trefferrate für die Vorhersage.
- Die False-Negativ-Rate $FNR = |FN|/|P|$ beschreibt die Rate der fehlenden Alarme für einen Ausfall und wird in dieser Arbeit als Nicht-Alarmrate benannt.
- Die False-Positiv-Rate $FPR = |FP|/|N|$ kann nicht berechnet werden, da sowohl die Anzahl der Fehlalarme $|FP|$ als auch die Anzahl der wahren Nicht-Ausfälle $|N|$ null ist. Die Fehlalarmrate FPR entspricht der Rate der als Ausfall vorhergesagten, aber in Wahrheit keinen Ausfall beschreibenden Daten.
- Die True-Negativ-Rate $TNR = |TN|/|N|$ beschreibt die sogenannte Spezifität, das heißt die korrekt klassifizierten Nicht-Ausfälle. In der hier entwickelten Methodik kommen ausschließlich Ausfälle vor, sodass diese Rate nicht bestimmt werden kann.
- Die Korrektklassifikationsrate $= |TP| + |TN|/|P| + |N|$ entspricht, da keine wahren Nicht-Ausfälle $|N|$ vorliegen, der Trefferrate TPR .
- Die Falschklassifikationsrate $= |FP| + |FN|/|P| + |N|$ entspricht, da keine wahren Nicht-Ausfälle $|N|$ vorhanden sind, der Nicht-Alarmrate FNR .
- Der Positive-Predictive-Value $PPV = |TP|/|P^*|$ beschreibt die Wirksamkeit der Vorhersagen. P^* ist die Menge und $|P^*|$ die Anzahl der als Ausfall vorhergesagten Daten. Durch die fehlende Anzahl der Fehlalarme $|FP|$ ist $PPV = 1$.
- Der Negative-Predictive-Value $NPV = |TN|/|N^*|$ beschreibt die Trennfähigkeit der Vorhersagen. N^* ist die Menge und $|N^*|$ die Anzahl der Daten, für welche kein Ausfall vorhergesagt wird. Da keine Nicht-Ausfälle $|TN|$ vorliegen ist der $NPV = 0$.

Da die Trefferrate TPR gleich der Korrektklassifikationsrate ist und die Nicht-Alarmrate FNR gleich der Falschklassifikationsrate ist, reicht es aus, im Weiteren nur die Trefferrate TPR und Nicht-Alarmrate FNR als Gütemaße zu betrachten. Die Fehlalarmrate FPR und auch die Spezifitätsrate TNR werden nicht betrachtet, da diese bei idealisiertem Klassifikator irrelevant sind. Einmalig kann der positive und negative Vorhersagewert PPV bzw. NPV berechnet werden. Allerdings muss dieser künftig nicht weiter berechnet werden, da dieser auf Basis der Daten unveränderlich sein wird. Die Wirksamkeit PPV für die Vorhersage von Ausfällen liegt bei $PPV = 1$. Gleichzeitig bedeutet die Trennfähigkeit $NPV = 0$, dass keine negativ-vorhergesagten Ausfälle auftreten werden.

Die entwickelte Methode muss daher einzig die Menge der Treffer TP und der Nicht-Alarme FN ermitteln, wodurch dann die Leistung und somit die bestmögliche Güte des Vorhersagemodells bestimmt werden kann:

- Trefferrate $TPR = |TP|/|P| \Rightarrow$ Korrektklassifikationsrate
- Nicht-Alarmrate $FNR = |FN|/|P| \Rightarrow$ Falschklassifikationsrate

4.2.2 Bestimmung von relevanten Gütemaßen

Zunächst sind die absoluten Gütemaße die Anzahl der Treffer TP und der Nicht-Alarme FN sowie die Anzahl der wahren Ausfälle P zu berechnen. Die Berechnungsvorschrift für einen Ausfall $a_{Kn} = L_K * n$ wurde bereits in Formel 4.2 vorgestellt, sodass deutlich wurde, dass ein Ausfall immer nach dem gleichen Zeitraum für die gleiche Komponente auftritt und der Ausfallzeitpunkt eindeutig ist. Die Ausfälle a werden als nächstes durch verschiedene Bedingungen in die unterschiedlichen Klassen aufgeteilt, umso die Gütemaße zu prognostizieren.

Begonnen wird mit der Bestimmung der wahren Ausfälle P über den Betriebszeitraum. Der Betriebszeitraum entspricht $t_B = [0, B[h$. Das heißt eine Bedingung für die zur Menge der wahren Ausfälle P gehörenden Ausfälle ist es, dass ein Ausfall a kleiner als die Betriebsdauer B ist, sonst würde der Ausfall auch in der Realität nicht mehr auftreten, da die Komponente K nicht mehr betrieben wird. Die Menge der wahren Ausfälle einer Komponente P_K kann daher wie folgt angegeben werden:

$$P_K = \{(a_{K1}, \dots, a_{Kn}, a_{Kn+1}) \mid n = \mathbb{N} \wedge a_{Kn} < B\}, \text{ mit } a_{Kn} = L_K * n \quad (4.4)$$

Im nächsten Schritt muss entscheiden werden, ob ein Ausfall a_{Kn} ein Treffer TP in der Klassifikation ist oder ob ein Nicht-Alarm FN vorliegt. Hierzu muss erstmal die Anzahl der notwendigen, bereits aufgetretenen Ausfälle s_K festgelegt werden. Je nach Klassifikations-Algorithmus ist eine unterschiedliche Anzahl an Ausfällen notwendig, auch trägt die Trennschärfe der Daten für einen Ausfall bzw. Nicht-Ausfall zur Anzahl s_K bei. In der entwickelten Methode wird die Anzahl der notwendigen Ausfälle für einen idealisiert betrachteten Treffer durch s_K parametrisiert. Beispielfhaft würde $s_K = 1$ bedeuten, dass der erste Ausfall einer Komponente ausreicht, um den zweiten Ausfall vorherzusagen. Sollte kein Ausfall zum Erlernen notwendig sein und sofort eine Vorhersagemodell zur Verfügung stehen, wäre $s_K = 0$ zu wählen.

Zur Verdeutlichung wird ein neudefiniertes Maß, die Menge der erlernbaren Ausfälle (Learn-Positive, kurz: LP), eingeführt. Dieses weitere Maß dient der verbesserten Nachvollziehbarkeit bei der Berechnung der weiteren Gütemaße. Für die Bestimmung der Menge an erlernbaren Ausfällen LP wird nur jeweils der n -te Ausfall einer Komponente, der genau s_K entspricht, herangezogen und zur entsprechenden Menge der erlernten Ausfälle LP hinzugefügt. Somit ergibt sich:

$$LP_K = \{(a_{Kn}) \mid a_{Kn} \in P_K \wedge n \equiv s_K\} \quad (4.5)$$

Nachdem Ausfälle erlernt wurden, können weitere Ausfälle einer Komponente vermieden werden. Die korrekt vorhergesagten Ausfälle werden auch Treffer oder True-Positiv (TP) genannt. Das bedeutet für die Menge der Treffer TP das alle Ausfälle a_{Kn} , welche nach dem n -ten-Ausfall einer Komponente bzw. nach der Anzahl für das Erlernen notwendiger Ausfälle s_K aufgetreten sind, vorhergesagt werden können. Die Zuordnung der Ausfälle a_{Kn} zur Menge der Treffer TP erfolgt unter den genannten Bedingungen:

$$TP_K = \{(a_{Kn}) \mid a_{Kn} \in P_K \wedge n > s_K\} \quad (4.6)$$

Andersherum bedeutet dies, dass sich nun die Menge für die Nicht-Alarme FN bestimmen lässt, denn alle Ausfälle einer Komponente a_{Kn} , welche bis einschließlich des erlernten Ausfalls auftreten können nicht korrekt klassifiziert werden. Somit erfolgt hierbei die Zuordnung eines Ausfalls a_{Kn} zur Menge FN wie folgt:

$$FN_K = \{(a_{Kn}) \mid a_{Kn} \in P_K \wedge n \leq s_K\} \quad (4.7)$$

Wie bereits in Abschnitt 4.1.2 dargestellt, sollen mittels der Methodik im späteren Verlauf der Arbeit für unterschiedliche Szenarien die Güte von Vorhersagemodellen für die voraussagende Instandhaltung bestimmt werden. Die bisher hergeleitete Methode berücksichtigt noch nicht den Beginn der Datenerfassung t_{data} . Folglich ist die Vorgehensweise zur Berechnung, der erlernbaren Ausfälle LP , der vermeidbaren Ausfälle TP und der verbleibenden Ausfälle FN auszubauen. Die Berechnung der wahren Ausfälle P bleibt unverändert.

Die Konsequenzen einer späteren Datenerfassung sind leicht nachvollziehbar, ein Ausfall wird erst später erlernt und somit reduziert sich die Anzahl der Treffer TP und mehr Ausfälle FN verbleiben. In der Methode lässt sich dies durch die Einführung eines weiteren Parameters o_K , welcher einen Offset für die Anzahl der zu erlernenden und verpassten Ausfälle wiedergibt, abbilden. Die Berechnung erfolgt nach Formel 4.8.

$$o_K = |\{(a_{Kn}) \mid a_{Kn} \in P_K \wedge a_{Kn} < t_{data}\}| \quad (4.8)$$

Die bisher eingefügten Formeln 4.5 bis 4.7 werden entsprechend erweitert, so dass diese implizit abhängig vom Beginn der Datenerfassung t_{data} sind.

$$LP_K = \{(a_{Kn}) \mid a_{Kn} \in P_K \wedge n \equiv s_K + o_K\} \quad (4.9)$$

$$TP_K = \{(a_{Kn}) \mid a_{Kn} \in P_K \wedge n > s_K + o_K\} \quad (4.10)$$

$$FN_K = \{(a_{Kn}) \mid a_{Kn} \in P_K \wedge n \leq s_K + o_K\} \quad (4.11)$$

Abschließend ist festzuhalten, dass korrekterweise eigentlich ein Wertepaar in den Mengen der wahren Ausfälle P , der erlernten Ausfälle LP , der Treffer TP und der Nicht-Ausfälle FN zu vermerken ist. Das soll heißen, dass das Element einer Menge mit (a_{Kn}, y_n) statt (a_{Kn}) einwandfrei ausgedrückt ist. Bisher wurde darauf verzichtet, da die Anzahl y_n eines Ausfalls a_{Kn} zu einem gegebenen Ausfall immer eins $y_n = 1$ ist.

$$P_K = \{(a_{K1}; y_1), \dots, (a_{Kn}; y_n), (a_{Kn+1}; y_{n+1}) \mid n = \mathbb{N} \wedge a_{Kn} < B\}, \quad (4.12)$$

mit $a_{Kn} = L_K * n$ und $y_n = 1$

$$LP_K = \{(a_{Kn}; y_n) \mid (a_{Kn}; y_n) \in P_K \wedge n \equiv s_K + o_K\} \quad (4.13)$$

$$TP_K = \{(a_{Kn}; y_n) \mid (a_{Kn}; y_n) \in P_K \wedge n > s_K + o_K\} \quad (4.14)$$

$$FN_K = \{(a_{Kn}; y_n) \mid (a_{Kn}; y_n) \in P_K \wedge n \leq s_K + o_K\} \quad (4.15)$$

4.2.3 Vorgehensweise zur Verwendung der Gütemaße

Eine besondere Herausforderung im Vergleich zu konventionellen Ansätzen der Leistungsbewertung ist es, dass nicht nur für einen gegebenen Zeitraum die mögliche Leistung bestimmt werden soll, sondern auch der zeitliche Verlauf von den Gütemaßen zu berücksichtigen und zu ermitteln ist. Die Methode erfordert also eine Bestimmung der Anzahl von Ausfällen zu jedem Zeitpunkt, demzufolge ist es nicht mehr ausreichend einfach die Mächtigkeit der Menge und damit die Anzahl der Wertepaare in einer Menge zu bestimmen, sondern eine Erweiterung der Methode durch eine zeitliche Angabe ist erforderlich. Folglich kann ein Data Scientist erkennen, zu welchem Zeitpunkt welcher Wert des Gütemaßes vorliegt oder auch der/die Anwenderin, welche Ausfälle noch zu erwarten sind.

Aus der Berechnungsvorgabe für einen Ausfall kann abgeleitet das dieser einem eindeutigen Zeitpunkt entspricht.

$$a_{Kn} = L_K * n \Rightarrow t \quad (4.16)$$

Da jedoch ein Ausfall einem Zeitpunkt entspricht $a_{Kn} = L_K * n \Rightarrow t$ kann a_{Kn} entsprechend ersetzt werden, wodurch sich die Formeln 4.17 bis 4.20 ergeben.

$$P_K(t) = \{(t_1; y_1), \dots, (t_n; y_n), (t_{n+1}; y_{n+1}) \mid n = \mathbb{N} \wedge t_n < B\}, \quad (4.17)$$

mit $a_{Kn} = L_K * n \Rightarrow t$ und $y_n = 1$

$$LP_K(t) = \{(t_n; y_n) \mid (t_n; y_n) \in P_K \wedge n \equiv s_K + o_K\} \quad (4.18)$$

$$TP_K(t) = \{(t_n; y_n) \mid (t_n; y_n) \in P_K \wedge n > s_K + o_K\} \quad (4.19)$$

$$FN_K(t) = \{(t_n; y_n) \mid (t_n; y_n) \in P_K \wedge n \leq s_K + o_K\} \quad (4.20)$$

Bringt man nun die zeitliche Einheit t in Abhängigkeit zu einem Gütemaß, so lässt sich für einen beliebigen Zeitraum $[a; b[$ die Anzahl von wahren Ausfällen P , erlernbaren Ausfällen LP , vermeidbaren Ausfällen TP oder verbleibenden Ausfällen FN angeben.

Im Weiteren wird außerdem eine Komponente immer zusammen mit anderen Komponenten eines Produktionssystems betrachtet. Dies ist insofern sinnvoll, da die Implementierung von einer voraussagenden Instandhaltung für ein System erfolgt und dann alle geeigneten Komponentenausfälle vorhergesagt werden sollen.

Die nächstfolgende Formel 4.21 definiert die Berechnung eines absoluten Gütemaßes (P , LP , FN oder TP) für ein System. Jetzt wird auch die Angabe von y_n im Wertepaar der Menge $(t_n; y_n)$ relevant. Sollte nämlich in mehr als einer Komponente ein Ausfall zu einem gleichen Zeitpunkt t auftreten, so erhöht sich der Wert von y entsprechend.

$$(P \vee LP \vee FN \vee TP)_{System}(t) = \sum_{i=1}^{|K|} (P \vee LP \vee FN \vee TP)_i(t) \quad (4.21)$$

Aus den bisherigen absoluten Gütemaßen der Anzahl der wahren Ausfälle P , der Treffer TP und der Nicht-Alarme FN lassen sich nun auch die relativen Gütemaße Trefferrate TPR und Nicht-Alarmrate FNR berechnen. Allerdings muss in der Formel für die Trefferrate TPR bzw. für die Nicht-Alarmrate FNR ebenso die zeitliche Abhängigkeit berücksichtigt werden, woraus sich ergibt:

$$TPR(t) = TP(0 \leq t) / P(0 \leq t < B) \quad (4.22)$$

$$FNR(t) = FN(0 \leq t) / P(0 \leq t < B) \quad (4.23)$$

Durch $P(0 \leq t < B)$ wird nun die Gesamtheit aller Ausfälle P beschrieben, denn durch die Einführung der zeitlichen Abhängigkeit ist die Mächtigkeit der Menge nicht mehr eindeutig definiert und damit deren Anwendung alleine nicht mehr ausreichend. Ebenso wird die Anzahl der Treffer TP und der Nicht-Alarme FN jetzt zeitabhängig im Intervall $[0; t]$, bestimmt.

Neben der Trefferrate TPR und der Nicht-Alarmrate FNR ist es sinnvoll, auch die Ausfallrate PR zu bestimmen. Hierzu wird die Anzahl der wahren Ausfälle P bis zu einem Zeitpunkt t bestimmt und durch die Gesamtanzahl der Ausfälle P über den Betriebszeitraum $t_B = [0, B[$ h bestimmt.

$$PR(t) = P(0 \leq t) / P(0 \leq t < B) \quad (4.24)$$

Für das neueingeführte absolute Gütemaß der Anzahl der erlernten Ausfälle LP lässt sich auch ein relatives Gütemaß LPR (Learn-Positive-Rate) bestimmen und somit der Bezug zur Anzahl der Komponenten $|K|$ und somit zur maximal erlernbaren Anzahl von Ausfalltypen herstellen. Die Formel hierzu lautet wie folgt:

$$LPR(t) = LP(0 \leq t) / |K| \quad (4.25)$$

Abschließend ist festzuhalten, dass neben den geforderten Gütemaßen – Anzahl der wahren Ausfälle P , Anzahl der vorhersagbaren Ausfälle TP , Anzahl der verbleibenden Ausfälle FN , Trefferrate TPR und Nicht-Alarmrate FNR – noch weitere Gütemaße – Ausfallrate PR , Anzahl der erlernbaren Ausfälle LP und Lernrate LPR – unter Berücksichtigung der zeitlichen Abhängigkeit bestimmt werden können.

4.3 Modellierung der Gütemaße

Die Betrachtung der Verteilung der Gütemaße über die Zeit t ist für die Bestimmung einer ausreichenden Datenbasis unerlässlich. Hierzu wurde bereits im Vorfeld die Abhängigkeit zwischen den Gütemaßen und der Zeit hergestellt. Im Folgenden geht es nun darum diese Gütemaße über die Zeit quantifizierbar und damit interpretierbar zu machen. Hierzu ist es notwendig, von einer zeit- und wertdiskreten Verteilung zu einer zeitkontinuierlichen Verteilung zu kommen. Die nachfolgenden Abschnitte stellen zwei Möglichkeiten vor dies zu erreichen.

4.3.1 Vorgehensweise zur Bestimmung einer empirischen Verteilungsfunktion

Die aktuelle Verteilung der mittleren Komponenten-Lebensdauern und der davon abgeleiteten Gütemaße sind zeit- und wertdiskret, d.h. es gibt nur zu bestimmten Zeitpunkten t ein entsprechendes Wertepaar $(t_n; y_n)$. Geht man davon aus, dass ein Ausfall allerdings bis zum nächsten Zeitpunkt t_{n+1} ansteht, ergibt sich eine empirische Verteilungsfunktion, welche aus einem eindeutigen Zeitpunkt t_n ein Intervall $[t_n; t_{n+1}[$ macht. Allgemein formuliert ergibt sich dementsprechend für jedes Wertepaar der Menge nachstehende Anpassung:

$$(P \vee LP \vee FN \vee TP)(t) = \{([t_n; t_{n+1}[; y_n)\} \quad (4.26)$$

Durch die Verwendung eines Intervalls wird also eine zeitkontinuierliche, aber noch wertdiskrete Verteilung erzeugt, bei der für jeden Zeitpunkt t im Definitionsbereich $D_f = t_B = [0, B[$ h ein Wert y bestimmt werden kann. Oftmals wird dieser auch als Treppenverlauf bezeichnet.

4.3.2 Vorgehensweise zur Bestimmung einer Näherungsfunktion

Eine weitere Möglichkeit zur Beschreibung der Gütemaße über die Zeit t ist die Verwendung einer Funktionsgleichung für die Näherung der jeweils zugrunde liegenden Verteilung. Dadurch kann eine Verteilung bestimmt werden, welche alle Wertepaare und damit Datenpunkte im zweidimensionalen Raum bestmöglich annähert. Durch eine solche Näherungsfunktion ist es möglich, die Verteilung wert- und zeitkontinuierlich zu bestimmen.

Grundsätzlich ist die Aufstellung einer Näherungsfunktion von Vorteil, da dann für jeden beliebigen Zeitpunkt t ein konkreter Wert der maximal erreichbaren Güte des entsprechenden Maßes abgeschätzt werden kann. Dahingegen würden die empirischen Verteilungen einen stufenförmigen Anstieg und damit die davon abgeleiteten Gütemaße einen diskretisierten Werteverlauf zeigen, welcher nicht der Realität entspricht.

Zur Bestimmung einer Näherungsfunktion ist eine Funktionsgleichung n-ten-Grades³⁸ mit Hilfe der Datenpunkte des entsprechenden Gütemaßes aufzustellen. Wie gut die Anpassung der Näherungsfunktion n-ten-Grades an die ursprüngliche Verteilung ist und welcher Grad der Gleichung zu wählen ist, lässt sich durch die Berechnung der Summe der quadrierten Abstände zwischen Funktionswerten und Datenpunkte (die allseits bekannte Methode der kleinsten Quadrate) bestimmen.

Durch solch eine parametrisierte Näherungsfunktion lässt sich die statistische Unsicherheit der Lebensdauern L_K und damit korrelierend der Ausfallzeitpunkte a_{Kn} berücksichtigen. Daher ist es sinnvoll, solche genäherten Verteilungsfunktionen für

³⁸ Bsp. für eine Funktionsgleichung 3. Grades: $f(x) = a_3x^3 + a_2x^2 + a_1x + a_0$

die Bestimmung der Vorhersagegüte zu einem beliebigen Zeitpunkt t in einem geeigneten Definitionsbereich D_f zu verwenden. Daher finden diese auch im nachfolgenden Teil der Arbeit Anwendung.

4.4 Fazit

Dieses Kapitel hat gezeigt, wie basierend auf entsprechenden Rahmenbedingungen eine Bestimmung der Vorhersagegüte für Systeme erfolgen kann. Es wurden dazu drei Szenarien eingeführt, welche wie folgt charakterisiert sind: Das erste Szenario ist definiert durch einen Beginn der Datenerfassung und -analyse zu Betriebsbeginn, also $t_{\text{data}} = 0 h$. Das zweite Szenario stellt einen späteren Beginn der Datenerfassung ($t_{\text{data}} > 0 h$) dar und das dritte Szenario nutzt die Daten bzw. das darauf basierende Vorhersagemodells aus anderen baugleichen Systemen.

Nachdem die Voraussetzungen für die Methodik definiert wurden, konnte im folgenden Abschnitt die Methode hergeleitet werden. Hierzu wurden die zu bestimmenden Gütemaße ausgewählt und anschließend entsprechende Berechnungsvorschriften aufgestellt. Damit eine Anwendbarkeit im zweiten Szenario und auch für eine zeitliche Betrachtung möglich ist, wurde die Methode nachfolgend entsprechend iterativ noch erweitert.

Eine zusätzliche Erweiterung der Methode erfolgte im dritten Abschnitt, welche die bisherige zeitdiskrete Verteilung in eine zeitkontinuierliche Verteilung überführte und somit die bisherigen Datenpunkte mittels Näherungsfunktionen modellierte.

5 Bestimmung der Vorhersagegüte für ein Bearbeitungszentrum

Im folgenden Kapitel wird die im vorherigen Kapitel 4 entwickelte Methodik zur Bestimmung der Vorhersagegüte an einem Bearbeitungszentrum mit 15 Komponenten angewendet. Damit lässt sich für dieses System bereits die bestmögliche Güte eines Vorhersagemodells prognostizieren.

5.1 Analyse der Ausgangslage

Die entwickelte Methode aus Kapitel 4 wird nun für die Bestimmung der Vorhersagegüte eines Bearbeitungszentrums (für die übersichtliche Benennung wird BAZ 1 verwendet) angewendet.

Hierzu sind zunächst die Lebensdauern der Komponenten zu bestimmen. Für das Bearbeitungszentrum 1 (BAZ 1) wurden 15 mittlere Lebensdauern (MTBF) von Komponenten (L_K) ermittelt. Die Ermittlung basiert auf Untersuchungen und Versuchen der Komponenten im Gesamtverbund des Systems und werden manchmal, so wie in diesem Fall, in TCO-Berichten (Total Cost of Ownership) zusätzlich dokumentiert. In Tabelle 5.1 sind die Komponenten und deren Lebensdauern sowie deren Funktion dokumentiert. Neben Antrieben für unterschiedliche Komponenten sind Führungen und ein Wegmesssystem dokumentiert worden.

Tabelle 5.1: Informationen über die Komponenten von System BAZ 1: Auflistung der Komponenten und deren Funktion sowie deren mittlere Lebensdauer (MTBF in Stunden h) im System BAZ 1

Komponente	Funktion	MTBF in h
Zuführtechnik	Fördern	10.000
Spindeleinheit	Kühlung	10.000
Werkstückspannung	Spanneinrichtung	10.000
Linearachsen	Antrieb	15.000
Werkstückwechseleinrichtung	Führung	18.000
Werkstückwechseleinrichtung	Antrieb	20.000
Revolver	Antrieb	20.000
Linearachsen	Führung	20.000
Linearachsen	Lagerung	20.000
Werkstückspannung	Spannfutter	20.000
Linearachsen	Wegmesssystem	25.000
Revolver	Lagerung	30.000
angetriebene Werkzeuge, Bohrköpfe	Lagerung	30.000
Spindeleinheit	Lagerung / Führung	30.000
Linearachsen	Absinksperr	100.000

Basierend auf Tabelle 5.1 kann für BAZ 1 folgende Menge der Lebensdauern von Komponenten L_K festgehalten werden.

BAZ 1 = {10.000, 10.000, 10.000, 15.000, 18.000, 20.000, 20.000, 20.000, 20.000, 20.000, 25.000, 30.000, 30.000, 30.000, 100.000}

Für einen späteren Vergleich aller Szenarien mit den zugrunde liegenden beobachteten Ausfällen in einer realen Produktion wird nachfolgend die Anzahl der wahren Ausfälle P berechnet (Formel 4.12). Hierbei wird davon ausgegangen, dass kein Vorhersagemodell eingesetzt wird und somit lediglich auf eine reaktive Instandhaltung zurückgegriffen wurde. Der Zeitraum für die Untersuchung von Ausfällen ist auf Basis eines Lebenszyklus einer Maschine bzw. der Betriebsdauer B gewählt, welche in der Regel bei 15 Jahren liegt [20]. Da die Lebensdauern von Komponenten L_K in der Praxis in Stunden angegeben werden (siehe [97]), wird für den weiteren Verlauf dieser Arbeit diese Einheit gewählt. Daraus abgeleitet beträgt die Betriebsdauer in der Regel $B = 131.400 h$. Dies bedeutet auch, dass Lebensdauern $L_K \geq B$ nicht berücksichtigt und aussortiert werden, da Komponenten mit Lebensdauern L_K größer gleich als die gewählte Betriebsdauer B für die nachfolgenden Untersuchungen ohne Bedeutung sind.

Zur Verdeutlichung der Vorgehensweise wurde ein Struktogramm (Bild 5.1) entsprechend Formel 4.12 erstellt. Damit wird verständlich, dass die bereits erhobenen Daten – die Lebensdauern L_K und die Betriebsdauer B – Eingangsgrößen für die Berechnung der Menge P_K sind. Berechnet werden die Ausfallzeitpunkte a_n nach Formel 4.2, wobei demnach nur Ausfallzeitpunkte a_n kleiner der Betriebsdauer B betrachtet werden müssen: $a_n < B$.

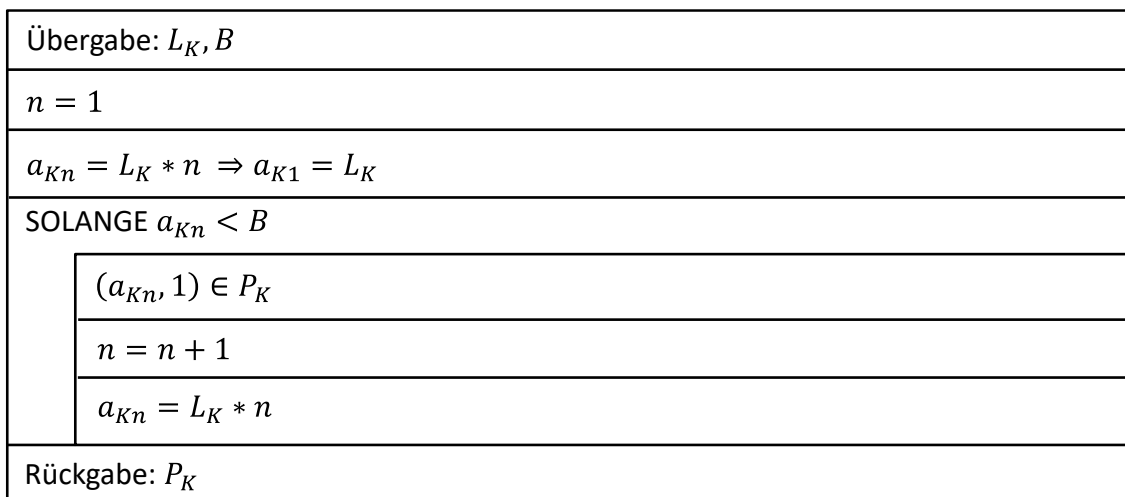


Bild 5.1: Struktogramm für die Berechnung der Menge der wahren Ausfälle P entsprechend Szenario 1: Dargestellt ist der einzelne Ausfall a indiziert durch die Anzahl n der bisherigen Ausfälle der Komponente K . In der Schleife wird entschieden, ob der untersuchte Ausfall a_{Kn} kleiner der Betriebsdauer B ist. Ist die Bedingung erfüllt, wird der Ausfall a_{Kn} zur Menge der wahren Ausfälle P_K hinzugefügt. Sobald die Bedingung $a_{Kn} < B$ nicht mehr erfüllt ist, wird die Schleife beendet und die Menge der wahren Ausfälle P_K ausgegeben.

Für eine erste Komponente des Systems BAZ 1 wird beispielhaft schrittweise die Menge der wahren Ausfälle P bestimmt.

$$L_{BAZ\ 1, Komponente\ 1} = 10.000\ h$$

Als Betriebsdauer B und somit Nutzungszeit T_N der Anlage wurden wie bereits erklärt 15 Jahre bzw. 131.400 Stunden angenommen.

$$B = 15\ Jahre = 15 * 365 * 24 = 131.400\ h$$

Nach Formel 4.12 bzw. dem Struktogramm aus Bild 5.1 ergibt sich folgendes Ergebnis. D.h. alle 10.000 Stunden tritt ein Ausfall auf.

Erg. 5.1: $P_{BAZ\ 1, Komponente\ 1}(t) = \{(10.000; 1), (20.000; 1), \dots, (130.000; 1)\}$

Führt man die Berechnung für alle 15 Komponenten K durch, so ergibt sich für das System BAZ 1 die vereinigte Menge der wahren Ausfälle $P_{BAZ\ 1}$ ³⁹, wobei nun mehr als ein Ausfall zu einem Zeitpunkt t vorliegen kann.

$$P_{BAZ\ 1}(t) = \{(10.000; 3), (15.000; 1), (18.000; 1), (20.000; 8), (25.000; 1), \\ (30.000; 7), (36.000; 1), (40.000; 8), (45.000; 1), (50.000; 4), (54.000; 1),$$

Erg. 5.2: $(60.000; 12), (70.000; 3), (72.000; 1), (75.000; 2), (80.000; 8), (90.000; 8), \\ (100.000; 10), (105.000; 1), (108.000; 1), (110.000; 3), (120.000; 12), \\ (125.000; 1), (126.000; 1), (130.000; 3)\}$

Damit liegt die Gesamtanzahl der Ausfälle in dem Betriebszeitraum $t_B = [0; B[$ für das System BAZ 1 bei 102.

Erg. 5.3: $P_{BAZ\ 1}(0 \leq t < B) = 102$

Anschließend ist es nun möglich, das relative Gütemaß $PR(t)$ nach Formel 4.24 ($PR(t) = P(0 \leq t) / P(0 \leq t < B)$) zu bestimmen und somit die Datenpunkte über die Zeit t kumuliert darzustellen. Bild 5.2 stellt neben der Datenpunkte der Menge $PR(t)$, auch zwei Möglichkeiten zur Modellierung des Gütemaßes vor (siehe Abschnitt 4.3). Zum einen ist dies eine empirische Verteilung der Daten durch die Annahme einer Delta-Verteilung. Zum anderen eine Beschreibung der Datenpunkte durch eine Näherungsfunktion bzw. genauer eine Polynomialverteilung. In diesem Fall reicht ein Polynom 1. Grades zur Approximation der Datenpunkte der Menge $PR_{BAZ\ 1}$ über der Zeit t aus. Das Residuum r beträgt 0,01. Auch bei Verwendung eines Polynoms höheren Grades lässt sich auf zwei Nachkommastellen gerundet kein geringes Residuum⁴⁰ erzielen und auch im Rahmen einer grafischen Analyse ist keine sichtbare Verbesserung erkennbar gewesen. Die Näherung kann also als hinreichend gut angenommen werden.

³⁹ $P_{BAZ\ 1}(t) = \{(t_n; y_n)\}$, ist die Menge der wahren Ausfälle P im System BAZ 1 beschrieben durch die Anzahl y_n eines Ausfallzeitpunkts t_n .

⁴⁰ 2. Grad $f(x) = -2,01 \cdot 10^{-13} x^2 + 8,27 \cdot 10^{-6} x - 0,07, r = 0,01$

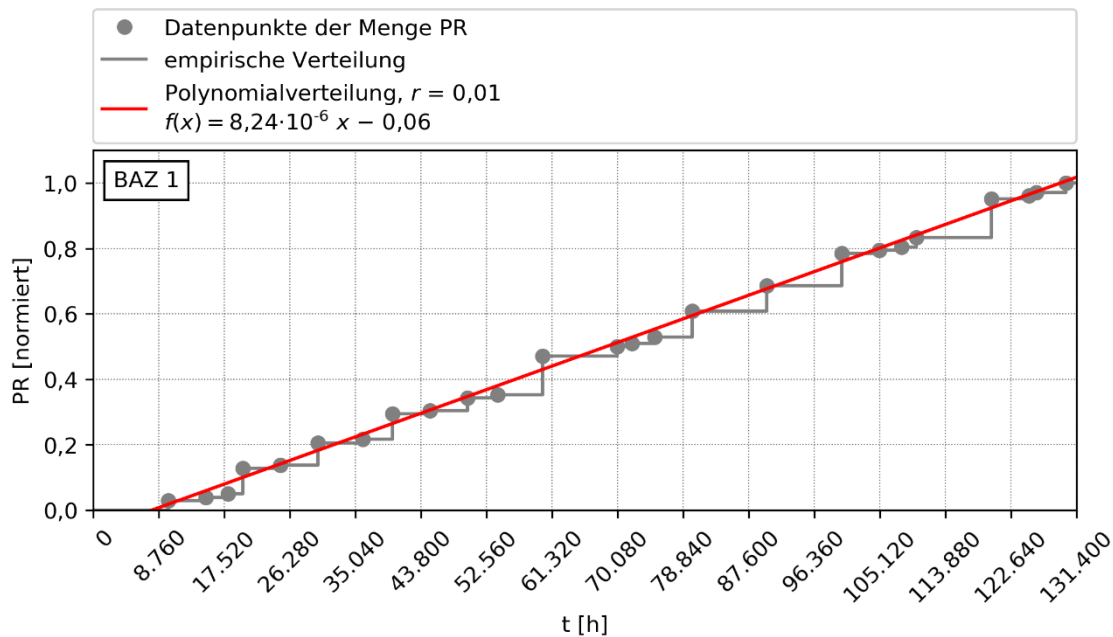


Bild 5.2: Normierte Datenpunkte, empirische und genäherte Verteilung der Ausfallrate PR des Systems BAZ 1 für wahre Ausfälle: Dargestellt sind die Datenpunkte, die normierte empirische Verteilung (in Grau) sowie die zugehörige Näherungsfunktion (in Rot) für die Ausfallrate PR des Systems BAZ 1 über den Betriebszeitraum t_B in Stunden h . Die Verteilungen wurden mittels der Gesamtanzahl der Ausfälle für BAZ 1 $P_{BAZ\ 1}(0 \leq t < B) = 102$ normiert. Für die Näherungsfunktion wurde ein Polynom 1. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400[h$.

Eine Interpretation von Bild 5.2 lässt den Schluss zu, dass die Ausfallrate PR einen kontinuierlichen Anstieg, welcher so oder so ähnlich zu erwarten war, aufzeigt. Dies lässt sich damit begründen, dass ein Vielfaches eines Ausfalls a über der Zeit t bei stetig gleichen Abständen kein anderes Verhalten hervorrufen kann. Allerdings ist dies generell nicht immer zu erwarten, da mehrere Komponenten in dem jeweiligen System (hier: System BAZ 1) zusammenspielen und deren Lebensdauern so unterschiedlich verteilt sein könnten, dass auch sehr viel größere Wertesprünge zwischen den einzelnen Datenpunkten auftreten könnten. Für dieses System kann jedoch von einem weitestgehend linearen Zusammenhang zwischen den Ausfällen und damit auch einer linearen Verteilungsfunktion $f(x) = 8,24 \cdot 10^{-6} x - 0,07$ ausgegangen werden.

5.2 Durchführung von Szenario 1 für ein Bearbeitungszentrum

Das erste Szenario, welches anhand von realen Daten des Systems BAZ 1 untersucht werden soll, zeigt die Auswirkungen des Einsatzes eines Datensammlers von Betriebsbeginn an, also $t_{data} = 0 h$, auf. Dieses Szenario soll die grundsätzliche Vorgehensweise bei der Umsetzung von voraussagender Instandhaltung demonstrieren. Daher wird davon ausgegangen, dass bei Beginn der Betriebsphase $t_{data} = 0 h$ eines Systems, also nach den Phasen der Planung und Inbetriebnahme (siehe Kapitel 2.1), direkt Daten erfasst, analysiert und für die Vorhersage genutzt werden.

5.2.1 Bestimmung der erlernbaren Ausfälle

Die wahren Ausfälle der Menge P zeigten bereits, dass Komponenten mehrfach ausfallen können. Somit ist es sinnvoll, Ausfälle zu beobachten und deren Wiederauftreten rechtzeitig vorherzusehen und dann entsprechend abzustellen. Dabei stellt sich zunächst die Frage, wie viele Ausfälle sind überhaupt beobachtbar und somit erlernbar. Auf die Beantwortung dieser ersten Fragestellung zielt im Folgenden die Bestimmung der Lernrate LPR .

Die Erhebung der Daten im Szenario 1 erfolgt von Betriebsbeginn an, also $t_{data} = 0 h$. Dadurch ist es bei idealisierter Betrachtung theoretisch sofort möglich mit dem Training zu beginnen. Zu welchem Zeitpunkt der Ausfall einer Komponente erlernt ist, ist abhängig von der Anzahl der dafür im Vorfeld notwendigerweise aufzutretenden Ausfälle s_K . Der datengetriebene Ansatz zur Erstellung eines Vorhersagemodells macht es notwendig, zunächst einen Ausfall samt seinen entsprechenden Symptomen wenigstens einmalig zu erfassen, bevor dieser vorhergesagt werden kann (d.h. erst bei Vorliegen von bekannten Ausfällen kann ein Klassifikator trainiert werden und nachfolgend nutzbringend angewandt werden). In dieser Arbeit wird die vereinfachende Annahme getroffen, dass ein Ausfall $s_K = 1$ ausreicht, um den Ausfall einer Komponente zu erlernen und künftig zu vermeiden. Diese Annahme lässt sich wie folgt begründen: Zum einen stellt das im Rahmen dieser Arbeit entwickelte und vorgestellte Konzept allein schon einen wissenschaftlichen Mehrwert dar und kann dann nachfolgend weiter an spezifische Gegebenheiten wie zwei unterschiedliche Ausfalltypen oder -ursachen auf triviale Weise (Bsp. Einführung von L_{Ka} für Ausfalltyp a einer Komponente und analog dazu L_{Kb} für Ausfalltyp b) angepasst werden. Zum anderen werden bereits die Ausfälle einzelner Komponenten eines komplexeren Gesamtsystems individuell betrachtet. Daher ist aufgrund der empirischen Erfahrung davon auszugehen, dass die einzelnen Komponenten bereits so hinreichend simpel sind, sodass die Ausfallursachen einer Komponente in der Regel ähnlich sind und sich durch ähnliche Symptome ankündigen. Zuletzt bleibt anzuführen, dass die Zielsetzung dieser Arbeit die Ableitung einer oberen Schranke für die bestmögliche Güte jeglicher Art von Vorhersagemodell ist. Dabei stellt die Annahme, dass nur ein Ausfall für das Erlernen notwendig ist, die höchstmögliche obere Schranke dar, während bei mehr als einem für das Erlernen notwendigen Ausfall die obere Schranke von sich aus niedriger liegen würde.

Die Vorgehensweise zur Berechnung der erlernbaren Ausfälle LP wurde bereits in Formel 4.13 definiert. Das Struktogramm in Bild 5.3 illustriert die Umsetzung. Laut diesem sind neben der Lebensdauer einer Komponente L_K und der Betriebsdauer B , auch die Anzahl des gleichen Ausfalls bis zum Erlernen des Ausfalls s_K anzugeben. In diesem Abschnitt werden die Gütemaße für das System BAZ 1 bestimmt, wobei die Lebensdauern $L_{BAZ 1}$ aus dem vorhergehenden Abschnitt verwendet werden. Ebenso wird eine Betriebsdauer $B = 15 \text{ Jahre}$ bzw. 131.400 Stunden angenommen. Bereits im vorherigen Abschnitt motiviert, wird von einer Anzahl an Ausfällen bis zum Erlernen eines Ausfalls $s_K = 1$ ausgegangen.

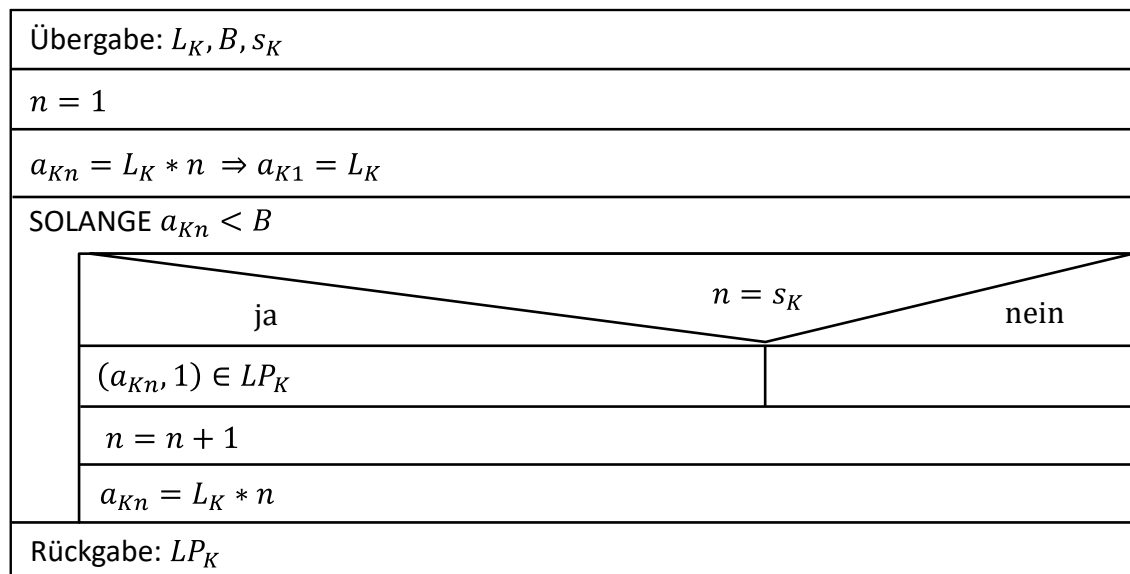


Bild 5.3: Struktogramm für die Berechnung der Menge der erlernbaren Ausfälle LP entsprechend Szenario 1: Dargestellt ist der einzelne Ausfall a indiziert durch die Anzahl n der bisherigen Ausfälle der Komponente K . In der Schleife wird entschieden, ob der untersuchte Ausfall a_{Kn} kleiner der Betriebsdauer B ist. Ist die Bedingung erfüllt, wird in der Verzweigung geprüft, ob die Anzahl der Ausfälle einer Komponente n mit der Anzahl zu erlernenden Ausfällen s_K identisch ist. Ist die Entscheidung „ja“, wird der Ausfall a_{Kn} zur Menge der erlernbaren Ausfälle LP_K hinzugefügt. Sobald die Bedingung $a_{Kn} < B$ nicht mehr erfüllt ist, wird die Schleife beendet und die Menge der erlernbaren Ausfälle LP_K ausgegeben.

Entsprechend der Übergabewerte ergibt sich bei einer Durchführung der Berechnung für alle Komponente des System BAZ 1 die Menge der erlernbaren Ausfälle $LP_{BAZ 1}$. Entsprechend dieser werden drei Ausfälle zu 10.000 Stunden, ein Ausfall zu 15.000 Stunden, ein Ausfall zu 18.000 Stunden, fünf Ausfälle zu 20.000 Stunden, ein Ausfall zu 25.000 Stunden, drei Ausfälle zu 30.000 Stunden und ein Ausfall zu 100.000 Stunden erlernt. Bereits an diesen Werten ist erkennbar, dass die Dauer bis Ausfälle für die voraussagende Instandhaltung gelernt worden sind, mehrere Jahre betragen kann.

Erg. 5.4:
$$LP_{BAZ 1}(t) = \{(10.000; 3), (15.000; 1), (18.000; 1), (20.000; 5), (25.000; 1), (30.000; 3), (100.000; 1)\}$$

Das System BAZ 1 hat insgesamt 15 Komponenten $|K| = 15$, die auf Grund ihrer mittleren Lebensdauer in der Betriebsdauer B ausfallen und deren Ausfälle somit im Bestfall erlernbar wären. Mit Hilfe dieser Information lässt sich nun auch die Lernrate $LPR(t) = LP(0 \leq t)/|K|$ (Formel 4.25) über der Zeit t berechnen. Eine Modellierung des Gütemaßes Lernrate $LPR_{BAZ 1}(t)$ über der Zeit t wird in Bild 5.4 vorgenommen, allerdings ist die Anzahl der Datenpunkte gering, wodurch die Näherungsfunktion auf einen passenden Definitionsbereich D_f eingeschränkt wurde. Grundsätzlich ist dennoch mit Hilfe der Datenpunkte und empirischen Verteilung aus Bild 5.4 zu erkennen, dass zwar nahezu alle Ausfälle bei System BAZ 1 im ersten Viertel des Betriebszeitraums t_B erlernt werden sollten. Eine weitere Analyse der empirischen Verteilungen zeigt aber auch, dass nach gut einem Jahr, also ~ 10.000 Stunden, erst rund 20 % der Ausfälle erlernt worden sind. Dies offenbart, dass für

die Erzeugung der für eine hohe Vorhersagegüte notwendigen Datenbasis ein mehrjähriger Zeitraum notwendig ist.

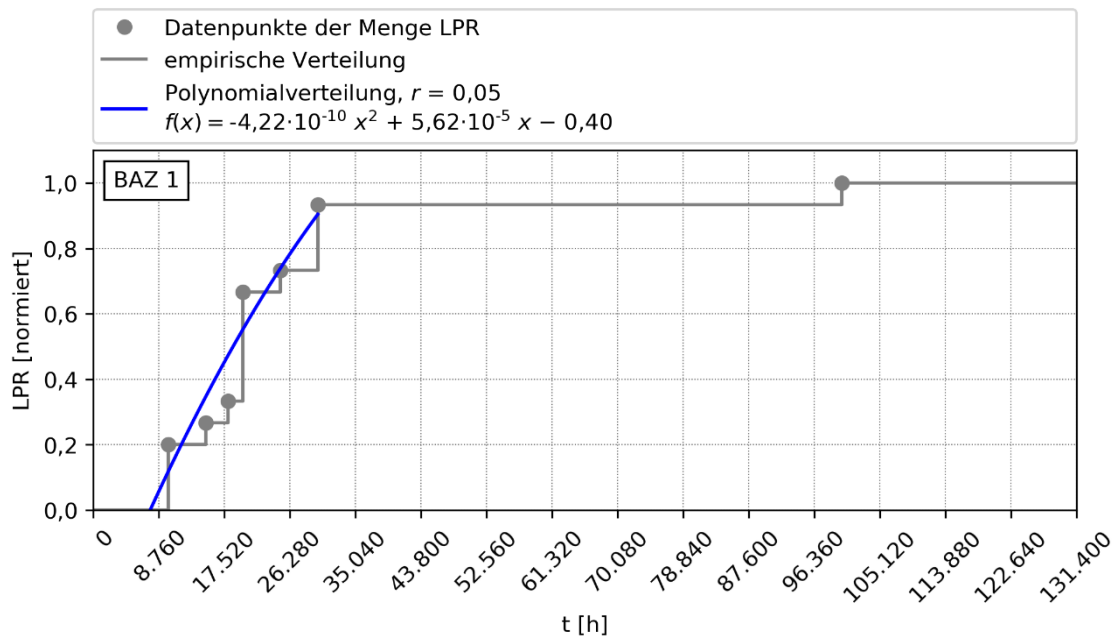


Bild 5.4: Normierte Datenpunkte, empirische und genäherte Verteilung der Lernrate LPR des Systems BAZ 1 für erlernbare Ausfälle entsprechend Szenario 1: Dargestellt sind die Datenpunkte, die normierte empirische Verteilung (in Grau) sowie die zugehörige Näherungsfunktion (in Blau) für die Lernrate LPR des Systems BAZ 1 über den Betriebszeitraum t_B in Stunden h . Die Verteilungen wurden mittels der Gesamtanzahl der Komponenten für BAZ 1 $|K| = 15$ normiert. Für die Näherungsfunktion wurde ein Polynom 2. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400[h$.

Wie bereits eingeführt, werden dauerhaft Daten erfasst und es besteht die Möglichkeit basierend auf diesen Daten zu lernen. Allerdings gibt es hierbei eine sinnvolle Grenze, denn ab der Hälfte der Betriebsdauer B können zwar noch neu zu erlernende Ausfälle eintreten, aber diese sollten – idealisiert betrachtet – nicht mehr innerhalb des verbleibenden Betriebszeitraums t_B ein zweites Mal auftreten. Somit kann eine maximale und anwendbare Lernrate LPR zum Zeitpunkt $t = 65.700 h$ angenommen werden. Daraus ergibt sich ein repräsentatives Ergebnis von rund 93 %⁴¹ erlernbaren und theoretisch wiederauftretenden Ausfällen basierend auf der empirischen Verteilung. Dies bedeutet genauer, dass nach der halben Betriebsdauer B von siebeneinhalb Jahren ($t = 65.700 h$) 93 % aller möglichen Ausfälle entsprechend der genäherten Verteilung beobachtet wurden. Im besten Fall kann dann für 93 % der beobachteten Ausfälle ein Vorhersagemodell abgeleitet werden, welches jeden weiteren Ausfall der bekannten Art zuverlässig und rechtzeitig vorhersagt. Man spricht in diesem Zusammenhang auch von den Treffern TP bzw. der Trefferrate TPR .

⁴¹ 14 von 15 Komponenten sind wenigstens einmal ausgefallen und damit erlernbar.

5.2.2 Bestimmung der vermeidbaren Ausfälle

Die Trefferrate (True-Positive-Rate, kurz TPR) beschreibt die Anzahl der richtigen Vorhersagen zu den wahren Ausfällen P . Der erste Ausfall bzw. die ersten Ausfälle einer Art können auf Grund des datengetriebenen Ansatzes nie vorhergesagt werden und dienen dazu den Ausfalltyp zu erlernen, wie im vorherigen Abschnitt 5.2.1 gezeigt wurde. Darauf basierend können danach alle weiteren Ausfälle im hier angenommenen Idealfall vermieden werden.

Unter der getroffenen Annahme ein Ausfall muss nur einmal auftreten ($s_K = 1$), damit dieser künftig vermieden werden kann, den Lebensdauern L_K und der Betriebsdauer B , wird nun die Menge der Treffer TP bei einer voraussagenden Instandhaltung berechnet. Zur Veranschaulichung wird ebenso ein Struktogramm, dargestellt in Bild 5.5 und basierend auf Formel 4.14, erstellt. Im Gegensatz zu der Menge der erlernbaren Ausfälle LP wird ein Ausfall a erst der Menge der Treffer TP zugeordnet, wenn die Anzahl der Ausfälle n größer der notwendigen Anzahl an Ausfällen s_K ist und somit der Ausfall a_{Kn} , also bereits erlernt wurde.

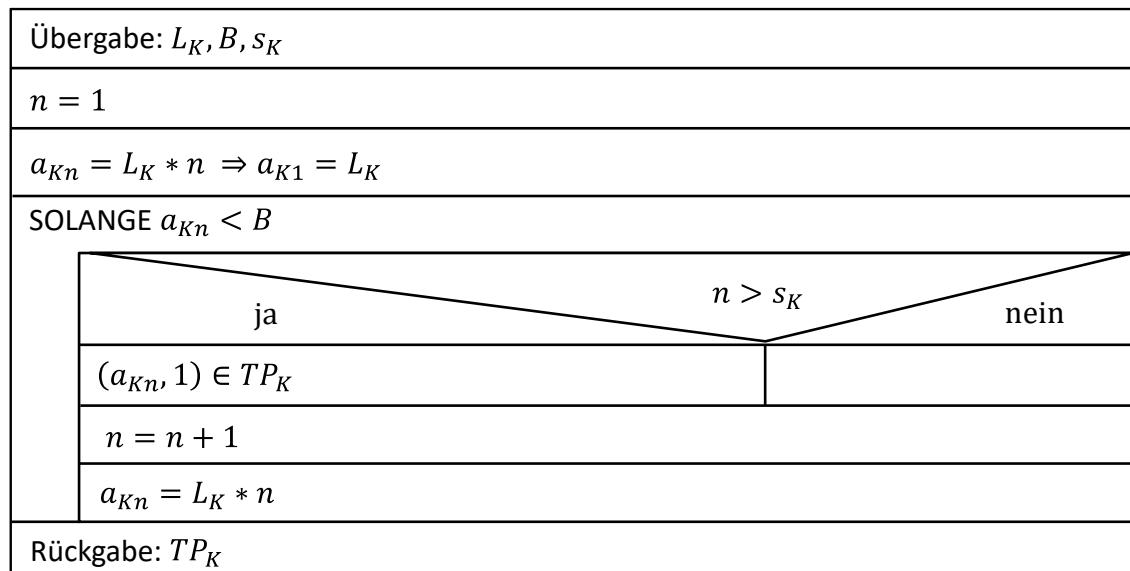


Bild 5.5: Struktogramm für die Berechnung der Menge der vermeidbaren Ausfälle TP entsprechend Szenario 1: Dargestellt ist der einzelne Ausfall a indiziert durch die Anzahl n der bisherigen Ausfälle der Komponente K . In der Schleife wird entschieden, ob der untersuchte Ausfall a_{Kn} kleiner der Betriebsdauer B ist. Ist die Bedingung erfüllt, wird in der Verzweigung geprüft, ob die Anzahl der Ausfälle einer Komponente n größer der Anzahl zu erlernenden Ausfällen s_K ist. Ist die Entscheidung „ja“, wird der Ausfall a_{Kn} zur Menge der vermeidbaren Ausfälle TP_K hinzugefügt. Sobald die Bedingung $a_{Kn} < B$ nicht mehr erfüllt ist, wird die Schleife beendet und die Menge der vermeidbaren Ausfälle TP_K ausgegeben.

Für das System BAZ 1 ergeben sich, entsprechend den Übergabewerten, 87 Treffer ($TPR_{BAZ\ 1}(0 \leq t < B) = 87$), welche im Ergebnis 5.5 aufgelistet sind.

$$TP_{BAZ\ 1}(t) = \{(20.000; 3), (30.000; 4), (36.000; 1), (40.000; 8), (45.000; 1), (50.000; 4), (54.000; 1), (60.000; 12), (70.000; 3),$$

Erg. 5.5: (72.000; 1), (75.000; 2), (80.000; 8), (90.000; 8), (100.000; 9), (105.000; 1), (108.000; 1), (110.000; 3), (120.000; 12), (125.000; 1), (126.000; 1), (130.000; 3)\}

Nach Formel 4.22 kann nun auch die Trefferrate zu jedem Zeitpunkt $TPR(t) = TP(0 \leq t) / P(0 \leq t < B)$ bestimmt werden, woraus sich auch eine grafische Darstellung wie Bild 5.6 ableiten lässt. Die Anzahl der Datenpunkte der Menge TPR ist hinreichend groß genug und entsprechend gut verteilt, sodass zusätzlich zur empirischen Verteilung eine adäquate Näherungsverteilung bestimmt werden kann. Entsprechend der Darstellung in Bild 5.6 kann über die komplette Betriebsdauer B von insgesamt 15 Jahren bzw. 131.400 Stunden eine maximale Trefferrate TPR von 85 % ausgehend von der empirischen Verteilung für das System BAZ 1 im Szenario 1 erreicht werden. Das heißt, bei z.B. 100 Ausfällen P über den gesamten Zeitraum würden rund 85 Ausfälle richtig vorhergesagt werden. Die Ergebnisse der Trefferrate TPR zeigen jedoch auch, dass nicht alle Ausfälle vermeidbar sind.

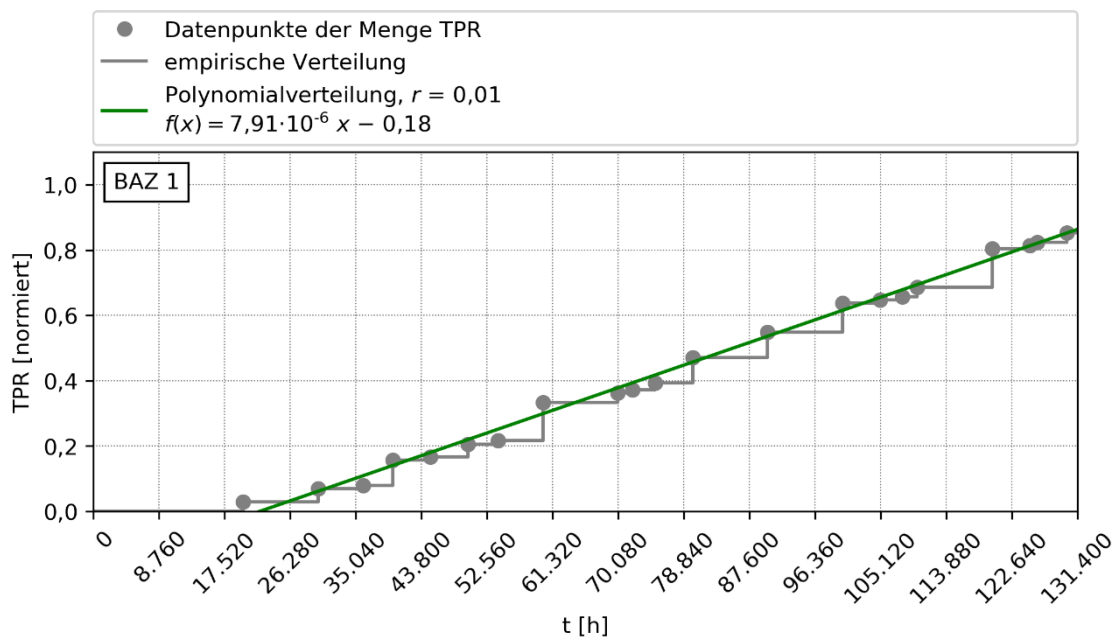


Bild 5.6: Normierte Datenpunkte, empirische und genäherte Verteilung der Trefferrate TPR des Systems BAZ 1 für vermeidbare Ausfälle entsprechend Szenario 1: Dargestellt sind die Datenpunkte, die normierte empirische Verteilung (in Grau) sowie die zugehörige Näherungsfunktion (in Grün) für die Trefferrate TPR des Systems BAZ 1 über den Betriebszeitraum t_B in Stunden h . Die Verteilungen wurden mittels der Gesamtanzahl der Ausfälle für BAZ 1 $P_{BAZ\ 1}(0 \leq t < B) = 102$ normiert. Für die Näherungsfunktion wurde ein Polynom 1. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400[h$.

5.2.3 Bestimmung der verbleibenden Ausfälle

Die vorherigen Abschnitte haben gezeigt, dass durch die Einführung einer voraussagenden Instandhaltung Ausfälle vermieden werden können, allerdings dauert es auch eine gewisse Zeit, bis mindestens ein Ausfall für jede Komponente aufgetreten ist und damit der Ausfall aller Komponenten erlernt ist. Durch den notwendigen Schritt, Ausfälle erst erlernen zu müssen, fehlt für die ersten Ausfälle ein Alarm. Somit kann die Anzahl der Ausfälle auch bei einer idealen Umsetzung einer voraussagenden Instandhaltung nicht vollständig auf null reduziert werden. Eine ideale Trefferrate $TPR(t = B) = 1$, also 100 % zum Betriebsende B , kann also nicht erzielt werden.

Eine Ausnahme ist lediglich, wenn das Training bereits im Vorfeld durch Verwendung von Ausfalldaten eines vergleichbaren Systems stattfinden konnte und dadurch der passende und bereits vollständig trainierte Klassifikator übertragen werden kann. Die Vorgehensweise hierzu ist in Abschnitt 3.2.2 erklärt. Eine solche Übertragung von Erkenntnissen zwischen mindestens zwei Systemen wird im Szenario 3 des Kapitels 5.4 diskutiert und die entsprechenden Ergebnisse werden dort vorgestellt.

Die Untersuchung der Nicht-Alarmrate (False-Negativ-Rate, kurz FNR) für das Szenario 1 des Systems BAZ 1 soll zeigen, wie viele Ausfälle noch über die Betriebsphase auftreten würden, wenn ein Vorhersagemodell während des Betriebszeitraumes t_B fortlaufend trainiert und kontinuierlich angewandt werden würde. Zunächst wurden hierzu die Datenpunkte der Menge der verbleibenden Ausfälle FN nach Formel 4.15 bestimmt und es ergibt sich:

Erg. 5.6:
$$FN_{BAZ\ 1}(t) = \{(10.000; 3), (15.000; 1), (18.000; 1), (20.000; 5), \\ (25.000; 1), (30.000; 3), (100.000; 1)\}$$

Eine Umsetzung der Formel 4.15 erfolgte entsprechend dem Struktogramm in Bild 5.7. Im Vergleich zu den vorhergehenden Struktogrammen wird ein Ausfall a_{Kn} nur zur Menge der verbleibenden Ausfälle FN hinzugefügt, wenn der Ausfall weniger oder gleich oft aufgetreten n ist wie die Anzahl der notwendigen Ausfälle zum Erlernen eines Ausfalls s_K . Für die Bestimmung in diesem Abschnitt wurde, wie auch in den anderen Abschnitten des Kapitels $s_K = 1$ angenommen, wonach sich auch Ergebnis 5.6. ergibt.

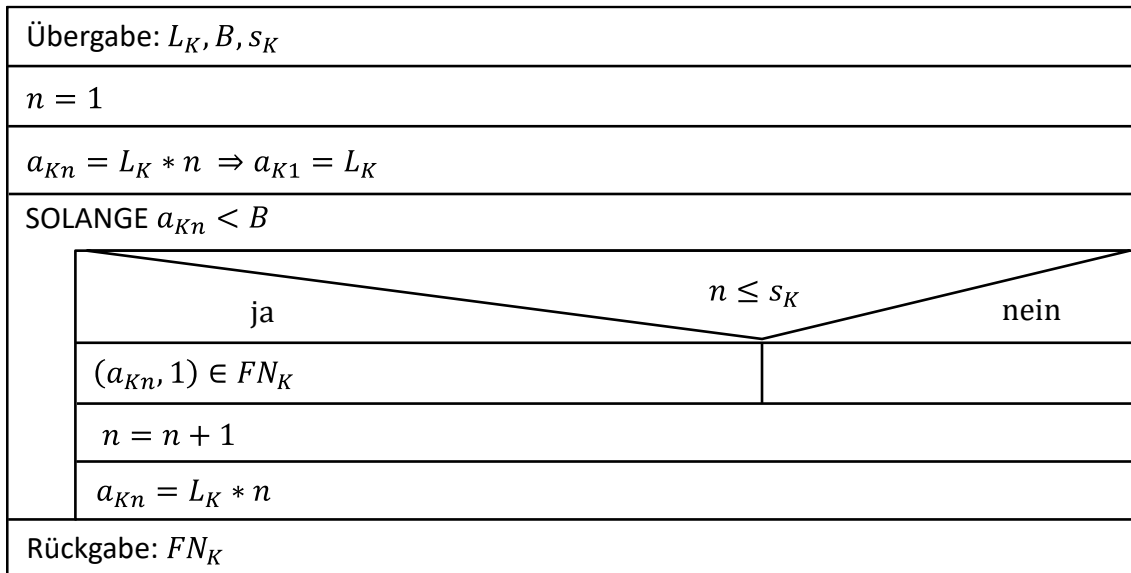


Bild 5.7: Struktogramm für die Berechnung der Menge der verbleibenden Ausfälle FN entsprechend Szenario 1: Dargestellt ist der einzelne Ausfall a indiziert durch die Anzahl n der bisherigen Ausfälle der Komponente K . In der Schleife wird entschieden, ob der untersuchte Ausfall a_{Kn} kleiner der Betriebsdauer B ist. Ist die Bedingung erfüllt, wird in der Verzweigung geprüft, ob die Anzahl der Ausfälle einer Komponente n kleiner oder gleich der Anzahl zu erlernenden Ausfällen s_K ist. Ist die Entscheidung „ja“, wird der Ausfall a_{Kn} zur Menge der verbleibenden Ausfälle FN_K hinzugefügt. Sobald die Bedingung $a_{Kn} < B$ nicht mehr erfüllt ist, wird die Schleife beendet und die Menge der verbleibenden Ausfälle FN_K ausgegeben.

Wie auch in den vorherigen Abschnitten wurde das Ergebnis für die Nicht-Alarmrate $FNR(t)$ grafisch in Bild 5.8 dargestellt und zeigt neben den Datenpunkten, auch die empirische Verteilung und die Näherungsverteilung 1. Grades. Aus Bild 5.8 lässt sich eine maximale Nicht-Alarmrate FNR von rund 15 % ablesen. Dies bedeutet, wenn z.B. 100 wahre Ausfälle P über den gesamten Betriebszeitraum von 15 Jahren bei rein reaktiver Wartung aufgetreten wären, so würden unter Verwendung des oben beschriebenen Vorhersagemodells bzw. im Rahmen einer voraussagenden Instandhaltung jetzt nur noch 15 Ausfälle ($FNR(t = B) = 15$) auftreten. Dies ist eine deutliche Verbesserung zu vorher und demonstriert den Mehrwert von voraussagender Wartung in der Produktion erstmalig.

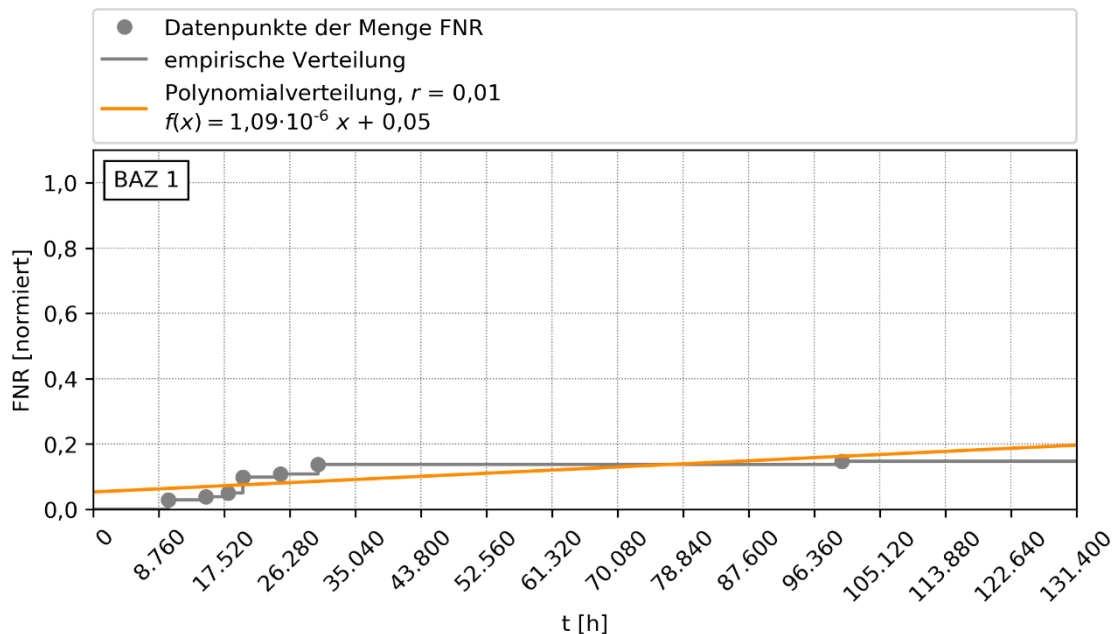


Bild 5.8: Normierte Datenpunkte, empirische und genäherte Verteilung der Nicht-Alarmrate FNR des Systems BAZ 1 für verbleibende Ausfälle entsprechend Szenario 1: Dargestellt sind die Datenpunkte, die normierte empirische Verteilung (in Grau) sowie die zugehörige Näherungsfunktion (in Orange) für die Nicht-Alarmrate FNR des Systems BAZ 1 den Betriebszeitraum t_B in Stunden h . Die Verteilungen wurden mittels der Gesamtanzahl der Ausfälle für BAZ 1 $P_{BAZ\ 1}(0 \leq t < B) = 102$ normiert. Für die Näherungsfunktion wurde ein Polynom 1. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400[h$.

Der Einsatz eines Vorhersagemodells hat folglich neben der Vermeidung von Ausfällen TP durch Vorhersage zusätzlich den Effekt, dass weitere verbleibende Ausfälle FN mit fortschreitender Betriebszeit eher weniger werden. So treten bereits nach 8 Jahren nur noch sehr wenig neue Ausfälle auf. Betrachtet man beispielhaft ein System mit insgesamt 100 Ausfällen, dann würde nach 8 Jahren nur noch ein weiterer Ausfall anstehen. Dies liegt vor allem an der Lebensdauer der Komponenten und der daraus resultierenden Lernrate LPR . Wie bereits Bild 5.4 zeigte, sind bei den erlernbaren Ausfällen bereits ein Großteil der Komponentenausfälle nach 8 Jahren bzw. 70.080 Stunden erlernt worden.

5.3 Durchführung von Szenario 2 für ein Bearbeitungszentrum

In diesem zweiten Szenario wird die Möglichkeit betrachtet, dass die Erfassung und Auswertung von Daten erst nach Beginn des Betriebs $t_{data} > 0 h$ startet. Zur Veranschaulichung wird in dieser Arbeit eine Untersuchung durchgeführt, bei der die Datenerfassung bei 35.040 Stunden bzw. 4 Jahren beginnt.

$$t_{data} = 35.040 h$$

Das Ende der Datenerfassung ist weiterhin durch die festgelegte Betriebsdauer B von 15 Jahren bzw. 131.400 Stunden charakterisiert.

Durch einen späteren Beginn der Datenerfassung werden einige erstmalig auftretende Ausfälle bereits verpasst. Folglich ist ein Offset o für die jeweilige Anzahl an

notwendigen Ausfällen einer Komponente K einzuführen, um dieses bei der Bestimmung der Gütemaße zu berücksichtigen. Die Bestimmung des Offsets o_K erfolgt entsprechend Formel 4.8, wobei die methodische Umsetzung in einem Struktogramm (siehe Bild 5.9) dargestellt wird.

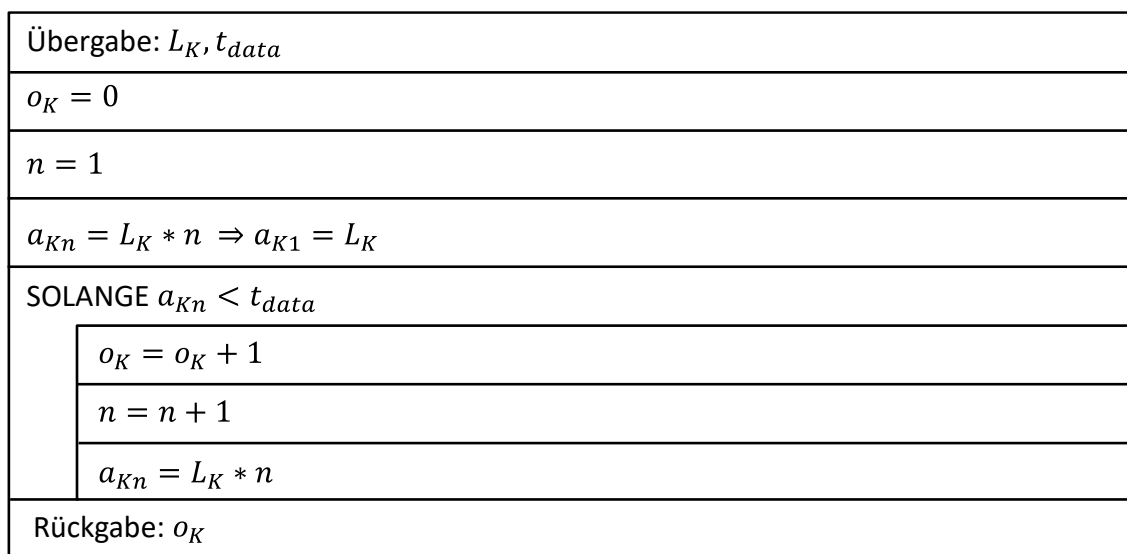


Bild 5.9: Struktogramm für die Berechnung eines Offsets bei späterem Beginn der Datenerfassung: Dargestellt ist der einzelne Ausfall a indiziert durch die Anzahl n der bisherigen Ausfälle der Komponente K . In der Schleife wird entschieden, ob der untersuchte Ausfall a_{Kn} kleiner dem Beginn der Datenerfassung t_{data} ist. Ist die Bedingung erfüllt, wird der Zähler o_K um eins erhöht. Sobald die Bedingung $a_{Kn} < t_{data}$ nicht mehr erfüllt ist, wird die Schleife beendet und die Anzahl der verpassten Ausfälle o_K ausgegeben.

Ergänzend ist festzuhalten, dass im vorherigen Abschnitt 5.2 auf die Verwendung des Offsets o_K zur Vereinfachung verzichtet wurde, obwohl die Formeln 4.12 bis 4.15 bereits diesen berücksichtigt haben. Daher wird nun zusammenfassend für die vier Mengen der wahren Ausfälle P , der erlernbaren Ausfälle LP , der vermeidbaren Ausfälle TP und der verbleibenden Ausfälle FN ein aktualisiertes Struktogramm in Bild 5.10 aufgeführt, wobei die Aktualisierung keine Auswirkungen auf die Menge der wahren Ausfälle P hat und somit diese für das Szenario 2 nicht neu bestimmt werden muss.

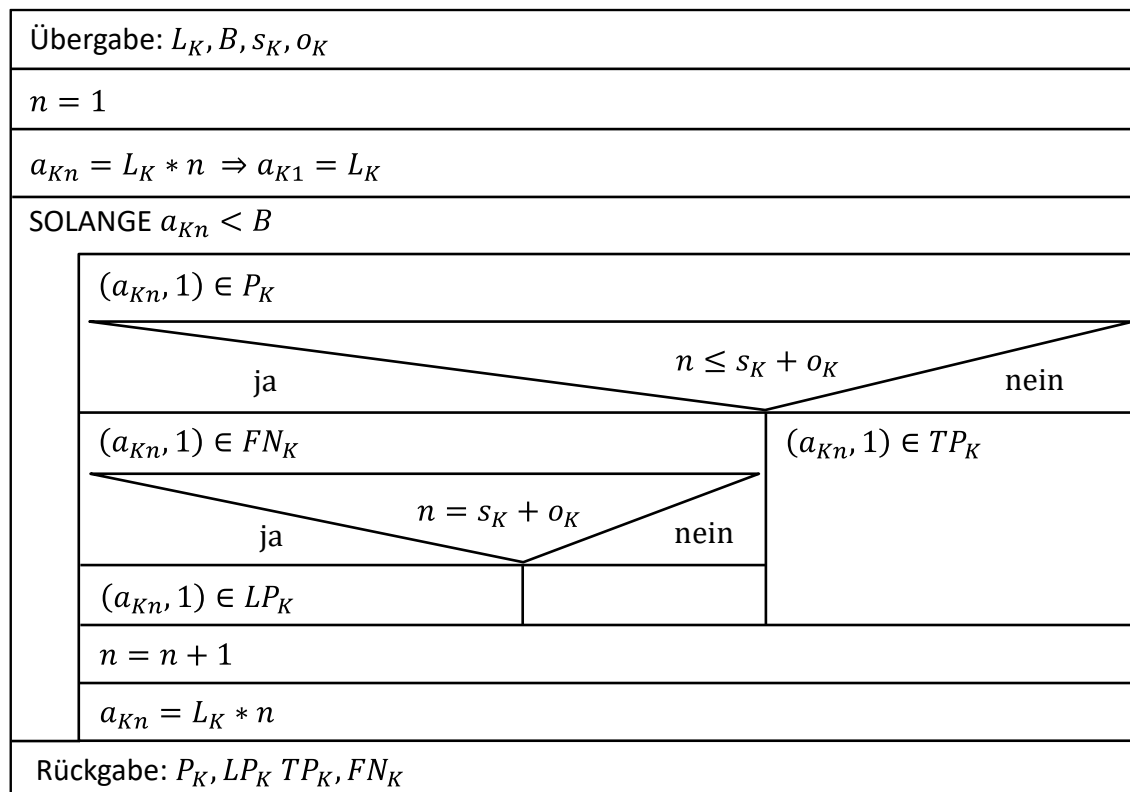


Bild 5.10: Generelles Struktogramm für die Berechnung der Mengen P, LP, TP, FN : Dargestellt ist der einzelne Ausfall a indiziert durch die Anzahl n der bisherigen Ausfälle der Komponente K . In der Schleife wird entschieden, ob der untersuchte Ausfall a_{Kn} kleiner der Betriebsdauer B ist. Ist die Bedingung erfüllt, wird der Ausfall a_{Kn} zur Menge der wahren Ausfälle P_K hinzugefügt. Anschließend wird in der ersten Verzweigung geprüft, ob die Anzahl der Ausfälle einer Komponente n kleiner oder gleich der Anzahl zu erlernenden Ausfällen s_K plus den Offset der Anzahl zu erlernenden Ausfällen o_K ist. Ist die Entscheidung „ja“, wird der Ausfall a_{Kn} zur Menge der verbleibenden Ausfälle FN_K hinzugefügt und in einer zweiten Verzweigung wird geprüft, ob die Anzahl der Ausfälle einer Komponente n identisch der Anzahl zu erlernenden Ausfällen s_K plus den Offset der Anzahl zu erlernenden Ausfällen o_K ist. Ist die Entscheidung der zweiten Verzweigung „ja“, wird der Ausfall a_{Kn} zur Menge der erlernbaren Ausfälle LP_K hinzugefügt. Bei „nein“ wird dieser nicht weiter berücksichtigt. Ist bei der ersten Verzweigung die Entscheidung „nein“, wird der Ausfall a_{Kn} zur Menge der vermeidbaren Ausfälle TP_K hinzugefügt. Sobald die Bedingung $a_{Kn} < B$ nicht mehr erfüllt ist, wird die Schleife beendet und die Menge der vermeidbaren Ausfälle TP_K ausgegeben.

Entsprechend dem Struktogramm in Bild 5.10 wurde nun für das System BAZ 1 die Menge der erlernbaren Ausfälle $LP_{BAZ 1}$ bestimmt.

Erg. 5.7:
$$LP_{BAZ 1}(t) = \{(36.000; 1), (40.000; 8), (45.000; 1), (50.000; 1), (60.000; 3), (100.000; 1)\}$$

Eine grafische Veranschaulichung des Ergebnisses $LPR_{BAZ 1}(t)$ inklusive der empirischen und angenäherten Verteilung ist in Bild 5.11 erfolgt. Die angenäherte Verteilung – eine Polynomialverteilung – wurde nur im geeigneten Definitionsbereich D_f dargestellt. Weiterhin ist deutlich erkennbar, dass die Ausfälle bereits schnell am Anfang erlernt werden, wobei es ein Jahr $t = 43.800 h$ dauert, bis die Lernrate LPR bei 60 % liegt.

Erg. 5.8:
$$LPR_{BAZ\ 1}(t = 43.800\ h) = \frac{LP(0 \leq t) = 9}{|K| = 15} = 0,6$$

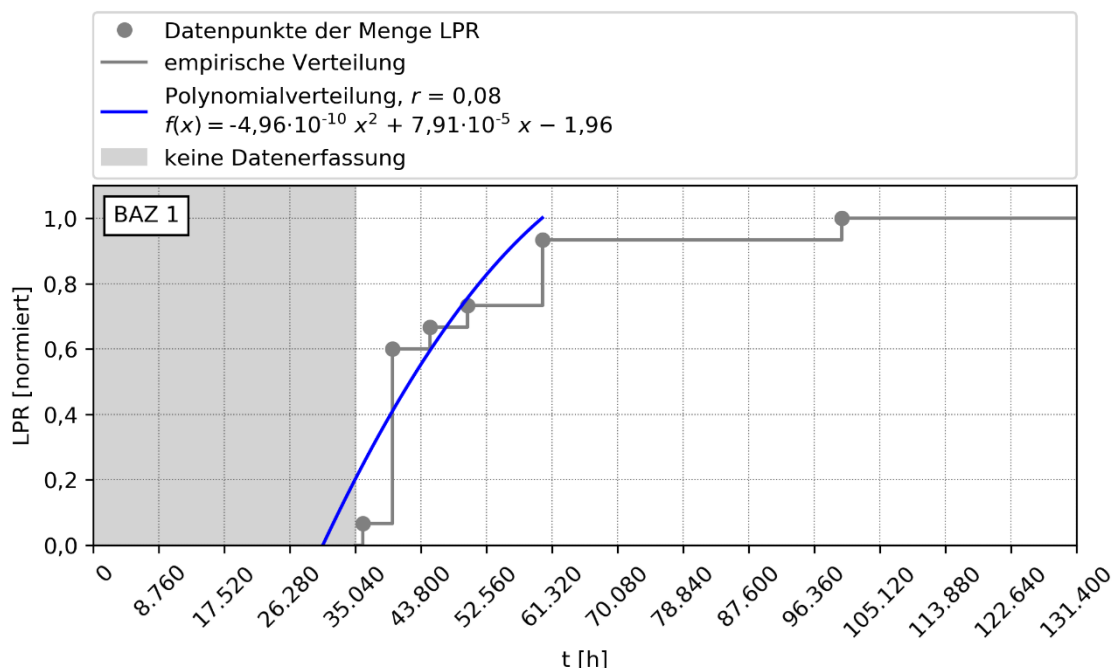


Bild 5.11: Normierte Datenpunkte, empirische und genäherte Verteilung der Lernrate LPR des Systems BAZ 1 für erlernbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb: Dargestellt sind die Datenpunkte, die normierte empirische Verteilung (in Grau) sowie die zugehörige Näherungsfunktion (in Blau) für die Lernrate LPR des Systems BAZ 1 über den Betriebszeitraum t_B in Stunden h . Die Verteilungen wurden mittels der Gesamtanzahl der Komponenten für BAZ 1 $|K| = 15$ normiert. Für die Näherungsfunktion wurde ein Polynom 2. Grades gewählt. Der ausgegraute Bereich stellt den Zeitraum $t = [0, 35.040[$ in Stunden h dar, in dem noch keine Datenerfassung erfolgt ist.

Die Lernrate $LPR_{BAZ\ 1}(t = B) = 1$ ist zwar die maximale Lernrate, allerdings umfasst diese auch erlernte Ausfälle, welche aufgrund des späteren Zeitpunkts des Erlernens kein zweites Mal mehr in der Betriebsdauer B auftreten werden. Dies sollte bei der Interpretation der Lernrate LPR berücksichtigt werden. Inwieweit die erlernten Ausfälle LP ausreichen, sodass die voraussagende Instandhaltung weiterhin einen Mehrwert bietet, zeigt daher die Auswertung der Treffer TP .

Erg. 5.9:
$$TP_{BAZ\ 1}(t) = \{(50.000; 3), (54.000; 1), (60.000; 9), (70.000; 3), (72.000; 1), (75.000; 2), (80.000; 8), (90.000; 8), (100.000; 9), (105.000; 1), (108.000; 1), (110.000; 3), (120.000; 12), (125.000; 1), (126.000; 1), (130.000; 3)\}$$

Aus Ergebnis 5.9. leitet sich ab, dass die Trefferrate TPR bei einem Beginn der Datenerfassung $t_{data} = 35.040\ h$ um 20 % auf 65 % im Vergleich zu Szenario 1 sinkt. Dieser Abstrich in Folge des späteren Einsatzes einer Datenerfassung bzw. in der Praxis durch die spätere Nachrüstung einer solchen Datenerfassung an einer Anlage konnte nun zum ersten Mal wissenschaftlich bestimmt und quantitativ angegeben werden.

Erg. 5.10:
$$TPR_{BAZ_1}(t = B) = \frac{TP(0 \leq t) = 66}{P(0 \leq t < B) = 102} = 0,65$$

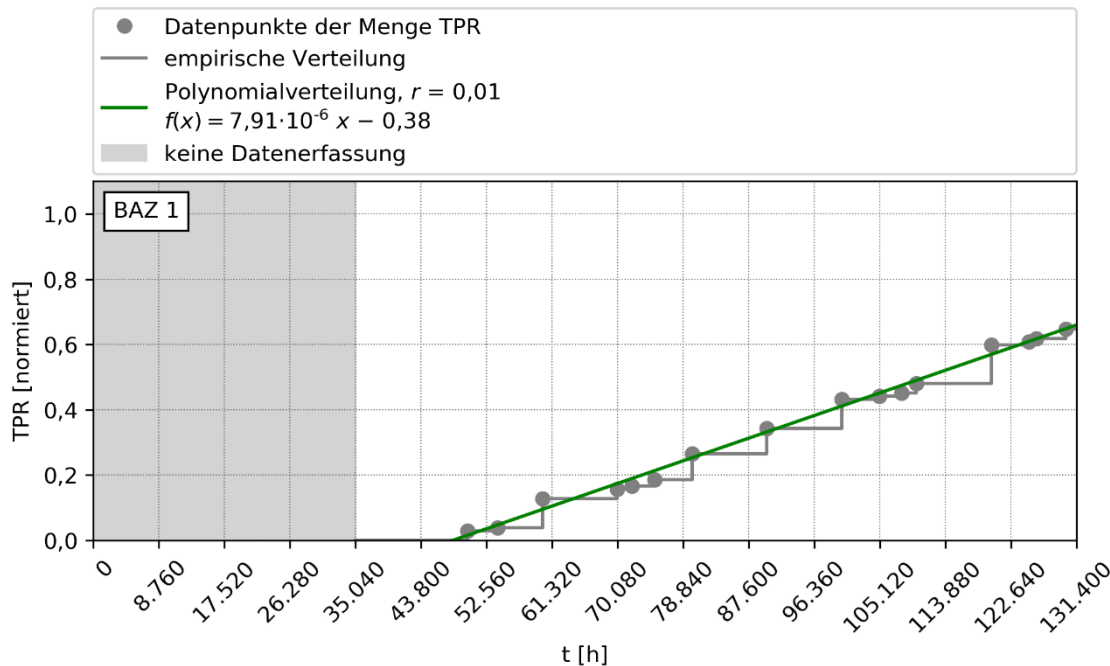


Bild 5.12: Normierte Datenpunkte, empirische und genäherte Verteilung der Trefferrate TPR des Systems BAZ 1 für vermeidbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb: Dargestellt sind die Datenpunkte, die normierte empirische Verteilung (in Grau) sowie die zugehörige Näherungsfunktion (in Grün) für die Trefferrate TPR des Systems BAZ 1 über den Betriebszeitraum t_B in Stunden h . Die Verteilungen wurden mittels der Gesamtanzahl der Ausfälle für BAZ 1 $P_{BAZ_1}(0 \leq t < B) = 102$ normiert. Für die Näherungsfunktion wurde ein Polynom 1. Grades gewählt. Der ausgegraute Bereich stellt den Zeitraum $t = [0, 35.040[$ in Stunden h dar, in dem noch keine Datenerfassung erfolgt ist.

Konträr zu den Ergebnissen der Trefferrate TPR verhält sich die Nicht-Alarmrate FNR . Diese steigt stark an, welches zunächst durch die Menge der verbleibenden Ausfälle FN wiedergegeben wird.

$$P_{BAZ_1}(t) = \{(10.000; 3), (15.000; 1), (18.000; 1), (20.000; 8),$$

Erg. 5.11: $(25.000; 1), (30.000; 7), (36.000; 1), (40.000; 8), (45.000; 1),$

$$(50.000; 1), (60.000; 3), (100.000; 1)\}$$

Durch den späteren Beginn der Datenerfassung t_{data} und damit bedingt durch die andere Lernrate LPR verändert sich die Verteilung für die verbleibenden Ausfälle FNR . Demzufolge verschlechtert sich unweigerlich und erwartungsgemäß die Nicht-Alarmrate FNR . Die Reduktion der Ausfälle für den Untersuchungszeitraum ab 4 Jahren bzw. 35.040 Stunden ist in Bild 5.13 dargestellt. Zur grafischen Demonstration wurde wie bisher neben den Datenpunkten auch die empirische und angenäherte Verteilung dargestellt. Hierbei fällt deutlich auf, dass die meisten Ausfälle in der Zeit t auftreten, bei der keine Daten erfasst wurden. Hierauf folgt dann noch ein kurzer Anstieg, bis dann nahezu alle Ausfälle relativ schnell erlernt wurden.

Auch hier wurde der Definitionsbereich D_f für die Polynomialverteilung passend beschränkt.

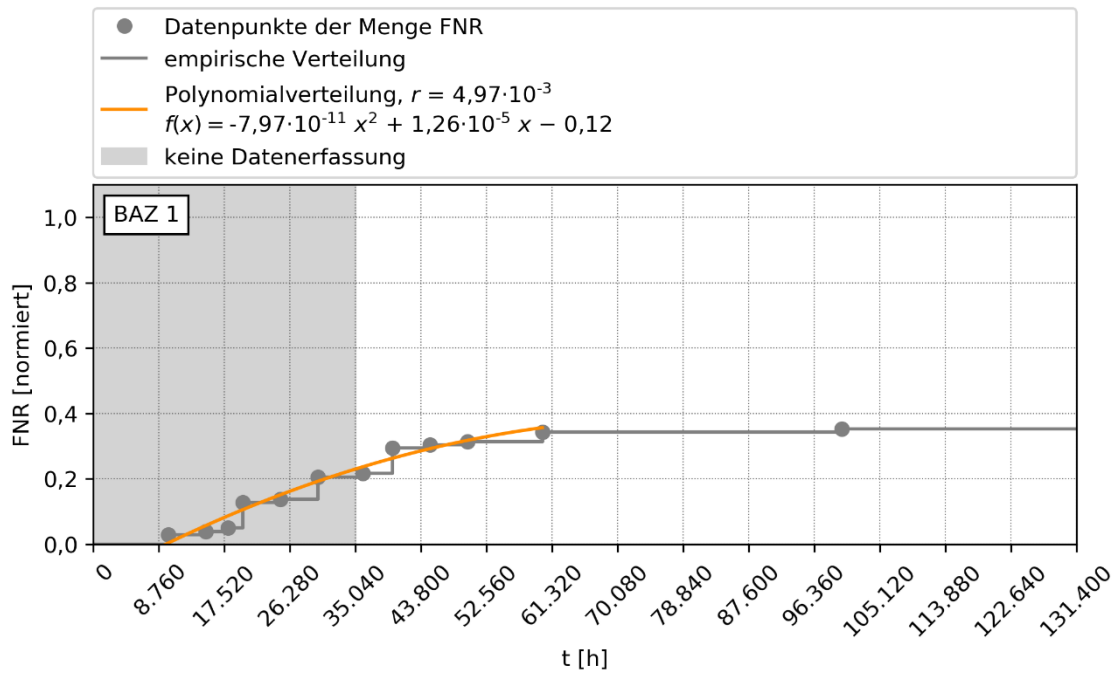


Bild 5.13: Normierte Datenpunkte, empirische und genäherte Verteilung der Nicht-Alarmrate FNR des Systems BAZ 1 für verbleibende Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb: Dargestellt sind die Datenpunkte, die normierte empirische Verteilung (in Grau) sowie die zugehörige Näherungsfunktion (in Orange) für die Nicht-Alarmrate FNR des Systems BAZ 1 über den Betriebszeitraum t_B in Stunden h . Die Verteilungen wurden mittels der Gesamtanzahl der Ausfälle für BAZ 1 $P_{BAZ 1}(0 \leq t < B) = 102$ normiert. Für die Näherungsfunktion wurde ein Polynom 2. Grades gewählt. Der ausgegraute Bereich stellt den Zeitraum $t = [0, 35.040[$ in Stunden h dar, in dem noch keine Datenerfassung erfolgt ist.

Zusammenfassend für Szenario 2 kann festgehalten werden, dass die Gütewerte deutlich schlechter ausfallen. Allerdings ist der Nutzen einer voraussagenden Instandhaltung für die angenommenen Parameter wie Startzeitpunkt der Datenerfassung t_{data} , Anzahl der für das Erlernen notwendiger Ausfälle s_K etc. noch gewährleistet. Für andere Systeme und vor allem andere Startzeitpunkte der Datenerfassung t_{data} mag dies aber nicht mehr zutreffen und muss dann anhand des in dieser Arbeit vorgestellten Verfahrens individuell betrachtet werden. In Kapitel 6 erfolgt noch eine Verallgemeinerung auf die Klasse der Produktionssysteme, sodass nur bei einem von dort abweichenden Startzeitpunkt der Datenaufnahme oder bei einer anderen Klasse an Systemen das Verfahren erneut angewandt werden muss. Andernfalls kann die Abschätzung der dortigen Gütemaße übertragend verwendet werden.

5.4 Anwendung der Methode auf ein baugleiches Bearbeitungszentrum

Eine andere Möglichkeit, um die Vorhersagegüte für Ausfälle von technischen Systemen zu verbessern, ist die Nutzung der Daten aus ähnlichen⁴² bzw. im besten Falle baugleichen und gleichgenutzten Maschinen. In der Produktion bedeutet dies, dass ein Transfer-Learning-Algorithmus grundsätzlich für baugleiche Maschinen wie Bearbeitungszentren angewendet werden kann, wobei deren Nutzung zusätzlich auch möglichst ähnlich sein sollte. Eine solche ähnliche Nutzung ist zwar vorteilhaft, aber in vielen Fällen nicht unbedingt zwingend erforderlich. Denn für das Transfer-Learning ist der Einsatz derselben Komponenten maßgeblich entscheidend und weniger deren unterschiedliche Beanspruchung. Die unterschiedlichen Einflussgrößen auf die Lebensdauer eines Systems sollten nämlich durch ein ausreichend gut gewähltes Machine-Learning-Verfahren ausgeglichen werden. Allerdings sind dennoch Grenzfälle denkbar, in denen Ausfälle aufgrund gewisser geringerer oder erhöhter oder allgemein andersartiger Beanspruchungen von Komponenten nur bei einzelnen Maschinen auftreten und nicht bei allen baugleichen Maschinen. Dies ist dann hinsichtlich der Übertragbarkeit von gelernten Ausfällen auf andere Maschine entsprechend zu berücksichtigen und kann gegebenenfalls negative Auswirkungen auf die Trefferrate TPR haben.

Das Szenario 3 untersucht die Vorteile des Transfer-Learning und somit die mögliche Verbesserung der Gütemaße Trefferrate TPR , Nicht-Alarmrate FNR aber auch der Lernrate LPR unter Betrachtung eines weiteren baugleichen Systems. Ein grundsätzliches Erfordernis ist es, dass ein zeitlicher Versatz zwischen den zwei baugleichen Systemen besteht, da ansonsten erstmalig auftretende Ausfälle an beiden Systemen (nahezu) gleichzeitig auftreten würden und das gelernte Erkennen von Ausfällen nicht rechtzeitig übertragen werden könnte. Somit sollte das erste System (System 1) um einen zeitlichen Versatz Δt früher den Betrieb aufnehmen und das zweite System wird erst nach diesem zeitlichen Versatz in Betrieb genommen, kann aber von Anfang direkt auf das von System 1 fortlaufend gelernte Vorhersagemodell zurückgreifen.

Bei einer idealen Betrachtung, also $s_K = 1$ für System 1 und sobald der Versatz der Maschinen $\Delta t \neq 0 h$ ist, würde das eine System (System 1 bei $\Delta t > 0 h$) immer zuerst die Ausfälle erlernen, sodass das andere System (System 2 im genannten Fall) voraussichtlich keine ungeplanten Ausfälle mehr hätte, da die entsprechenden Ausfälle im Vorfeld bereits mindestens einmal im ersten System aufgetreten sind, dadurch bereits erlernt wurden und nun korrekt vorhergesagt werden würden. Ein größerer zeitlicher Versatz Δt ist erst bei einer nicht-idealisierten Betrachtung relevant, da dadurch unterschiedliche Einflüsse auf die baugleichen Systeme und

⁴² Als Ähnlichkeit wird im Zusammenhang mit maschinellem Lernen die gleiche Anzahl an Features und deren gleiche Ausprägung bezeichnet. Erst dann kann ein Transfer-Learning, also das Übertragen der gelernten Ausfälle auf andere Systeme, angewendet werden. Die Funktionsweise von Transfer-Learning wurde in Abschnitt 3.2.2 vorgestellt.

insbesondere auf deren einzelnen Komponenten ausgeglichen werden können. Denn durch verschiedene Einflussfaktoren, welche in Abschnitt 2.1.3 beschrieben wurden, können Veränderung der Lebensdauern der Komponenten L_K entstehen, wodurch möglicherweise eine Komponente K aus einem System mit später startender Datenerfassung (System 2 in diesem Kapitel) stärker beansprucht wird und dadurch früher ausfällt oder weniger beansprucht wird und später ausfällt.

Sollten beide Systeme jedoch zeitgleich die Betriebsphase beginnen, also $\Delta t = 0 h$, so ist bei einer idealisierten Betrachtung kein Vorteil aus der gemeinsamen Nutzung des Vorhersagemodells zu erwarten und keine Veränderung der Lernrate LPR zu erzielen. In der Realität würde sich aber vermutlich die Vorhersagegenauigkeit durch die gemeinsame Nutzung des Vorhersagemodells dennoch erhöhen, da gleiche Ausfälle in den beiden baugleichen Maschinen, unter anderem aufgrund deren unterschiedlicher Nutzung, abweichende Merkmalsausprägungen aufweisen würden, die die Generalisierbarkeit des zugrunde liegenden Machine-Learning-Modells erhöhen würden. Hinzu kommt natürlich, dass in der Praxis Ausfälle nie genau zum gleichen Zeitpunkt auftreten würden. Somit wäre ein gleicher Ausfalltyp in einem baugleichen System eventuell bereits aufgetreten und in das Vorhersagemodell eingeflossen, bevor er kurz danach erst beim anderen System auftreten würde. Somit kann es also hilfreich sein, die Daten aus mehreren Systemen zu erfassen und anschließend für das Training eines Vorhersagemodells zu nutzen. Denn dies bietet die Möglichkeit unterschiedliche Ausfallursachen und auch deren Symptome, welche auf einen potenziellen Ausfall schließen lassen, zu erlernen.

Durch die Übertragung des Wissens von System 1 auf System 2 ist es also möglich, die Trefferrate TPR zu steigern und ungeplante Ausfälle P zu vermeiden, selbst wenn diese noch nie im System 2 aufgetreten sind. Die Trefferrate TPR ist wie gesagt fast unabhängig vom zeitlichen Versatz Δt , dieser muss aber wenigstens so groß sein, dass der Ausfall bereits in System 1 erlernt werden konnte und noch rechtzeitig auf System 2 übertragen werden konnte. In dieser Arbeit wird angenommen, dass ein Ausfall beim ersten Mal ($s_K = 1$) im System 1 erlernt wird und im System 2 kein Ausfall ($s_K = 0$) mehr für das Erlernen erforderlich ist. Dies ermöglicht die Betrachtung der bestmöglichen Steigerung der Trefferrate TPR durch Transfer-Learning. Die Ergebnisse für das System 2 BAZ 1 im Szenario 3 sind nachfolgend als Mengen der Lernrate LPR , Trefferrate TPR und Nicht-Alarmrate FNR beschrieben.

Erg. 5.12: $LPR_{BAZ\ 1}(t) = \{ \}$

$$TPR_{BAZ\ 1}(t) = \{(10.000; 3), (15.000; 1), (18.000; 1), (20.000; 8), \\ (25.000; 1), (30.000; 7), (36.000; 1), (40.000; 8), (45.000; 1), (50.000; 4),$$

Erg. 5.13: $(54.000; 1), (60.000; 12), (70.000; 3), (72.000; 1), (75.000; 2), (80.000; 8), \\ (90.000; 8), (100.000; 10), (105.000; 1), (108.000; 1), (110.000; 3), \\ (120.000; 12), (125.000; 1), (126.000; 1), (130.000; 3)\}$

Erg. 5.14: $FNR_{BAZ\ 1}(t) = \{ \}$

Die Lernrate LPR für das System 2 BAZ 1 ist eine leere Menge, da vereinfacht betrachtet keine unbekannt neuen Ausfälle mehr erlernt werden. Daher wird auch auf einen grafischen Vergleich zwischen System 1 und System 2 des BAZ 1 verzichtet. Die Trefferrate TPR hingegen wurde in Bild 5.14 für System 1 und System 2 vergleichend dargestellt. Die Bestimmung der mittleren Differenz zwischen den beiden Polynomialverteilungen für System 1 und System 2 des BAZ 1 ergibt eine mittlere Differenz von 13 %. Es ist kaum erkennbar, allerdings ist die Differenz am Anfang erst bei 11 % und steigert sich dann zum Betriebsende auf über 15 %. D.h. für das System 2 konnten folglich 15 % mehr Ausfälle als bei System 1 durch den Einsatz einer voraussagenden Instandhaltung vermieden werden. Die Steigerung der Trefferrate TPR bewirkt bei System 2 zu $t = 131.400 \text{ h}$ letztendlich, dass alle Ausfälle vermeidbar sind. Daraus lässt sich auch schlussfolgern, dass die Nicht-Alarmrate FNR bei 0 % liegt und daher kein Ausfall mehr auftritt.

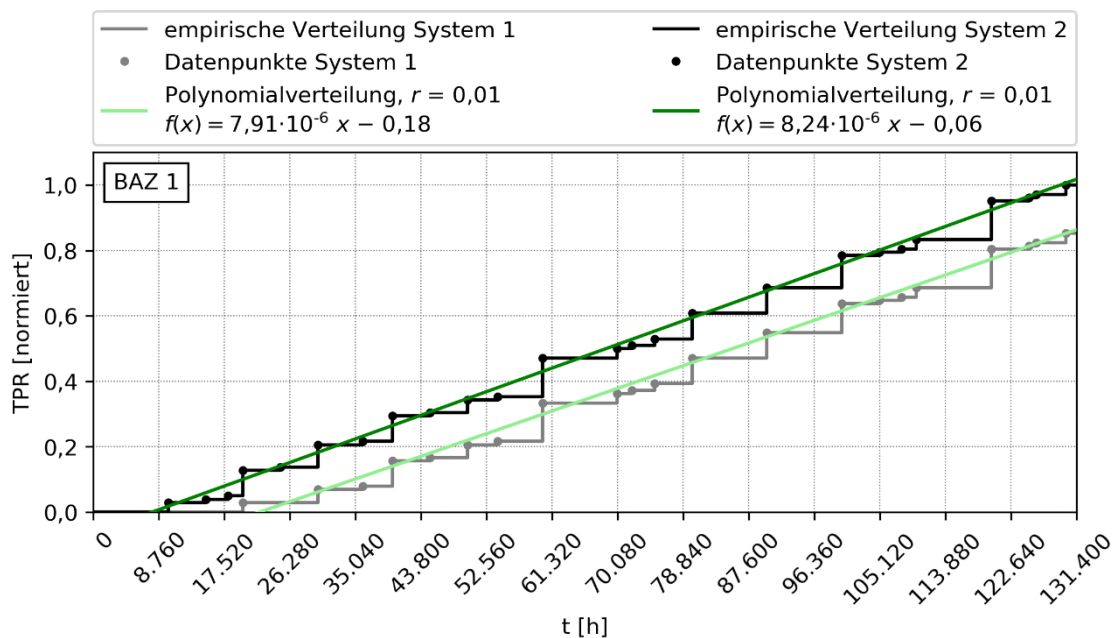


Bild 5.14: Normierte Datenpunkte, empirische und genäherte Verteilung der Trefferrate TPR des Systems BAZ 1 für vermeidbare Ausfälle entsprechend Szenario 3 mit zeitlichem Versatz der Systeme: Dargestellt sind die Datenpunkte (in Grau), die normierte empirische Verteilung (in Grau) sowie die zugehörige Näherungsfunktion (in Hellgrün) für die Trefferrate TPR des Systems BAZ 1 im System 1 (siehe Bild 5.6) über den Betriebszeitraum t_B in Stunden h . Zusätzlich sind die Datenpunkte (in Schwarz), die normierte empirische Verteilung (in Schwarz) sowie die zugehörige Näherungsfunktion (in Grün) für die optimierte Trefferrate TPR des Systems 2 von BAZ 1, die durch Optimierung der Treffer mittels eines baugleichen Systems mit einem zeitlichen Versatz $\Delta t \neq 0 \text{ h}$ berechnet wurde, dargestellt. Die Verteilungen wurden mittels der Gesamtanzahl der Ausfälle für BAZ 1 $P_{BAZ 1}(0 \leq t < B) = 102$ normiert. Für die Näherungsfunktionen wurde ein Polynom 1. Grades gewählt.

Abschließend kann für das Szenario 3 festgehalten werden, dass eine prinzipielle Steigerung der Trefferrate TPR durch die Nutzung von Daten eines anderen Systems erzielt werden kann. Allerdings sollte der Aufwand für die Implementierung einer Datenerfassung im Rahmen einer voraussagenden Instandhaltung für mehrere baugleiche Systeme und die Schaffung der gemeinsamen Nutzung eines

Vorhersagemodells dessen Nutzen gegenübergestellt werden. An einem konkreten Beispiel soll verdeutlicht werden, dass dieser Aufwand in der industriellen Produktion häufig gerechtfertigt sein kann, weil der Nutzen meist deutlich überwiegt. Werden ausgehend von 100 Ausfällen pro baugleiches System bis zu 15 Ausfälle ($TPR_{System\ 2}(t = 131.400\ h) - TPR_{System\ 1}(t = 131.400\ h) = 15\ \%$) mehr über den Betriebszeitraum t_B vorhergesagt, dann kann sich der Aufwand rechnen. Denn ein Ausfall bei einer kosten- und zeitintensiven Komponente kann gerne einmal zu einem Produktionsausfall von einer Woche führen. Auch Folgeschäden sind möglich, welche zusätzliche Aufwände und Kosten erzeugen.

5.5 Fazit

Dieses Kapitel hat neben der Anwendbarkeit der Methodik zur Bestimmung der Vorhersagegüte, eine Implementierung anhand von Struktogrammen vorgestellt. Weiterhin wurde eine erste Abschätzung der Gütemaße – Ausfallrate PR , Lernrate LPR , Trefferrate TPR und Nicht-Alarmrate FNR – für drei unterschiedliche Szenarien am Beispiel des Systems BAZ 1 vorgenommen. Dabei wurde nicht nur jeweils ein Wert zum Zeitpunkt des Betriebsendes angeben, sondern es wurde auch ein kompletter zeitlicher Verlauf je Gütemaß für jedes Szenario ausgearbeitet.

Die Datenbasis für die Gütemaße sind dabei im eigentlichen Sinne die Lebensdauern der Komponenten. Für das Training von Vorhersagemodellen ist die maßgebliche Datenbasis jedoch die Anzahl der erlernbaren Ausfälle LP bzw. die Lernrate LPR . Durch den Vergleich mit den weiteren Gütemaßen Trefferrate TPR und Nicht-Alarmrate FNR konnte nun der Informationsgehalt der Datenbasis für jeden Zeitpunkt t bestimmt werden. Die Arbeit und insbesondere dieses Kapitel konnten dabei zeigen, dass je Szenario zu einem unterschiedlichen Zeitpunkt t die bestmögliche Datenbasis für ein Vorhersagemodell erreicht wird. Zum Beispiel wird für Szenario 1 die bestmögliche Datenbasis nach der halben Betriebsdauer B , sprich 7,5 Jahre, erreicht. Dabei gilt weiterhin die Annahme, dass ein erstmaliger Ausfall ($s_K = 1$) zum Erlernen des Ausfalls ausreichend ist. Denn dann ist jeder erlernbare und innerhalb der Betriebsdauer B wiederauftretende Ausfall auch eingetreten. Für Szenario 2 und 3 liegt der Zeitpunkt der bestmöglichen Datenbasis dahingegen anders, weil die Datenerfassung später bzw. am anderen System früher beginnt und damit die voraussagbaren und wiederkehrenden Ausfälle bzw. das dies beschreibende Gütemaß Nicht-Alarmrate FNR einen anderen zeitlichen Verlauf bzw. einen geringeren oder gar keinen Anstieg aufweist.

Anhand der grafischen Darstellungen der Gütemaße Ausfallrate PR und Nicht-Alarmrate FNR bzw. genauer dem Vergleich zwischen diesen konnten die Unterschiede zwischen einer reaktiven und einer voraussagenden Instandhaltung vergleichbar gemacht werden. Hierauf basierend konnten die Vorteile durch den Einsatz von voraussagender Instandhaltung wie die Reduktion der Ausfälle deutlich herausgearbeitet werden. Zusammenfassend kann daher für die drei unterschiedlichen Szenarien dabei nachfolgendes Fazit gezogen werden.

Das erste Szenario hat gezeigt, dass der Einsatz der voraussagenden Instandhaltung sehr sinnvoll ist. Eine bestmögliche Steigerung der Trefferrate TPR auf 85 % über den kompletten Betriebszeitraum t_B ist für das System BAZ 1 zu erwarten. Das Szenario 1 ist sicherlich ein wünschenswerter Fall, wobei dann bereits in der Planungsphase der Einsatz von Predictive Maintenance zu berücksichtigen ist.

Das Szenario 2 hat die Auswirkungen bei einem späteren Einsatz von voraussagender Instandhaltung beleuchtet und gezeigt, dass nicht nur neue Systeme, sondern auch Bestandssysteme durchaus von voraussagender Instandhaltung profitieren können. Bei einem Beginn der Datenerfassung ab 4 Jahre bzw. 35.040 Stunden könnte noch eine Trefferrate TPR von 65 % bewirkt werden. Dieses zweite Szenario ist in der Praxis sehr gängig, denn in der Regel befinden sich (in der Motoren-Komponenten-Produktion) bereits viele Systeme seit längerer Zeit im Feld und damit in der Betriebsphase und werden erst sukzessive auf neuere Technologien wie die Erfassung von IoT-Daten aufgerüstet. Die Betrachtung dieses Szenarios verdeutlicht, dass für den nachträglichen Einsatz einer voraussagenden Instandhaltungsstrategie die verbleibende Betriebsdauer des jeweiligen Produktionssystems zu berücksichtigen ist und dann der Mehrwert beim Einsatz von Predictive Maintenance individuell für das System zu bewerten ist.

Das Szenario 3 hat gezeigt, dass für die Verbesserung der Ergebnisse der Gütemaße Trefferrate TPR und Nicht-Alarmrate FNR lediglich ein geringer zeitlicher Versatz $\Delta t \neq 0 h$ zwischen dem Start der Datenerfassung von baugleichen Systemen notwendig ist. Denn hierdurch lässt sich das Vorhersagemodell bei einem System erlernen und dann nutzbringend auf die weiteren Systeme übertragen. Eine Untersuchung unter idealisierten Bedingungen ergab, dass diese Art der Verwendung von Transfer-Learning die Trefferrate TPR auf 100 % steigert und die Nicht-Alarmrate FNR auf 0 % für ein baugleiches System reduziert. Das heißt alle Ausfälle innerhalb des Datenerfassungszeitraums können künftig vermieden werden. Auch wenn eine Nutzung von Daten aus baugleichen Maschinen eine Verbesserung der Trefferrate auf $TPR = 100 \%$ erwarten lässt, ist es fraglich, ob der für die Ermöglichung der Übertragbarkeit von Vorhersagemodellen notwendige Aufwand wie Sicherstellung der gleichen Abnutzung zu rechtfertigen wäre. Auch die heterogene Produktion von Motoren-Komponenten spricht dagegen das Transfer-Learning im großen Maßstab weiter voranzutreiben. Lediglich bei der Verwendung von Vorhersagemodellen, bereitgestellt durch einen Lieferanten, scheint dieser Ansatz interessant und sinnvoll, da dort möglicherweise weiteres domänenspezifisches Wissen nutzbringend Anwendung finden kann. Der potenzielle Hebel⁴³ für dieses Szenario kann sehr einfach bestimmt werden. Daher wird im weiteren Verlauf der Arbeit keine zusätzliche Untersuchung des Szenario 3 durchgeführt.

⁴³ Der potenzielle Hebel für die Trefferrate TPR des zweiten Systems $\Delta TPR_{System 2}$ ist die Differenz zwischen der Trefferrate TPR des ersten Systems und 100 % bei Betriebsende (131.400 Stunden). Also: $\Delta TPR_{System 2} = 100 \% - TPR_{System 1}(t = 131.400 h)$

Es sollte nochmals erwähnt werden, dass es sich bei dem angegebenen Wert für die Trefferrate TPR um eine obere Schranke handelt, die ein ideales Vorhersagemodell annimmt und durch einen realen Klassifikator überhaupt erst einmal erreicht werden muss. D.h. unter Umständen fallen die real erreichbare Trefferrate TPR und damit die Mehrwerte von Predictive Maintenance geringer aus. Kapitel 7 beleuchtet diese Mehrwerte nochmals genauer.

Bei der Bestimmung der Näherungsfunktion für dieses Beispiel wurde festgestellt, dass bei einer geringen Anzahl an Datenpunkten wie hier beim System BAZ 1 die Näherungsverteilung nur unzureichend approximiert werden kann. Dies könnte durch die Analyse einer größeren Menge an Datenpunkten verbessert werden. Im nachfolgenden Kapitel findet deshalb mit einer Stichprobe von 34 Produktionssystemen aus der Motoren-Komponenten-Produktion eine deutlich größere Grundgesamtheit Anwendung.

6 Generalisierung und Validierung der Vorhersagegüte

In diesem Kapitel wird für eine Klasse von Produktionssystemen die theoretisch maximal erreichbare Vorhersagegüte bestimmt. Hierzu wird eine Stichprobe von Produktionssystemen und deren Komponenten erhoben und gemeinsam ausgewertet. Durch die Analyse einer großen Menge an Lebensdauerdaten werden die Werte aus Kapitel 5 für die Vorhersagegüte eines Bearbeitungszentrums validiert. Die Ergebnisse aus diesem Kapitel ermöglichen künftig eine Vorhersage der Gütewerte von Produktionssystemen ohne eine weitere Erhebung von Lebensdauerdaten.

6.1 Motivation der Generalisierung und Validierung

Im vorhergehenden Kapitel 5 wurde die Anwendbarkeit für ein System sehr gut verdeutlicht. Dieses Ergebnis ist vorerst jedoch nur für das System BAZ 1 gültig, da ein Vergleich zu anderen Systemen bislang ausblieb. Aus diesem Grund sollten die prognostizierten Gütewerte vorerst besser nicht für andere Produktionssysteme herangezogen werden, sondern die Ergebnisse generalisiert und dann nochmals validiert werden.

In diesem Kapitel wird die entwickelte Methodik zur Bestimmung der Vorhersagegüte aus Kapitel 4 daher auf eine größere Stichprobe angewendet. Dadurch erfolgt eine generelle Abschätzung der Gütewerte eines Vorhersagemodells für eine Klasse von Produktionssystemen. Dies vereinfacht für sehr viele Anwender die Prognose der Gütewerte, denn vorher hätte jeder von ihnen erst die Lebensdauerdaten der Komponenten L_K des jeweiligen Systems erheben und die Gütewerte entsprechend der innerhalb dieser Arbeit entwickelten Vorgehensweise für das einzelne System bestimmen müssen. Die dafür notwendigen Lebensdauerdaten liegen in der Regel aber gar nicht vor, so dass die Methode damit oftmals gar keine Anwendung finden könnte oder der Aufwand für deren Durchführung in Folge der notwendigen Erhebung von Lebensdauerdaten L sehr hoch wäre.

Als Stichprobe werden in dieser Arbeit von 536 Komponenten, aus 34 möglichst unterschiedlichen Produktionssystemen der Motoren-Komponenten-Produktion, die Lebensdauern L verwendet. Mit Hilfe der bereits erhobenen Lebensdauern von den Komponenten L_K werden die folgenden Analysen zur Bestimmung der Vorhersagegüte durchgeführt. Durch die Bestimmung von Näherungsverteilungen lassen sich die Gütewerte für eine Klasse von Produktionssystemen generalisieren. Allerdings ist bei der Verwendung der allgemeingültigen Gütewerte eine statistische Unsicherheit zu berücksichtigen, welche durch die Varianz der einzelnen Produktionssysteme gegeben ist. In dieser Arbeit wird deshalb für die allgemeingültigen Gütewerte eine Hüllkurve bestimmt, welche deren Standardabweichung s angibt.

Zur abschließenden Validierung der generalisierten Gütewerte wird mittels des Systems BAZ 1 aus Kapitel 5 noch bereits direkt überprüft, ob die Gütewerte der Klasse

von Produktionssystemen repräsentativ für ein einzelnes, von der Stichprobe unabhängiges Produktionssystem (hier: das unabhängige System BAZ 1) sind.

6.2 Stichprobenumfang für eine Klasse von Produktionssystemen

Zur Bestimmung der Vorhersagegüte unter Annahme einer idealen Klassifikation von potenziellen Ausfällen ist die Information über die Lebensdauern von Komponenten L_K in Produktionssystemen relevant. Der folgende Abschnitt zeigt die erhobenen Daten und ordnet diese den entsprechenden Produktionssystemen zu.

Entgegen den Ausführungen aus Abschnitt 2.1.5 (dass die Arbeit sich nur auf Fertigungszellen konzentriert) wurden bei der folgenden Analyse unterschiedlichste Produktionssysteme, einschließlich Fertigungszellen, verwendet. Damit verändert bzw. vergrößert sich zwar auch die Grundgesamtheit, aber es ergibt sich auch die Möglichkeit, eine allgemeingültige Aussage für die Vorhersagegüte im Kontext von Produktionssystemen abzuleiten und eine verallgemeinerte Vorgehensweise zur Bestimmung der idealisierten Vorhersagegüte darzulegen.

In diesem Zusammenhang ist es besonders notwendig, eine Stichprobe mit einem ausreichend großen Umfang zu wählen, sodass auf die Grundgesamtheit geschlossen werden kann. Dazu wurde eine eigene Erhebung der mittleren Lebensdauern von Komponenten (L_K) aus unterschiedlichen Produktionssystemen vorgenommen. Diese basiert auf Informationen aus TCO-Berichten von Produktionssystemen einer Motoren-Komponenten-Produktion und bildet neben den Fertigungsprozessen Umformung, Zerspanung und Härtung auch weitere grundlegendere Aufgaben wie Handhabung, Montage und Messen ab. Die mittleren Lebensdauern der Komponenten sind in Bild 6.1 – deren jeweiligen Produktionssystemen zugeordnet – dargestellt. Die exakten Werte für die Lebensdauern je System sind als Menge $L_{System} = \{L_1, \dots, L_K\}$ im Anhang A.1 angegeben. Die Gesamtheit der Lebensdauern aller Systeme L – ausgenommen BAZ 1 – wird als Stichprobe 1 bezeichnet.

In Bild 6.1 sind die erhobenen Lebensdauern L_K gezeigt. Dabei ist auch noch das Bearbeitungszentrum 1 (BAZ 1) aufgeführt, allerdings wird dieses nicht zur Generalisierung der Gütewerte herangezogen und ist nur der Vollständigkeit aufgelistet worden. Das System BAZ 1, dient wie bereits eingangs erläutert, später zur Validierung der Stichprobe 1. Die ausgewählten 536 Komponenten sind in 34 Produktionssystemen verbaut. Da die Informationen der mittleren Lebensdauern L_K exakt aus den zugehörigen TCO-Berichten des jeweiligen Produktionssystems übernommen wurden, wurde keine Prüfung auf Vollständigkeit der Angaben gemacht, wodurch Komponenten K vernachlässigt worden sein könnten. Auch sind fehlerhafte Werte für die Lebensdauern L_K möglich, welche zu Ausreißern in den folgenden Gütewerten führen könnten. Dies wird jedoch als irrelevant erachtet, da dies nur eine geringe Anzahl der Daten betrifft und somit in der Gesamtheit der Daten nur einen geringen statistischen Einfluss hat. Bild 6.1 enthält entsprechend der

bereits aus früheren Kapiteln bekannten, methodischen Vorgehensweise nur Lebensdauern L_K , welche kleiner als die Betriebsdauer B sind. Jeder Punkt im Bild stellt eine Lebensdauer für eine Komponente L_K dar und ist zur Unterscheidung der Systeme farblich unterschieden.

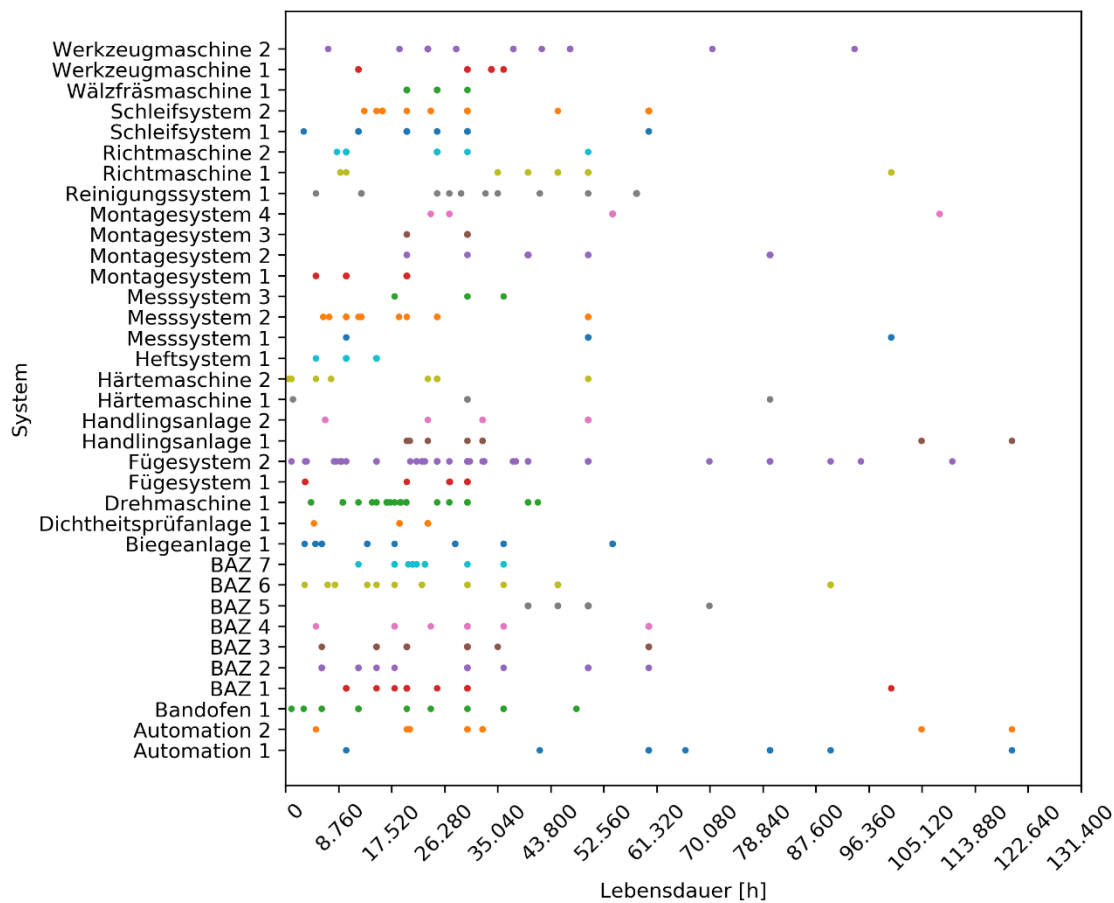


Bild 6.1: Erhobene mittlere Lebensdauern der Komponenten je Produktionssystem: Dargestellt sind in Form von Punkten die Lebensdauern von Komponente L_K für das jeweilige System. Angegeben werden die Komponenten-Lebensdauern L_K in Stunden h , wobei der Zeitraum $t = [0, 131.400[$ auf die Betriebsdauer B begrenzt wurde. Die Lebensdauern L_K entspringen einer eigenen Erhebung basierend auf TCO-Berichten von Produktionssystemen in der Motoren-Komponenten-Fertigung. Eine identische farbliche Hervorhebung der Lebensdauern je System erfolgte für eine bessere Übersicht. Die Farbe selbst hat keine Bedeutung.

6.3 Generalisierung der Vorhersagegüte für eine Klasse von Produktionssystemen

Im Folgenden wird zunächst die Vorgehensweise zur Generalisierung der Vorhersagegüte eingeführt. Dazu wird neben der Bestimmung einer Näherungsfunktion auch noch die Ermittlung der statistischen Unsicherheit der Prognose durch Berechnung der Standardabweichung hergeleitet. Basierend auf der oben eingeführten Stichprobe 1 wurde anschließend das erste Szenario untersucht. Die Ergebnisse für das zweite Szenario mit einem späteren Beginn der Datenerfassung bei $t_{data} = 35.040 h$ sind im Anhang A.2 aufgeführt.

6.3.1 Vorgehensweise der Generalisierung

Für jedes System aus der Stichprobe 1 werden unabhängig von den anderen Systemen zunächst die Ausfallrate PR , für die Lernrate LPR , für die Trefferrate TPR oder für die Nicht-Alarmrate FNR bestimmt. Im Weiteren werden ausschließlich die Datenpunkte verwendet und keine empirische einzelne oder genäherte Verteilung modelliert. Stattdessen wird aufbauend auf den Datenpunkten der 34 Systeme eine gemeinsame Näherungsfunktion je Gütemaß bestimmt.

Weiterhin ist für die ermittelten Gütewerte eine statistische Unsicherheit zu berücksichtigen. Diese bezieht sich auf den Wertebereich der Gütewerte und nicht auf den Zeitbereich. Die Unsicherheit wird dabei sowohl grafisch, in Form von Hüllkurven, als auch quantitativ, in Form eines Wertes, angegeben. Herkömmlicherweise werden entsprechend GUM⁴⁴ zusätzlich zu dem eigentlichen Ergebniswert sowohl die systematische (in diesem Fall nicht zu berücksichtigende⁴⁵) als auch die statistische Unsicherheit in Form eines positiven und negativen Wertes, also einem Wertepaar, aufgeführt. Für die Angabe der Streuung der Werte wird herkömmlicherweise die Standardabweichung verwendet, diese beschreibt inwieweit die Beobachtungswerte von dem Mittelwert abweichen. In dieser Arbeit wird zunächst die als Differenz d bezeichnete Abweichung zwischen dem Erwartungswert $Gütemaß_{System}(t)$ und dem Wert der Polynomialverteilung zu einem Zeitpunkt $f_{Stichprobe\ 1}(x = t)$ berechnet. Dabei wurden nur die Zeitpunkte t verwendet, an denen ein Datenpunkt für das Gütemaß vorliegt.

$$d(t | t \in Gütemaß_{System}) = f_{Stichprobe\ 1}(x = t) - Gütemaß_{System}(t) \quad (6.1)$$

Aus den Differenzen d wird dann die Standardabweichung s bestimmt. Da es sich um die Standardabweichung s einer Stichprobe handelt wird $n - 1$ angewendet.

$$s = \sqrt{\frac{\sum d^2}{n - 1}} \quad (6.2)$$

6.3.2 Generalisierung der Vorhersagegüte im Szenario 1

Als Erstes soll mittels Bild 6.2 ein grafischer Überblick über die Datenpunkte der einzelnen Systeme für die wahren Ausfälle bzw. relativ betrachtet der Ausfallrate PR gegeben werden. Im Bild ist neben den Datenpunkten die Näherungsverteilung der Stichprobe 1 als rote Linie dargestellt, wobei der Näherungsverteilung ein Polynom 1. Grades zugrunde liegt. Zusätzlich verdeutlicht die in

⁴⁴ Guide on uncertainties in measurement [100].

⁴⁵ Eine Berücksichtigung der systematischen Unsicherheit ist nicht notwendig, da die Lebensdauerdaten von verschiedenen Messeinrichtungen (Herstellern bzw. Lieferanten) stammen und somit sich über der Grundgesamtheit ausgleichen sollten.

Hellrot dargestellte Hüllkurve die Streuung der Stichprobe 1 und basiert auf der zuvor ermittelten Standardabweichung s . Allerdings ist diese in Bild 6.2 aufgrund des geringen Wertes von $s = 0,028$ und der Vielzahl an Datenpunkten kaum erkennbar.

Die lineare Verteilung für die Ausfallrate PR kann damit begründet werden, dass die Ausfälle durch ein Vielfaches der Lebensdauern L_K beschrieben werden und somit neue Ausfälle, unter der Annahme einer Delta-Verteilung, in äquidistanten Abständen auftreten. Dies entspricht auch den Erwartungen für die sogenannten Abnutzungsausfälle, bereits dargestellt in Bild 2.3. Die Konsequenz aus der linearen Ausfallrate eines Systems ist, dass in gleichen Abständen gleich viele Ausfälle auftreten und somit die technische Verfügbarkeit V_T erwartungsgemäß gleichbleibend ist.

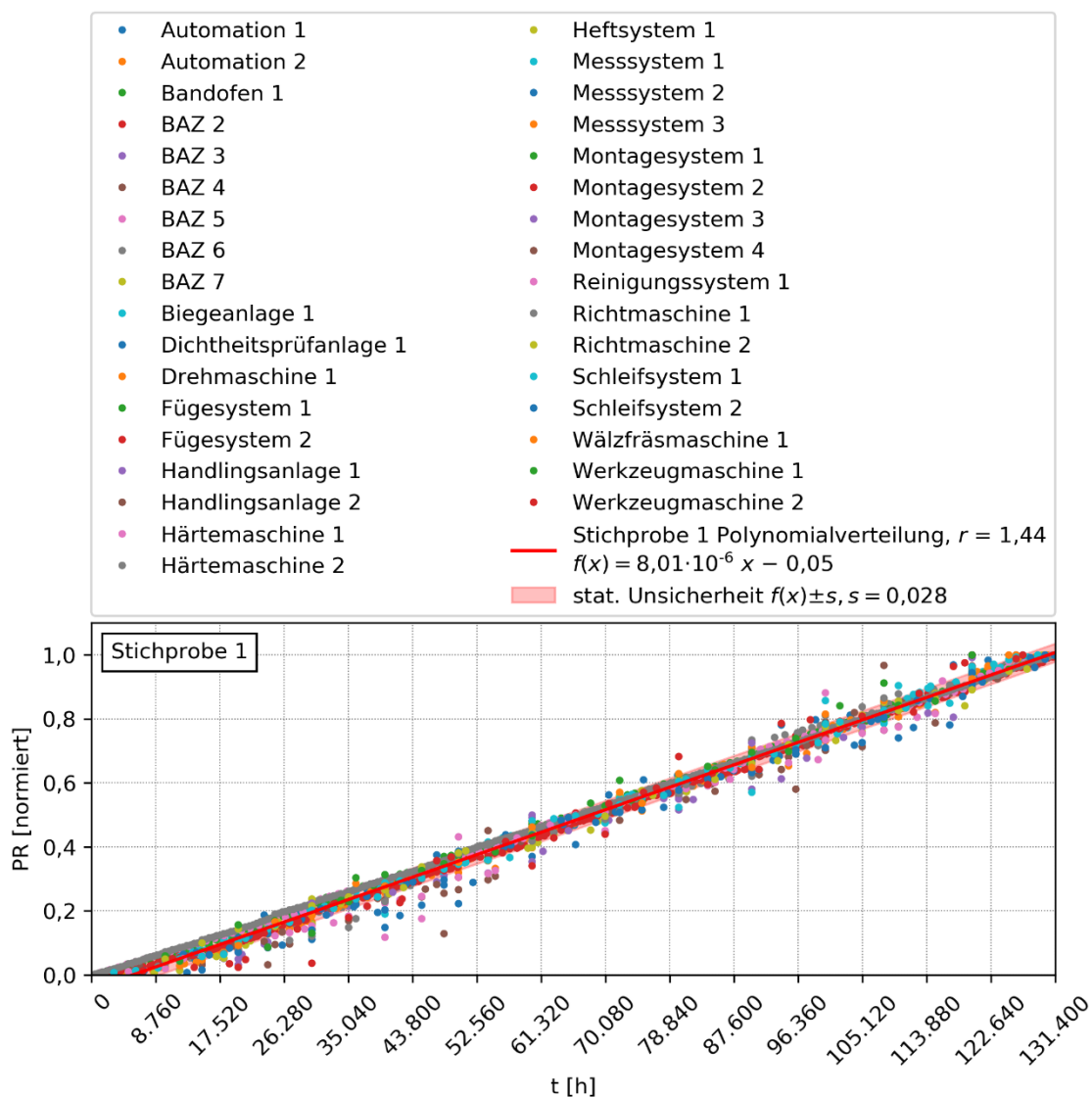


Bild 6.2: Normierte Datenpunkte aller einzelnen Systeme und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Ausfallrate PR für wahre Ausfälle: Dargestellt sind die Datenpunkte der Systeme (in verschiedenen Farben) sowie die Näherungsfunktionen der Stichprobe 1 (in Rot) und deren statistische Unsicherheit s (Hüllkurve in Hellrot) für die Ausfallrate PR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Ausfälle des jeweiligen Systems $P(0 \leq t < B)$ normiert. Für die Näherungsfunktion der Stichprobe 1 wurde ein Polynom 1. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400] h$.

Aus dem Bild 6.2 ist außerdem erkennbar, dass erwartungsgemäß nicht jeder Datenpunkt jedes einzelnen Systems zur Gesamtheit der Systeme passt und von der roten Linie, welche die Näherungsverteilung der Stichprobe 1 und somit die Klasse der Produktionssysteme beschreibt, sowie deren Hüllkurve abweicht. Diese Abweichung ist aber damit erklärbar, dass für die Hüllkurve eine Standardabweichung von $\pm 1 \sigma$ verwendet wurde. Unter Annahme einer gaußförmigen Verteilung der statistischen Unsicherheit ist dann zu erwarten, dass nur 68,2 % aller Datenpunkte innerhalb der Hüllkurve liegen. Die geringe Streuung s von ca. 2,8 % ist ein gutes Ergebnis und lässt eine genaue Prognose der Ausfallrate PR erwarten.

Nachdem der grundlegende Gütewert der Ausfallrate PR zu jedem Zeitpunkt t bestimmt wurde, wird als nächstes eine Generalisierung der Werte der Lernrate LPR vorgenommen. In Bild 6.3 ist analog zur Ausfallrate PR der Gütewert der Lernrate LPR über der Zeit t grafisch dargestellt. Durch die größere Anzahl an Datenpunkten ist es nun möglich für die Stichprobe 1 eine bessere Approximation der genäherten Verteilung als bei dem System BAZ 1 zu erzielen. Allerdings spricht das im Verhältnis zur Ausfallrate PR größere Unsicherheitsband (Standardabweichung $s = 0,187$) für eine größere Streuung der Daten und damit einer höheren statistischen Unsicherheit bei der Prognose der Lernrate LPR .

Die Lernrate LPR über den betrachteten Betriebszeitraum t_B lässt die Schlussfolgerung zu, dass am Anfang der Betriebsdauer die meisten Ausfälle erlernt werden. Dies ist vorteilhaft, denn nur ein Ausfall, welcher vor 7,5 Jahren bzw. 65.700 Stunden erlernt wurde, tritt innerhalb des Betriebszeitraums t_B erneut auf und kann dann vorhergesagt werden. Im statistischen Mittel werden allerdings rund 5 Jahre bzw. 43.800 Stunden benötigt, um 80 % der Ausfälle zu erlernen.

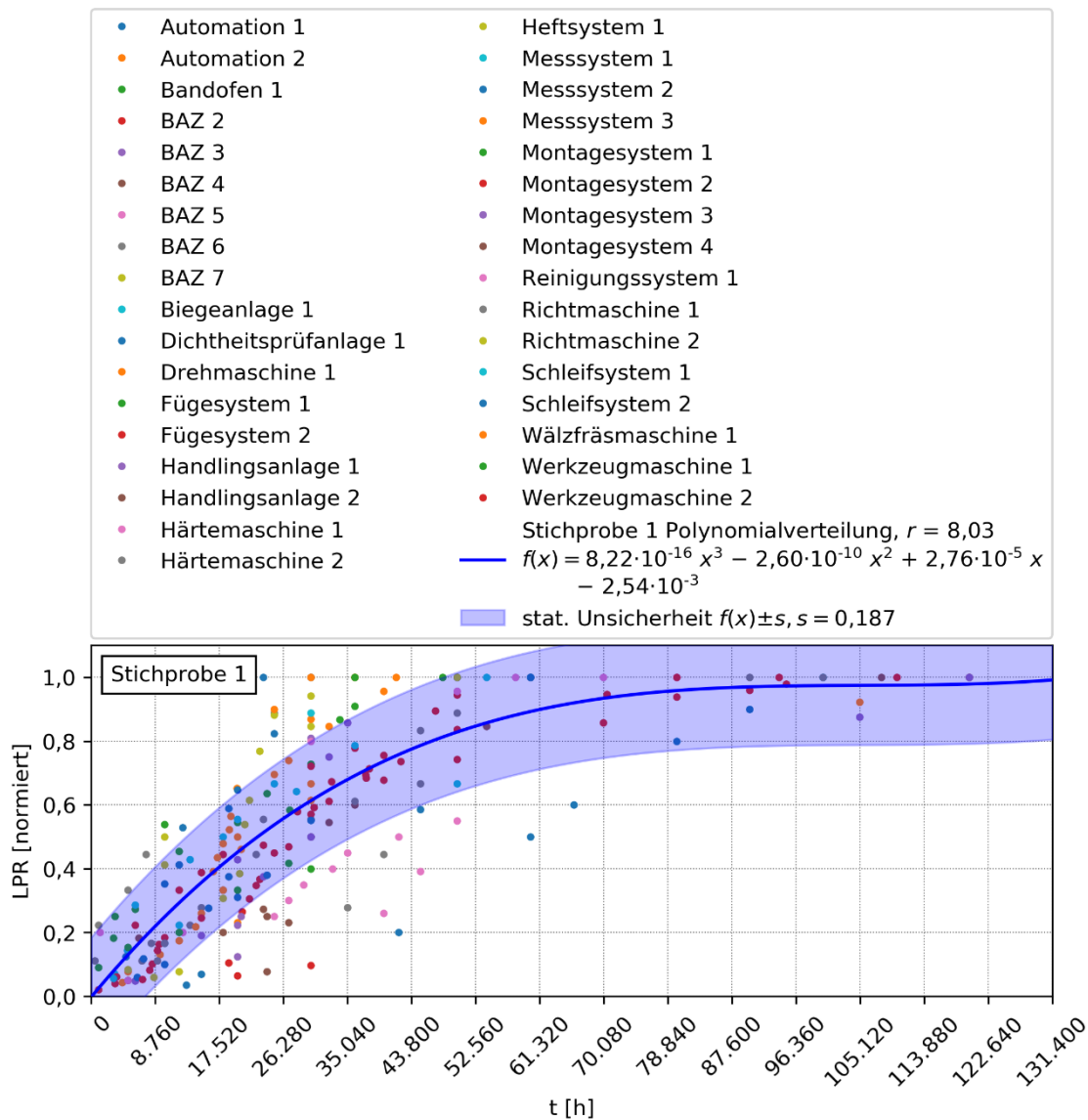


Bild 6.3: Normierte Datenpunkte aller einzelnen Systeme und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Lernrate LPR für erlernbare Ausfälle entsprechend Szenario 1: Dargestellt sind die Datenpunkte der Systeme (in verschiedenen Farben) sowie die Näherungsfunktionen der Stichprobe 1 (in Blau) und deren statistische Unsicherheit s (Hüllkurve in Hellblau) für die Lernrate LPR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Komponenten des jeweiligen Systems $|K|$ normiert. Für die Näherungsfunktion der Stichprobe 1 wurde ein Polynom 3. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400[h$.

Inwieweit sich die statistische Unsicherheit der Lernrate LPR auf die Trefferrate TPR überträgt wurde anschließend untersucht. Das Bild 6.4 gibt einen ganzheitlichen Überblick über die Datenpunkte der einzelnen 34 Systeme für die Trefferrate TPR und stellt neben der Näherungsfunktion der Stichprobe 1 in Grün auch die statistische Unsicherheit als hellgrüne Hüllkurve dar. Die schlechte Standardabweichung s der Lernrate LPR konnte sich scheinbar nicht auf die der Trefferrate TPR übertragen, sondern entspricht nur $s = 0,072$.

Für die Trefferrate TPR ist interessant, dass einzelne Datenpunkte eher nach unten von der Näherungsfunktion abweichen, während einige Systeme anhand ihrer Datenpunkte generell eine höhere Trefferrate TPR erwarten lassen.

Insgesamt ist festzuhalten, dass der zeitliche Verlauf der Trefferrate TPR weitestgehend linear ist mit einem geringen quadratischen Beitrag. Dies bedeutet für die Praxis, dass am Anfang des Betriebszeitraums t_B eher weniger Ausfälle vorhergesagt werden können und somit die Treffer TP noch geringer sind. Die Ursache liegt in den notwendigen Erstaussfälle $s_K = 1$.

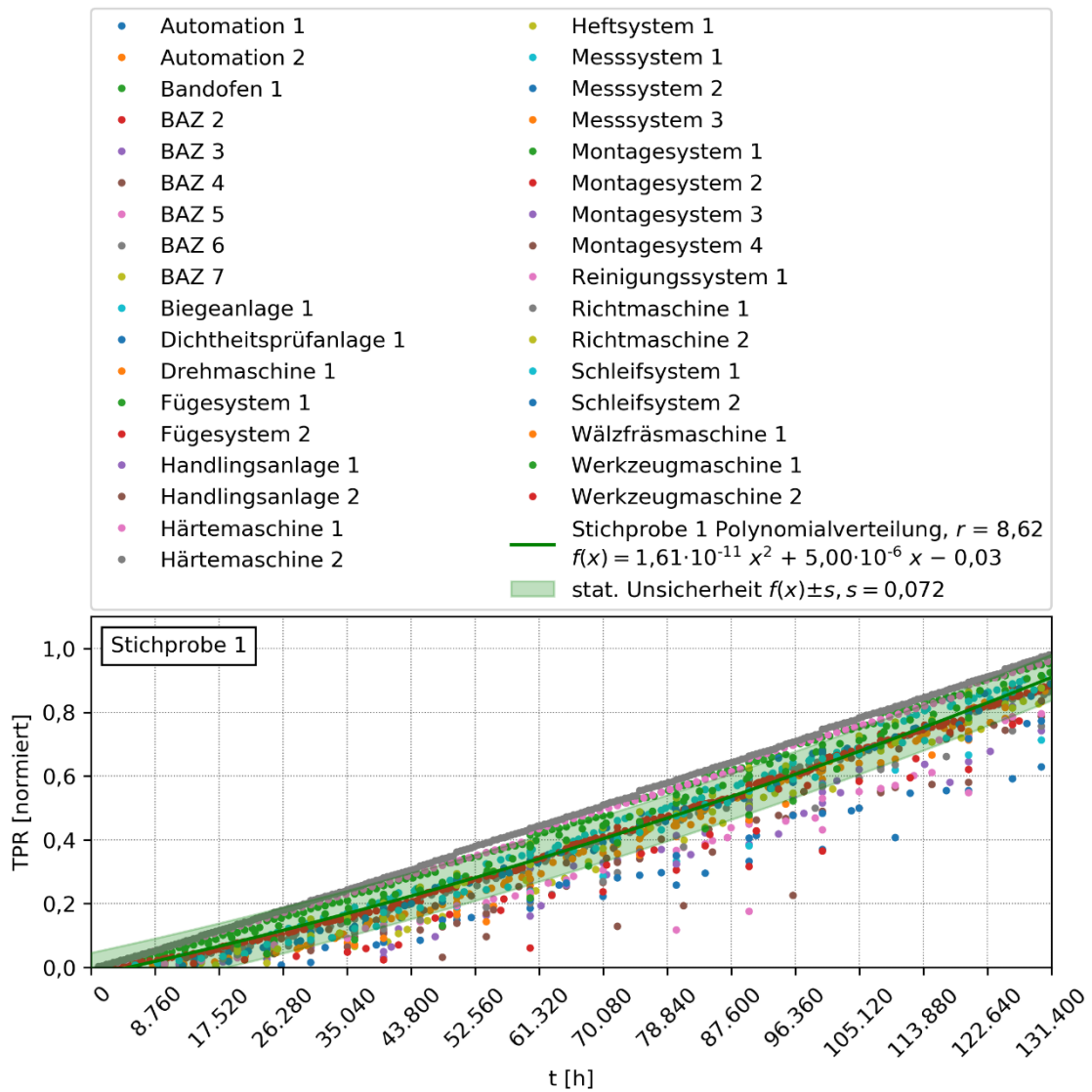


Bild 6.4: Normierte Datenpunkte aller einzelnen Systeme und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Trefferrate TPR für vermeidbare Ausfälle entsprechend Szenario 1: Dargestellt sind die Datenpunkte der Systeme (in verschiedenen Farben) sowie die Näherungsfunktionen der Stichprobe 1 (in Grün) und deren statistische Unsicherheit s (Hüllkurve in Hellgrün) für die Trefferrate TPR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Ausfälle des jeweiligen Systems $P(0 \leq t < B)$ normiert. Für die Näherungsfunktion der Stichprobe 1 wurde ein Polynom 2. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400] h$.

Die vereinzelt schlechte Beschreibung der Trefferrate TPR durch die Näherungsfunktion und ihre Hüllkurve wirkt sich jedoch nicht konträr auf die Nicht-Alarmrate FNR , in Bild 6.5 dargestellt, aus. Die statistische Unsicherheit $s = 0,057$ für die Nicht-Alarmrate FNR der Stichprobe 1 fällt verhältnismäßig gering aus. Auch die Lage der einzelnen Datenpunkte lässt darauf schließen, dass dieses für die Arbeit

wesentliche Gütemaß sehr gut durch die Näherungsverteilung beschrieben wird und damit eine sehr gute Prognose der Nicht-Alarmrate FNR für die Klasse der Produktionssysteme gegeben ist.

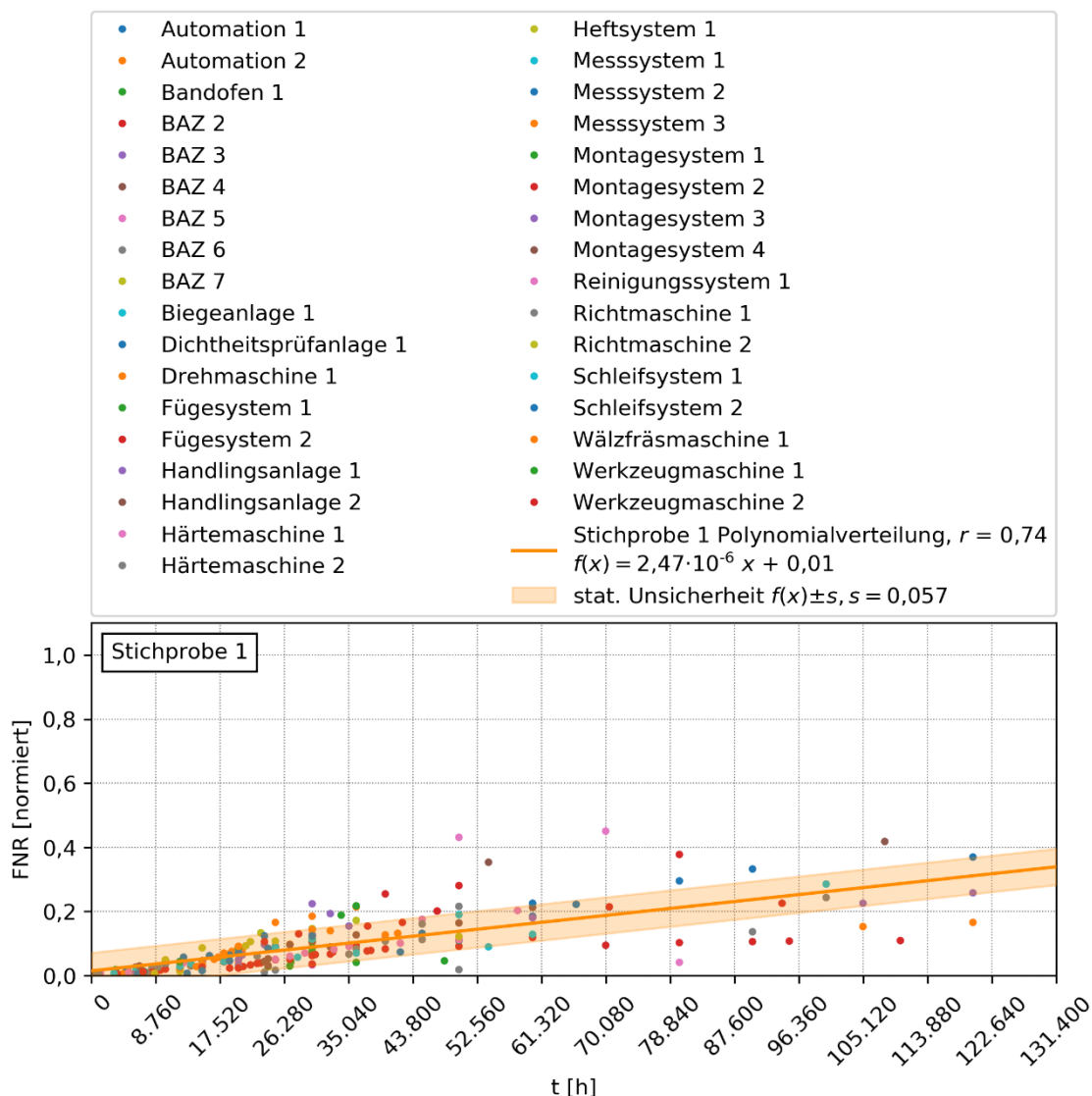


Bild 6.5: Normierte Datenpunkte aller einzelnen Systeme und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Nicht-Alarmrate FNR für verbleibende Ausfälle entsprechend Szenario 1: Dargestellt sind die Datenpunkte der Systeme (in verschiedenen Farben) sowie die Näherungsfunktionen der Stichprobe 1 (in Orange) und deren statistische Unsicherheit s (Hüllkurve in Hellorange) für die Nicht-Alarmrate FNR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Ausfälle des jeweiligen Systems $P(0 \leq t < B)$ normiert. Für die Näherungsfunktion der Stichprobe 1 wurde ein Polynom 1. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400[h$.

Dieser Abschnitt konnte sehr schön demonstrieren wie die Generalisierung der Datenpunkte durch eine Näherungsverteilung und damit der Gütewerte zu jedem Zeitpunkt t erfolgen kann. In dieser Arbeit wurde dadurch eine allgemeingültige Vorhersagegüte für eine Klasse von Produktionssystemen erzielt, welche künftig Data Scientists, Instandhalter(inne)n und Unternehmen helfen kann, eine Abschätzung für die Reduktion von Ausfällen in einem Produktionssystem durch voraussagende Instandhaltung durchzuführen.

6.4 Validierung der Vorhersagegüte für eine Klasse von Produktionssystemen

Das vorgehende Kapitel 6.3 hat für eine Klasse von Produktionssystemen eine allgemeingültige Ableitung der Werte der Gütemaße Ausfallrate PR , Lernrate LPR , Trefferrate TPR und Nicht-Alarmrate FNR und deren entsprechende statistische Unsicherheit s erzielt. Nun stellt sich die Frage, inwieweit die Gütewerte der Stichprobe 1 die Grundgesamtheit der Gütewerte von Produktionssystemen widerspiegelt. Hierzu wird in diesem Kapitel eine Validierung mithilfe der bereits in Kapitel 5 ausgewerteten Gütewerte des Bearbeitungszentrums BAZ 1 durchgeführt.

Die Vorgehensweise zur Validierung der in Kapitel 6.3 bestimmten Gütemaße wird dabei wie folgt konzipiert. Es wird geprüft, inwieweit die Datenpunkte von dem System BAZ 1 (in Schwarz abgebildet) innerhalb der Hüllkurve der genäherten Verteilung von Stichprobe 1 plus/minus der Standardabweichung ($\pm s$) liegen.

In Bild 6.6 ist die Validierung der Ausfallrate PR für die Stichprobe 1 – Klasse der Produktionssysteme – abgebildet. Deutlich wird die sehr gute Übereinstimmung der Datenpunkte für System BAZ 1 und der Stichprobe 1 einschließlich der Standardabweichung $f(x) \pm s$. Damit lässt sich für dieses Gütemaß bereits das allgemeingültige Ergebnis positiv bestätigen.

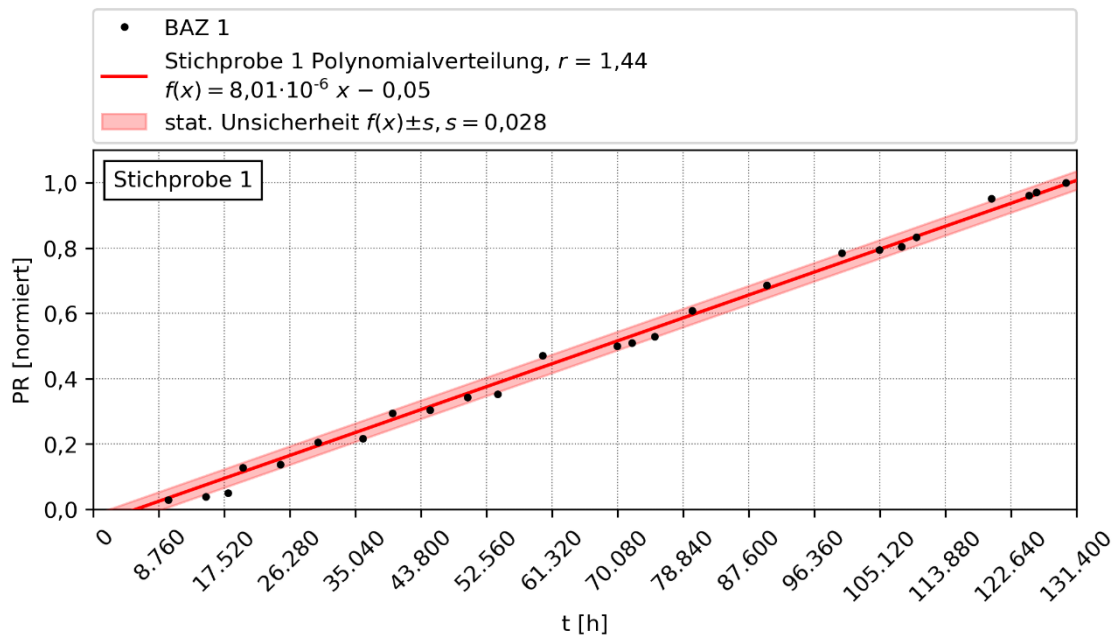


Bild 6.6: Normierte Datenpunkte des Systems BAZ 1 und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Ausfallrate PR für wahre Ausfälle: Dargestellt sind die Datenpunkte des Systems BAZ 1 (in Schwarz) sowie die Näherungsfunktionen der Stichprobe 1 (in Rot) und deren statistische Unsicherheit s (Hüllkurve in Hellrot) für die Ausfallrate PR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Ausfälle für BAZ 1 $P_{BAZ1}(0 \leq t < B) = 102$ normiert. Für die Näherungsfunktion der Stichprobe 1 wurde ein Polynom 1. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400] h$.

Nachstehend wird ausführlich die Analyse des Szenario 1 detailliert durchgeführt. Die Ergebnisse für Szenario 2 sind dem Anhang A.3 zu entnehmen. Für Szenario 3

erscheint keine extra Validierung notwendig, weil die Ergebnisse der Validierung von Szenario 1 auf dieses weitestgehend übertragbar sind.

Begonnen wird mit der Analyse der Lernrate LPR . Bereits die Generalisierung der Lernrate LPR im vorhergehenden Abschnitt 6.3.2 hat eine eher höhere statistische Unsicherheit angedeutet. In Bild 6.7 sind neben der abgeschätzten Lernrate $LPR \pm s$ für die Klasse von Produktionssystemen (Stichprobe 1) auch die Datenpunkte des Systems BAZ 1 abgebildet. Trotz der vergleichsweise großen statistischen Unsicherheit von $s = 0,186$, also $\sim 19\%$, liegen einige Datenpunkte von BAZ 1 außerhalb der Hüllkurve. Dies ist nicht unbedingt verwunderlich, denn die Standardabweichung beschreibt nur 68,26 % der Datenpunkte von Stichprobe 1. Somit kann es auch für System BAZ 1 zu Datenpunkten außerhalb der Hüllkurve kommen. Dieses Ergebnis wirkt sicherlich erstmal nicht zufriedenstellend, allerdings bietet das Ergebnis im Vergleich zu allen bisherigen Methoden die Möglichkeit für den Data Scientist abzuschätzen, wie viele Ausfälle in der Datenbasis bereits erfasst wurden und damit dann zukünftig, einen perfekt funktionierenden Klassifikator vorausgesetzt, vermieden werden können. Somit können z.B. 50 % der Ausfälle durchschnittlich nach 2,5 Jahren erfasst und entsprechend vermieden werden. Allerdings kann dies auf Grund der statistischen Unsicherheit auch schon nach 1,5 Jahren oder leider auch erst nach 4 Jahren erzielt werden. Wer also auf der sicheren Seite sein möchte, wartet einfach 4 Jahre ab. Daran wird auch wieder deutlich, dass die Erfassung einer ausreichenden Datenbasis viele Jahre dauern kann.

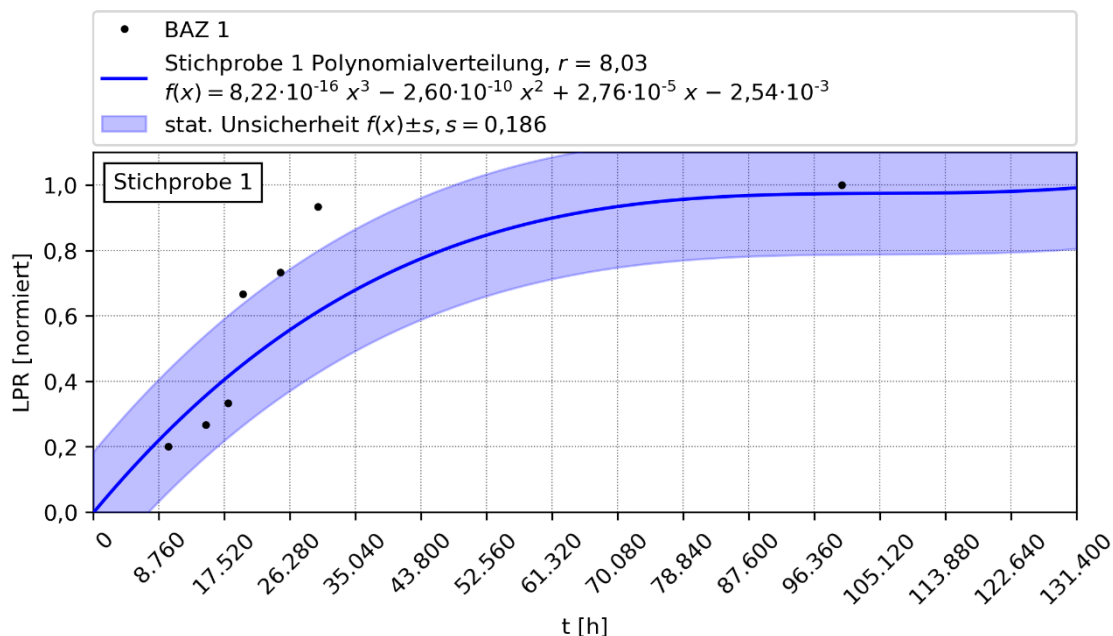


Bild 6.7: Normierte Datenpunkte des Systems BAZ 1 und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Lernrate LPR für erlernbare Ausfälle entsprechend Szenario 1: Dargestellt sind die Datenpunkte des Systems BAZ 1 (in Schwarz) sowie die Näherungsfunktionen der Stichprobe 1 (in Blau) und deren statistische Unsicherheit s (Hüllkurve in Hellblau) für die Lernrate LPR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Komponenten für BAZ 1 $|K| = 15$ normiert. Für die Näherungsfunktion der Stichprobe 1 wurde ein Polynom 3. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400[h$.

Inwieweit die Trefferrate TPR generalisiert abgeschätzt werden kann, zeigt die Validierung durch das System BAZ 1 in Bild 6.8. Die Abschätzung des Gütemaßes durch die Stichprobe 1 inklusive der Hüllkurve passt optisch sehr gut zu den Datenpunkten von System BAZ 1. Tendenziell liegen die Datenpunkte unter oder auf der Polynomialverteilung von Stichprobe 1 und es befinden sich nahezu alle innerhalb der Hüllkurve. Damit kann also gesagt werden, dass auch wenn die Trefferrate TPR für das System BAZ 1 anhand der Ergebnisse der Stichprobe 1 abgeschätzt werden würde, die eigentliche Trefferrate TPR des Systems BAZ 1 zwar überschätzt wird, aber immer noch im Rahmen der statistischen Unsicherheit s zutreffend ist.

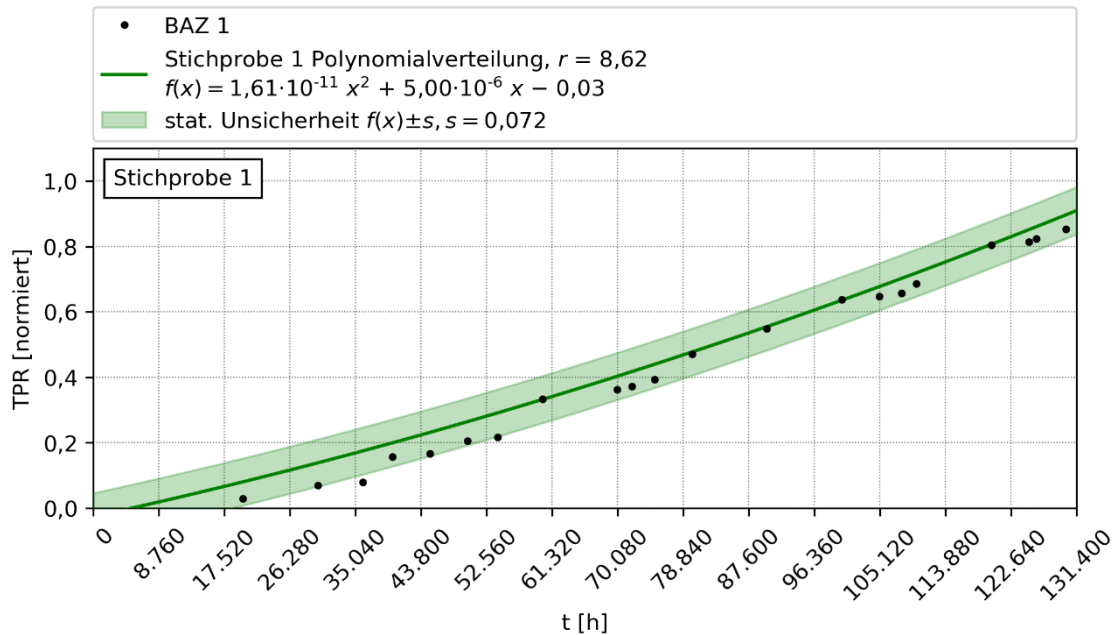


Bild 6.8: Normierte Datenpunkte des Systems BAZ 1 und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Trefferrate TPR für vermeidbare Ausfälle entsprechend Szenario 1: Dargestellt sind die Datenpunkte des Systems BAZ 1 (in Schwarz) sowie die Näherungsfunktionen der Stichprobe 1 (in Grün) und deren statistische Unsicherheit s (Hüllkurve in Hellgrün) für die Trefferrate TPR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Ausfälle für BAZ 1 $P_{BAZ\ 1}(0 \leq t < B) = 102$ normiert. Für die Näherungsfunktion der Stichprobe 1 wurde ein Polynom 2. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400[h$.

Die Validierung der Stichprobe 1 des Gütemaßes Nicht-Alarmrate FNR in Bild 6.9 stellt erwartungsgemäß eine ähnlich gute Abschätzung des Gütemaßes dar. Allerdings gibt einen Ausreißer zwischen 96.360 und 105.120 Stunden, welcher möglicherweise auf die Approximation der Verteilung durch ein Polynom 1. Grades zurückzuführen ist. Nichtsdestotrotz ist das Ergebnis in Bild 6.9 vertretbar und kann für weitere Auswertungen herangezogen werden. Zumal darauf hinzuweisen ist, dass eine Überschätzung der Nicht-Alarmrate FNR einem konservativen Verhalten entspricht. Denn ein solches Überschätzen würde bedeuten, dass in der Praxis bei einem (Produktions-)System weniger nicht-vorhergesagte Ausfälle auftreten würden als hiermit abgeschätzt wurde.

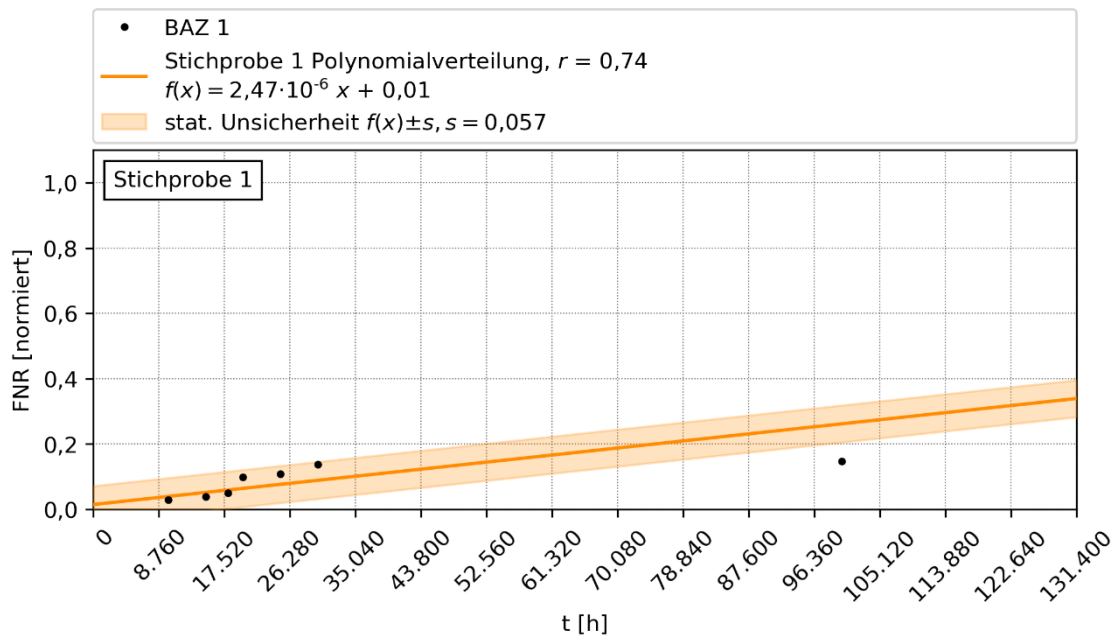


Bild 6.9: Normierte Datenpunkte des Systems BAZ 1 und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Nicht-Alarmrate FNR für verbleibende Ausfälle entsprechend Szenario 1: Dargestellt sind die Datenpunkte des Systems BAZ 1 (in Schwarz) sowie die Näherungsfunktionen der Stichprobe 1 (in Orange) und deren statistische Unsicherheit s (Hüllkurve in Hellorange) für die Nicht-Alarmrate FNR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Ausfälle für BAZ 1 $P_{BAZ\ 1}(0 \leq t < B) = 102$ normiert. Für die Näherungsfunktion der Stichprobe 1 wurde ein Polynom 1. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400[h$.

Zusammenfassend hat dieser Abschnitt sehr gut verdeutlicht, dass eine Generalisierung der Bestimmung von Gütewerten unter Berücksichtigung einer statistischen Unsicherheit vorgenommen werden kann. Aus didaktischen Gründen wurde auf eine Darstellung der Kreuzvalidierung mit einer Aufteilung von 70/30 % der Systeme in diesem Kapitel allerdings verzichtet. Die entsprechende Kreuzvalidierung ist jedoch zusätzlich im Anhang A.4 zu finden und bestätigt die ebenso validen Gütewerte für die Klasse des Produktionssystems.

6.5 Fazit

Die Methodik zur Bestimmung möglichst allgemein anwendbarer und aussagekräftiger Gütewerte für Produktionssysteme wurde auf eine breit gewählte Stichprobe von Produktionssystemen angewendet. Zusätzlich wurde zur weiteren Verbesserung der Aussagekraft der Gütewerte auch noch die statistische Unsicherheit der Ausfallrate PR , Lernrate LPR , Trefferrate TPR und Nicht-Alarmrate FNR für diese Stichprobe bestimmt.

Im Anschluss daran wurde die Allgemeingültigkeit durch Vergleich mit dem aus Kapitel 5 bekannten und unabhängigen Produktionssystem (BAZ 1) umfassend validiert.

Dieses Kapitel hat also die Anwendbarkeit der selbstentwickelten Methode und vor allem die Generalisierbarkeit ihrer Gütewerte bestätigt. Es kann somit festgehalten werden, dass eine objektive und breit einsetzbare Methode geschaffen wurde. Mittels dieser können basierend auf einer Stichprobe von Systemen (bspw. Produktionssysteme) sowie deren Beschaffenheit und der Lebensdauer ihrer einzelnen Einheiten allgemeingültige Gütewerte für diese System-Klasse bestimmt werden. Diese generalisierten Gütewerte können dann zur Abschätzung der oberen Schranke für die Gütemaße von Vorhersagemodellen für ein System aus dieser Klasse an Systemen herangezogen werden.

7 Ableitung der Mehrwerte

Das Potenzial der voraussagenden Instandhaltung in der Produktion kann durch die Steigerung der technischen Verfügbarkeit beziffert werden, hierzu werden die Gütemaße Ausfallrate PR und Trefferrate TPR aus Kapitel 6 herangezogen. Basierend auf verschiedenen zugrunde liegenden technischen Verfügbarkeiten wird das entsprechende Potenzial in diesem Kapitel berechnet. Anschließend wird der Gewinn mit Hilfe exemplarischer Werte für verschiedene technische Verfügbarkeiten bestimmt. Somit zeigt dieses Kapitel den wirtschaftlichen Mehrwert einer voraussagenden Instandhaltung basierend auf der gesteigerten technischen Verfügbarkeit und in einem ersten Beispiel gemessen am Gewinn auf.

7.1 Vorgehensweise zur Berechnung der gesteigerten technischen Verfügbarkeit

Eine wirtschaftliche Betrachtung der durch die voraussagende Instandhaltung gesteigerten technischen Verfügbarkeit kann die positiven Effekte einer solchen Instandhaltungsstrategie für die Industrie deutlicher zu Tage bringen. Die Vorgehensweise zur Berechnung der technischen Verfügbarkeit V_T wurde bereits in Kapitel 2.1.1 erläutert, dabei gilt folgende Formel⁴⁶:

$$V_T = \left(1 - \frac{T_T}{T_B}\right) * 100 \% \quad (7.1)$$

In der Produktion liegt in der Regel bereits die Kennzahl der technischen Verfügbarkeit V_T vor, sodass die Bestimmung dieser Kennzahl nicht weiter vertieft werden muss.

Aus der technischen Verfügbarkeit V_T ergibt sich auch die technische Ausfallrate A_T .

$$A_T = 100 \% - V_T \quad (7.2)$$

Bei einer technischen Verfügbarkeit V_T von 90 %, ergibt sich somit eine technische Ausfallrate A_T von 10 %.

$$A_T = 100 \% - 90 \% = 10 \%$$

Ausgehend von einer bisherigen technischen Ausfallrate A_T lässt sich, mittels der nachfolgenden, eigenständig erarbeiteten Formel 7.3, die Steigerung der technischen Verfügbarkeit infolge einer voraussagenden Instandhaltung (Predictive Maintenance, kurz PM) $\Delta V_{T,PM}$ errechnen.

⁴⁶ T_T beschreibt die technische Ausfallzeit und T_B die Belegungszeit, wobei T_B die Summe der Nutzungszeit T_N , der organisatorischen Ausfallzeit T_O , der technischen Ausfallzeit T_T und der Wartungszeit T_W ist. Resultierend ergibt sich: $T_B = T_N + T_O + T_T + T_W$ [22].

$$\Delta V_{T,PM}(t) = \frac{A_T * TPR(t)}{PR(t)} \quad (7.3)$$

Der Zähler enthält die technische Ausfallrate A_T , die zusätzlich mit dem prozentualen Faktor der vermeidbaren Ausfälle – also der Trefferrate TPR^{47} – multipliziert wird. Durch Division durch die Ausfallrate PR^{48} berechnet sich daraus die gesteigerte technische Verfügbarkeit als Folge einer voraussagenden Instandhaltung $\Delta V_{T,PM}$ für einen beliebigen Zeitpunkt t . Es wird somit eine neue Normierung eingeführt.

Würde das Betriebsende (in diesem Fall 131.400 Stunden bzw. 15 Jahre) betrachtet werden, so vereinfacht sich die Formel wie folgt, da die Ausfallrate PR in diesem Fall 100 % beträgt:

$$\Delta V_{T,PM}(t = 131.400 h) = \frac{A_T * TPR(t)}{1} = A_T * TPR(t) \quad (7.4)$$

Basierend auf der technischen Verfügbarkeit V_T und der gesteigerten technischen Verfügbarkeit als Folge von voraussagender Instandhaltung $\Delta V_{T,PM}$ lässt sich die neue gesteigerte Verfügbarkeit $V_{T,PM}$ errechnen. Hierzu sind lediglich die aktuelle technische Verfügbarkeit V_T und die Steigerung der technischen Verfügbarkeit $\Delta V_{T,PM}$ miteinander zu addieren.

$$V_{T,PM}(t) = V_T + \Delta V_{T,PM}(t) \quad (7.5)$$

Wenn eine Steigerung der technischen Verfügbarkeit $V_{T,PM}$ erzielt wird, so hat dies gegebenenfalls auch Auswirkungen auf die Wartungszeit T_W bzw. die wartungsbedingte Ausfallrate⁴⁹ A_W und der Organisation von Wartungsmaßnahmen im Allgemeinen: Zwar reduziert sich die technische Ausfallrate A_T , aber nun müsste häufiger eine geplante Instandhaltung erfolgen, wodurch sich die Wartungszeit T_W erhöht. Im Rahmen dieser Arbeit werden die Auswirkungen auf die Wartungszeit T_W und damit auf die wartungsbedingte Ausfallrate A_W sowie die sich daraus ergebenden organisatorischen Veränderungen nicht weiter analysiert. Denn die geplante Instandhaltung erfolgt in der Regel an Wochenenden sowie je nach Schichtplanung innerhalb freier Produktionsschichten und fällt damit nicht in die Belegungszeit T_B .

⁴⁷ nach Formel 4.22 $TPR(t) = TP(0 \leq t) / P(0 \leq t < B)$

⁴⁸ nach Formel 4.24 $PR(t) = P(0 \leq t) / P(0 \leq t < B)$

⁴⁹ wartungsbedingte Ausfallrate: $A_W = \frac{T_W}{T_B} * 100,0 \%$, wobei T_W für die Wartungszeit und T_B für die Belegungszeit steht.

7.2 Technische Verfügbarkeit – Stichprobe 1

Auf Basis der bisherigen Ergebnisse für die Stichprobe 1 – also der Polynomialverteilung⁵⁰ im ersten Szenario – (siehe Abschnitt 6.3.2) würden sich folgende Ergebnisse für die Steigerung der technischen Verfügbarkeit $\Delta V_{T,PM}$ ergeben. Ausgehend von einer bisherigen Verfügbarkeit V_T von 90 % würde eine Verbesserung der Verfügbarkeit $\Delta V_{T,PM}$ um 4,6 % nach zwei Jahren $t = 17.520 h$ erzielt werden.

$$\text{Erg. 7.1: } \Delta V_{T,PM}(t = 17.520 h) = \frac{A_T * TPR(t)}{PR(t)} = \frac{10 \% * 4,4 \%}{9,5 \%} \approx 4,6 \%$$

Hierbei wurde die Ausfallrate A_T von 10 %, die in Abschnitt 6.3.2 bestimmte Näherungsfunktion der Trefferrate $TPR(t = 17.520 h) = 0,044$ und der Ausfallrate $PR(t = 17.520 h) = 0,095$ verwendet. Damit läge die gesteigerte technische Verfügbarkeit in Folge von voraussagender Instandhaltung $V_{T,PM}$ bei 94,6 %.

$$\text{Erg. 7.2: } V_{T,PM}(t = 17.520 h) = 90,0 \% + 4,6 \% \approx 94,6 \%$$

Weiterführend stellt sich die Frage, inwieweit sich die gesteigerte technische Verfügbarkeit je nach Zeitpunkt t verändert. Nach Ablauf der Betriebsdauer B von 15 Jahren bzw. 131.400 Stunden wäre die gesteigerte technische Verfügbarkeit $V_{T,PM}$ sogar auf 98,7 % angestiegen, sofern voraussagende Instandhaltung eingesetzt worden wäre.

$$\text{Erg. 7.3: } \Delta V_{T,PM}(t = 131.400 h) = 10,0 \% * 0,868 \approx 8,7 \%$$

$$\text{Erg. 7.4: } V_{T,PM}(t = 131.400 h) = 90,0 \% + 8,7 \% \approx 98,7 \%$$

Dieser Anstieg von 8,7 % gegenüber der bislang realisierten technischen Verfügbarkeit V_T ohne Einsatz voraussagender Instandhaltung geht dabei davon aus, dass der bisherige Verlust der Verfügbarkeit, bislang 10 % der Produktionszeit, vollständig durch die technische Ausfallrate A_T bestimmt wird. Dies ist im Sinne der Definition der technischen Verfügbarkeit eine valide Annahme.

Ausgehend von sechs verschiedenen gewählten technischen Verfügbarkeiten $V_T = [50 \%, 75 \%, 85 \%, 90 \%, 95 \%, 100 \%]$ zeigt das Bild 7.1 die Steigerung der technischen Verfügbarkeit $V_{T,PM}$ über der Zeit t als Folge voraussagender Instandhaltung. Bild 7.1 geht dabei von Szenario 1 aus, dass kontinuierlich Daten erfasst und analysiert werden. Wenn keine Datenerfassung und damit auch Datenanalyse erfolgen würde, so würde die gesteigerte technische Verfügbarkeit $V_{T,PM}$ auf den ursprünglichen Wert einer reaktiven Instandhaltung zurückgehen.

Die verschiedenen gesteigerten technischen Verfügbarkeiten $V_{T,PM}$ zeigen das größere, positive Auswirkungen erzielt werden können, wenn die ursprüngliche

⁵⁰ $f(x) \Rightarrow TPR(t) = 1,61 \cdot 10^{-11} t^2 + 5,00 \cdot 10^{-6} t - 0,03$. Für den Wertebereich der Funktion ist zu beachten, dass $W_f = [0; 1]$ ist.

$f(x) \Rightarrow PR(t) = 8,01 \cdot 10^{-6} t - 0,05$

technische Verfügbarkeit im Vorfeld gering gewesen wäre bzw. gegebenenfalls sogar wie bei einer bislang realisierten technischen Verfügbarkeit V_T von 50 % geradezu schlecht gewesen wäre. Ein solche schlechte technische Verfügbarkeit $V_T = 50 \%$ ist genauso wie eine technische Verfügbarkeit $V_T = 75 \%$ in der Praxis allerdings eher selten anzutreffen und stellt damit eher ein negatives Extrembeispiel dar. Verfügbarkeiten, die in der Produktion üblicher sind, sind $V_T = [85 \%, 90 \%, 95 \%]$. Für diese drei angenommenen technischen Verfügbarkeiten V_T lässt sich immer noch eine merkliche, wenn auch abnehmende Steigerung der technischen Verfügbarkeit aufgrund der voraussagenden Instandhaltung $\Delta V_{T,PM}$ erkennen. Bei einer ursprünglichen Verfügbarkeit von $V_T = 100 \%$ kann jedoch keine Verbesserung mehr erzielt werden, da diese ja schon optimal ist. Diese technische Verfügbarkeit V_T wird in der Praxis in aller Regel jedoch nicht auftreten und stellt eher einen rein hypothetischen Fall dar.

Vergleicht man die drei technischen Verfügbarkeiten $V_T = [85 \%, 90 \%, 95 \%]$, so ist zum Betriebsende hin nur noch ein geringer Unterschied zwischen deren jeweiligen gesteigerten technischen Verfügbarkeit infolge der voraussagenden Instandhaltung $V_{T,PM}$ erkennbar. Damit erzielt der Einsatz der voraussagenden Instandhaltung weitestgehend unabhängig von der ursprünglichen technischen Verfügbarkeit bei $V_T = [85 \%, 90 \%, 95 \%]$ eine ähnliche gesteigerte technische Verfügbarkeit zu Betriebsende $V_{T,PM}(t = 131.400 \text{ h})$.

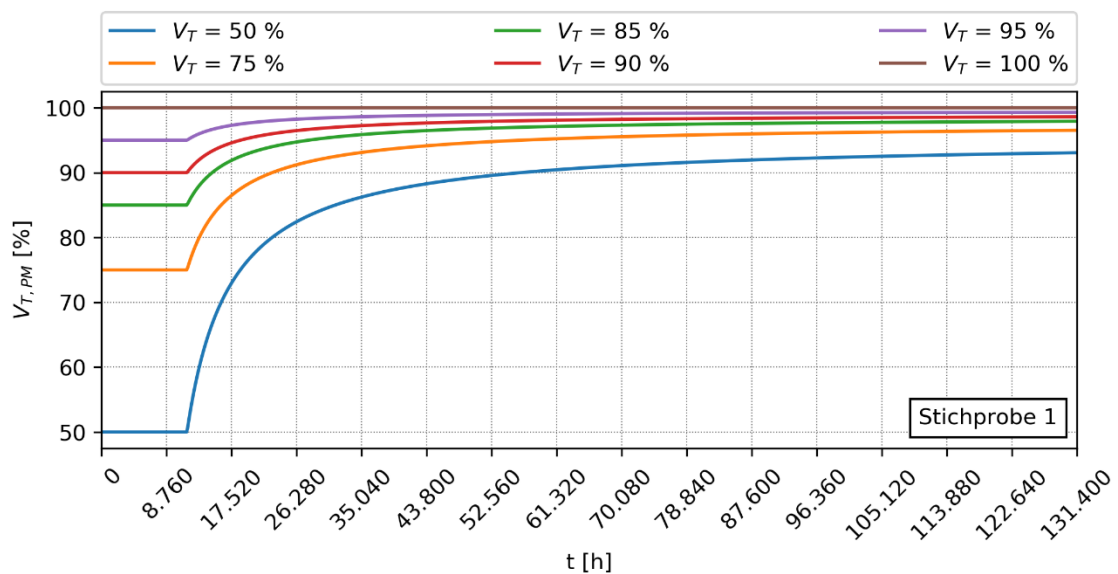


Bild 7.1: Berechnete technische Verfügbarkeit im Falle von voraussagender Instandhaltung für die Stichprobe 1 entsprechend Szenario 1: Dargestellt sind sechs verschiedene gesteigerte technische Verfügbarkeiten $V_{T,PM}$ (in verschiedenen Farben) basierend auf den verschiedenen angenommenen technischen Verfügbarkeiten V_T . Für die Berechnung wurden die Ergebnisse für die Ausfallrate PR und Trefferrate TPR der Stichprobe 1 entsprechend Szenario 1 verwendet. Die Datenerfassung erfolgt über den kompletten Betriebszeitraum $t_B = [0, 131.400] \text{ h}$.

7.3 Exemplarische Berechnung des Gewinns – Stichprobe 1

Nachfolgend wird der sich aus der gesteigerten technischen Verfügbarkeit $V_{T,PM}$ ergebende Gewinn in Euro ermittelt, um die Ergebnisse der Arbeit noch greifbarer zu machen. Hierzu muss zuerst der Erlös E nach Formel 7.6 berechnet werden, wobei p den Gewinn pro Stück und x die Stückzahl beschreibt.

$$E(x) = p * x \quad (7.6)$$

Die Stückzahl x ergibt sich aus der Betriebsdauer B und der Taktzeit pro Stück T_x . Im weiteren Verlauf werden die folgenden exemplarischen, aber realitätsnahen Werte angenommen:

$$B = 15 \text{ Jahre} * 365 \text{ Tage} * 24 \text{ h} * 60 \text{ min} = 7.884.000 \text{ min}$$

$$T_x = 15 \text{ min/Stück}$$

$$x = \frac{B}{T_x} = 525.600 \text{ Stück}$$

$$p = 10,00 \text{ €/Stück}$$

Damit ergibt sich ein Erlös E von rund 5,3 Mio. € über die Betriebsdauer B von 7.884.000 Minuten.

Erg. 7.5:

$$\begin{aligned} E(525.600 \text{ Stück}) &= 10,00 \text{ €/Stück} * 525.600 \text{ Stück} \\ &= 5.256.000 \text{ €} \end{aligned}$$

Dieser entspricht jedoch noch nicht dem Gewinn G . Hierzu sind noch die verschiedenen Kosten vom Erlös E zu subtrahieren. Dafür wird die Kostenfunktion $K(x)$ wie folgt aufgestellt, wobei k die variablen Kosten und F die Fixkosten beschreibt.

$$K(x) = k * x + F \quad (7.7)$$

Die variablen Kosten k je Stück x werden in dieser Vorgehensweise anhand der technischen Ausfallrate A_T berechnet. Diese entspricht $A_T = 100 \% - V_T$ (siehe Formel 7.2).

$$k = \frac{p * A_T}{100 \%} \quad (7.8)$$

Ausgehend von den Ergebnissen aus Abschnitt 7.2 wird beispielhaft eine realitätsnahe technische Verfügbarkeit V_T von 90 % angenommen, sodass die technische Ausfallrate $A_T = 10 \%$ beträgt. Dies würde für die variablen Kosten k Kosten von 1 € pro Stück bedeuten.

Erg. 7.6:

$$k_{\text{ohne PM}} = \frac{10,00 \text{ €/Stück} * 10 \%}{100 \%} = 1,00 \text{ €/Stück}$$

Für die Fixkosten F wird kein Wert angenommen, da dies den Vergleich zwischen dem reaktiven und vorhersagenden Ansatz der Instandhaltung vereinfacht. Somit

entstehen bei einer reaktiven Instandhaltung, also ohne voraussagende Instandhaltung, rund 0,5 Mio. € an Kosten $K_{ohne PM}$ über eine komplette Betriebsdauer $B = 131.400 h$.

$$\begin{aligned} \text{Erg. 7.7: } K_{ohne PM}(525.600 \text{ Stück}) &= 1,00 \text{ €/Stück} * 525.600 \text{ Stück} + 0 \text{ €} \\ &= 525.600 \text{ €} \end{aligned}$$

Vergleichend hierzu sind es bei der voraussagenden Instandhaltung in Folge der geringeren Ausfallrate $A_{T,PM} = 100 \% - 98,7 \% = 2,3 \%$ (vgl. Ergebnis 7.4) nur rund 0,15 Mio. € an Kosten K_{PM} über die Betriebsdauer B von 15 Jahren. Dies entspricht damit einer Kostensenkung vom 71 % gegenüber den Kosten $K_{ohne PM}$ im Falle einer reaktiven Instandhaltung.

Die Kosten im Falle einer voraussagenden Instandhaltung K_{PM} berechnen sich dabei unter Berücksichtigung, dass sich die technische Verfügbarkeit $V_{T,PM}$ über die Zeit t wie folgt steigert. Der stündliche, jährliche oder auch monatliche Wert der in Folge der voraussagenden Instandhaltung gesteigerten technischen Verfügbarkeit $V_{T,PM}$ kann zu jedem Zeitpunkt t durch die Näherungsverteilung der Trefferrate $TPR(t)$ bestimmt werden. Da keine lineare Verteilung vorliegt, kann das Ergebnis jedoch nicht einfach z.B. durch 15 Jahre dividiert werden, um die jährliche Einsparung zu erhalten. Daher sind die variablen Kosten je Stück $k_{PM}(x)$ einzeln zu berechnen und anschließend zu summieren. Da sich die technische Ausfallrate bei Anwendung einer voraussagenden Instandhaltung $A_{T,PM}$ abhängig von der Zeit t ändert, muss die Stückzahl x mittels der Taktzeit pro Stück T_x in einen zeitlichen Bezug gesetzt werden. Dies erfordert eine Anpassung der Formel 7.8 wie folgt:

$$k_{PM}(x) = \frac{p * A_{T,PM}(t = T_x * x)}{100 \%} \quad (7.9)$$

Fortführend sind die variablen Kosten k_{PM} der Kostenfunktion $K_{PM}(x)$ je Stück individuell zu berechnen und aufzusummieren:

$$K_{PM}(x) = \left(\sum_{i=0}^x k_{PM}(i) \right) + F \quad (7.10)$$

Exemplarisch werden für zwei Zeitpunkte t bzw. deren Stückzahl x die variablen Kosten k_{PM} berechnet. Für die Zeitpunkte $t = 17.520 h$ und $t = 131.400 h$ wurde die gesteigerte technische Verfügbarkeit bereits im vorhergehenden Abschnitt 7.2 berechnet und in den Ergebnissen 7.2 und 7.4 wiedergegeben. Für die Umrechnung des Zeitpunkts t in die Stückzahl x gilt:

$$t = T_x * x \Leftrightarrow x = \frac{t}{T_x} \quad (7.11)$$

Erg. 7.8: $k_{PM}(t = 17.520 h \Leftrightarrow x = 70.080 \text{ Stück})$

$$\begin{aligned} &= \frac{p * (100 \% - V_{T,PM}(t = 17.520 h))}{100 \%} \\ &= \frac{10,00 \text{ €/Stück} * (100 \% - 94,6\%)}{100 \%} \\ &= 0,54 \text{ €/Stück} \end{aligned}$$

Erg. 7.9: $k_{PM}(t = 131.400 h \Leftrightarrow x = 525.600 \text{ Stück})$

$$\begin{aligned} &= \frac{p * (100 \% - V_{T,PM}(t = 131.400 h))}{100 \%} \\ &= \frac{10,00 \text{ €/Stück} * (100 \% - 98,7\%)}{100 \%} \\ &= 0,13 \text{ €/Stück} \end{aligned}$$

Damit ergibt sich für die Kostenfunktion $K_{PM}(x)$ in Folge voraussagender Instandhaltung ohne die Berücksichtigung von Fixkosten F :

Erg. 7.10:
$$K_{PM}(525.600 \text{ Stück}) = \left(\sum_{i=0}^x k_{PM}(i) \right) + F = 156.875 \text{ €}$$

Der Gewinn G berechnet sich nach Formel 7.12 zu:

$$G(x) = E(x) - K(x) \tag{7.12}$$

Daraus ergibt sich somit ein Gewinn $G_{ohne PM}$ von 4,7 Mio. € im Falle einer reaktiven Instandhaltung und ein Gewinn G_{PM} in Höhe von 5,1 Mio. € unter Einsatz einer voraussagenden Instandhaltung. Dies ist eine Steigerung des Gewinns um 0,4 Mio. € über 15 Jahre für ein einzelnes beispielhaft betrachtetes, aber realitätsnahes System. Insgesamt würde der Gewinn über die Betriebsdauer $B = 15 \text{ Jahre}$ damit also um ca. 8 % gesteigert werden. Dies wäre eine Gewinnsteigerung, die in der heutigen hochtechnisierten und bereits hochgradig optimierten Produktion kaum noch auf anderem Wege erreicht werden könnte. Gerade hieran zeigt sich, warum die Implementierung und die Anwendung von voraussagender Instandhaltung einen wesentlichen Einfluss auf die zukünftige Wettbewerbsfähigkeit eines Unternehmens haben kann (vgl. Kapitel 1 und 2).

Erg. 7.11:
$$\begin{aligned} G_{ohne PM}(525.600 \text{ Stück}) &= 5.256.000 \text{ €} - 525.600 \text{ €} \\ &\approx 4,7 \text{ Mio. €} \end{aligned}$$

Erg. 7.12:
$$\begin{aligned} G_{PM}(525.600 \text{ Stück}) &= 5.256.000 \text{ €} - 156.875 \text{ €} \\ &\approx 5,1 \text{ Mio. €} \end{aligned}$$

In Bild 7.2 sind die Kosten $K(x)$ ohne Einsatz von und mit Einsatz von voraussagender Instandhaltung vergleichend für jeden Zeitpunkt t abgebildet. Zusätzlich wird auch der Gewinn $G(x)$ für beide Vorgehensweisen basierend auf dem jeweiligen Erlös $E(x)$ dargestellt. Daraus erkennbar ist auch die zunehmende Differenz zwischen den beiden zugehörigen Kosten-Verteilungen $K(x)$.

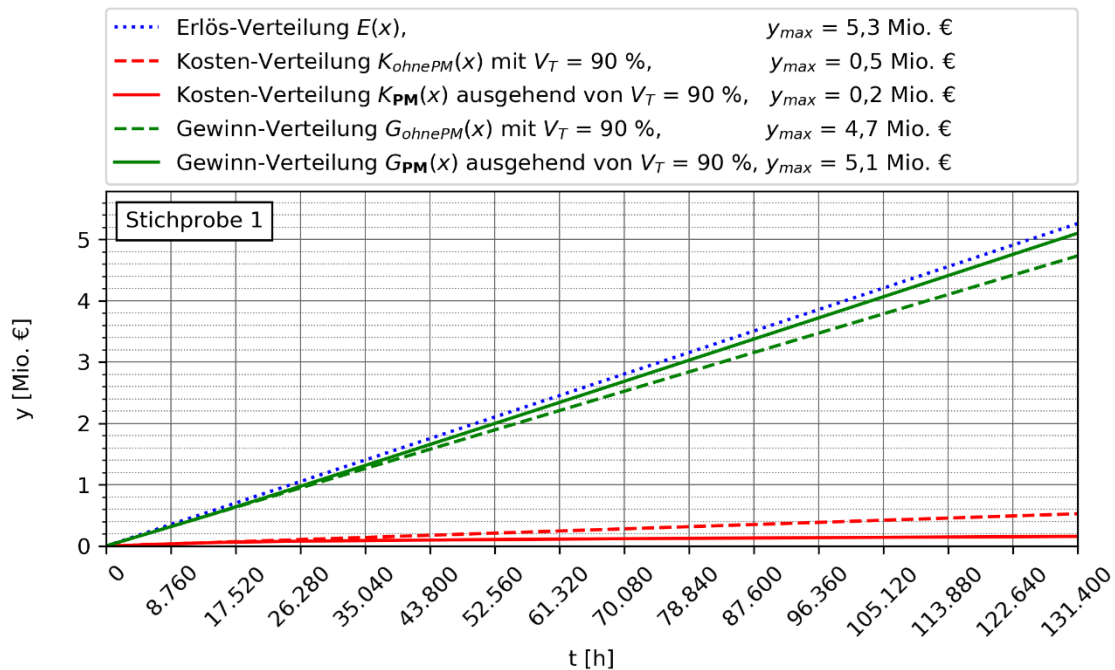


Bild 7.2: Berechnete Erlös-, Kosten- und Gewinn-Verteilungen ohne Fixkosten für die Stichprobe 1 entsprechend Szenario 1: Dargestellt ist die Erlös-Verteilung $E(x)$ (in Blau gepunktet), die Kosten-Verteilungen im Falle einer reaktiven Instandhaltung $K_{ohnePM}(x)$ (in Rot gestrichelt) und im Falle einer voraussagenden Instandhaltung $K_{PM}(x)$ (in Rot) sowie die entsprechenden Gewinn-Verteilungen $G_{ohnePM}(x)$ (in Grün gestrichelt) und $G_{PM}(x)$ (in Grün) basierend auf der technischen Verfügbarkeit $V_T = 90\%$ bzw. der gesteigerten technischen Verfügbarkeit $V_{T,PM}$.

Zusätzlich zu den bisherigen Kosten k können beim Einsatz einer voraussagenden Instandhaltung bzw. beim Einsatz von Predictive Maintenance (PM) auch Fixkosten F für die Nutzung einer Softwarelösung entstehen. Diese muss zum Beispiel implementiert, betrieben und weiterentwickelt werden. Je mehr Systeme eine solche Software nutzen, umso geringer können jedoch die Fixkosten F pro System ausfallen. Im Folgenden wird von Fixkosten $F = 750.000$ € über die Betriebsdauer B ausgegangen. Dies ist lediglich ein beispielhafter Wert und hat keinen Bezug zu einem real-existierenden Beispiel.

$$\text{Erg. 7.13: } K_{PM}(525.600 \text{ Stück}) = 156.875 \text{ €} + 750.000 \text{ €} = 906.875 \text{ €}$$

Bild 7.3 stellt ebenso wie Bild 7.2 den Erlös $E(x)$, die Kosten $K(x)$ und den Gewinn $G(x)$ dar, allerdings berücksichtigt die Kostenfunktion im Falle einer voraussagenden Instandhaltung $K_{PM}(x)$ zusätzliche Fixkosten für die Umsetzung derselben in Höhe $F = 750.000$ €. Dadurch reduziert sich der zu erwartende Gewinn bei Einsatz einer voraussagenden Instandhaltung $G_{PM}(x)$ und würde sogar unter dem Gewinn ohne Einsatz von PM $G_{ohnePM}(x)$ liegen. Resultierend lässt sich

somit festhalten, dass die Kosten $K_{PM}(x) < K_{ohne PM}(x)$ sein sollten, damit ein finanzieller Mehrwert erzielt wird.

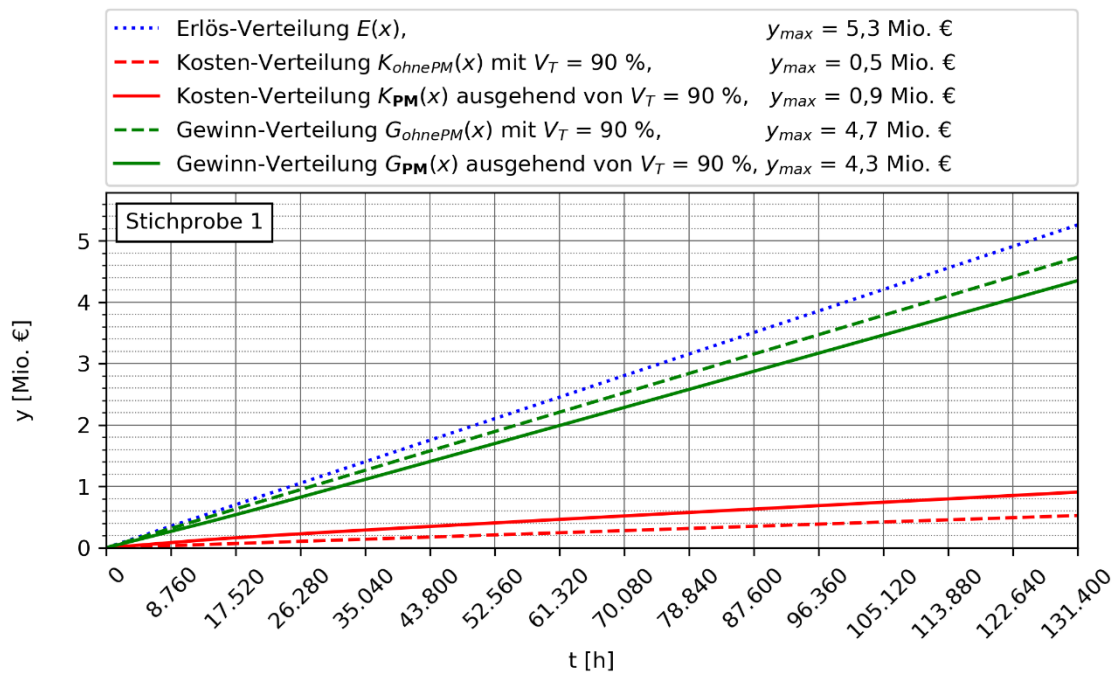


Bild 7.3: Berechnete Erlös-, Kosten- und Gewinn-Verteilungen mit Fixkosten für die Stichprobe 1 entsprechend Szenario 1: Dargestellt ist die Erlös-Verteilung $E(x)$ (in Blau gepunktet), die Kosten-Verteilungen im Falle einer reaktiven Instandhaltung $K_{ohne PM}(x)$ (in Rot gestrichelt) und im Falle einer voraussagenden Instandhaltung $K_{PM}(x)$ (in Rot) sowie die entsprechenden Gewinn-Verteilungen $G_{ohne PM}(x)$ (in Grün gestrichelt) und $G_{PM}(x)$ (in Grün) basierend auf der technischen Verfügbarkeit $V_T = 90\%$ bzw. der gesteigerten technischen Verfügbarkeit $V_{T,PM}$. Als Fixkosten wurden $F = 750.000$ € über den kompletten Betriebszeitraum $t_B = [0, 131.400]$ h angenommen.

Die genannten Auswirkungen beziehen sich auf das freigewählte Beispiel. Die Ergebnisse sind weiterhin abhängig von der Betriebsdauer B , der Taktzeit pro Stück T_x , dem Gewinn pro Stück p und natürlich maßgeblich von der ursprünglichen technischen Verfügbarkeit V_T .

7.4 Fazit

Das Kapitel hat gezeigt, wie aus den Gütemaßen Ausfallrate PR und Trefferrate TPR quantitative Mehrwerte abgeleitet werden können. Hierzu wurde zuerst die gesteigerte technische Verfügbarkeit im Falle des Einsatzes einer voraussagenden Instandhaltung auf Basis von verschiedenen, bislang realisierten technischen Verfügbarkeiten im Falle der Anwendung einer reaktiven Instandhaltung berechnet. Mittels dieser Kennzahl war es anschließend möglich, für ein exemplarisches System die Kosten sowie den Erlös und damit auch den Gewinn zu kalkulieren. Dabei hat sich herausgestellt, dass ohne Berücksichtigung von Fixkosten für die Implementierung, den Betrieb und die Weiterentwicklung einer voraussagenden Instandhaltung eine deutliche Kostenreduktion (in dem gewählten Beispiel Kosteneinsparung von rund 71 %, also $\sim 0,4$ Mio.€ pro exemplarisch betrachtetem System, über 15 Jahre bei einer ursprünglichen technischen Verfügbarkeit von 90 %) erzielt werden kann.

Die verwendeten Gütemaße – Ausfallrate PR , Trefferrate TPR , Nicht-Alarmrate FNR – sind in bisherigen Veröffentlichungen meist im Rahmen einer Klassifikation von Bedeutung. Jedoch sind die Gütemaße für die Industrie nur untergeordnete Informationen. Dieses Kapitel hat deutlich gemacht, dass durch den Einsatz von voraussagender Instandhaltung die technische Verfügbarkeit maßgeblich und damit indirekt die Wettbewerbsfähigkeit gesteigert werden kann. Bereits Bild 2.8, die Zusammenfassung einer Studie von Roland Berger über die Mehrwerte von Predictive Maintenance, hat diesen Nutzen von voraussagender Instandhaltung in Form von der Erhöhung der Anlagenverfügbarkeit dargelegt. Es darf aber nicht verschwiegen werden, dass gerade auch der finanzielle Mehrwert des Einsatzes von voraussagender Instandhaltung davon abhängt, in welchem Maße zusätzliche (Fix-)Kosten für die Implementierung, den Betrieb und die Weiterentwicklung derselben anfallen.

Resultierend lässt sich festhalten, dass der Einsatz von voraussagender Instandhaltung für jeden einzelnen Fall bewertet werden sollte. Eine pauschale Ableitung des Gewinns ist aufgrund der individuellen Gewinnfunktion, basierend auf den jeweiligen Kosten des Systems, nicht ohne weiteres möglich. Diese Arbeit ermöglicht jedoch eine aufwandsarme Schätzung des Gewinns. Durch die Bestimmung der allgemeingültigen Verteilung der gesteigerten technischen Verfügbarkeit in Folge des Einsatzes von Predictive Maintenance kann nun für weitere ähnliche Systeme (entsprechend dem Stichprobenumfang) der Unterschied zwischen den anfallenden Kosten ohne und mit Predictive Maintenance abgeleitet werden. Damit kann der Einsatz von voraussagender Instandhaltung künftig bereits im Vorfeld umfassend individuell bewertet und eine Obergrenze für dessen Mehrwert ermittelt werden.

8 Zusammenfassung und Ausblick

Abschließend wird in diesem Kapitel eine Zusammenfassung zum Inhalt und Ergebnis der Arbeit gegeben. Danach wird fokussiert der wissenschaftliche Beitrag der Arbeit beschrieben. Im Ausblick werden weitere Themenfelder, welche im Rahmen der Arbeit von Relevanz sind, angeführt.

8.1 Zusammenfassung

Die vorliegende Arbeit hat die Möglichkeiten und Herausforderungen zur Realisierung einer voraussagenden Instandhaltung (Predictive Maintenance) in der diskreten Fertigung von Motoren-Komponenten herausgearbeitet. Motiviert aus den neuen Möglichkeiten zur Steigerung der Wettbewerbsfähigkeit und damit dem Einsatz von Predictive Maintenance, ausgeführt in Abschnitt 1.1, setzt sich die Arbeit in Abschnitt 1.2 das Ziel, eine datengetriebene Methodik für die Bestimmung der theoretischen Güte von Vorhersagemodellen für Ausfälle von Komponenten eines Produktionssystems auszuarbeiten, zu untersuchen und festzulegen. Die dazu in Kapitel 4 entwickelte Methodik erlaubt es, verschiedene Szenarien beim Einsatz von Predictive Maintenance unter Verwendung von Vorhersagemodellen zu bewerten und zu vergleichen.

Zur Erreichung der Zielstellung war jedoch erstmal eine Definition der Grundlagen und eine Analyse des Standes der Technik – zusammengestellt in Kapitel 2 – erforderlich. Fußend auf den Kenntnissen zu Produktionssystemen in der Motoren-Komponenten-Produktion, den bekannten Instandhaltungsstrategien und Werkzeugen auf dem Markt wurde im Anschluss die Notwendigkeit der Arbeit erörtert. Resultierend aus der Marktanalyse in Abschnitt 2.2, dass heutige Topunternehmen aus dem Werkzeugmaschinenbau noch keine schlüsselfertige Lösung für voraussagende Instandhaltung anbieten, wurde in Abschnitt 0 dementsprechend ausgearbeitet, welche Herausforderungen beim Einsatz von Predictive Maintenance bezüglich Anwendung statistischer Methoden und ausreichend großer Datenbasis noch bestehen.

Für die erste Herausforderung – der Anwendung von statistischen Methoden – wurde demzufolge ein Systementwurf für die Generierung von Vorhersagemodellen in Kapitel 3 erstellt. Dieser stellt auf verständliche Weise vor, wie eine Auswahl und Verfügbarkeit von Daten erreicht werden kann, welche Optionen zur Kennzeichnung von Datenpunkten bestehen und schließlich mit welcher Vorgehensweise ein Vorhersagemodell trainiert werden kann. Deutlich wurde dabei, dass nur sich wiederholende Ausfälle detektiert und damit vorhergesagt werden können, da historische Daten eine zwingende Voraussetzung für das Antrainieren von Vorhersagemodellen eines üblichen Black-Box-Modells – also bei einer datengetriebenen Modellierung – sind.

Die zweite Herausforderung – die ausreichend große Datenbasis – stellte dann den Hauptfokus der Arbeit dar. Basierend auf der Fragestellung, wie eine ausreichend große Datenbasis überhaupt definiert sein kann, wurde zuerst in Kapitel 4 eine Methode zur Bestimmung der bestmöglichen Güte von Vorhersagemodellen bei optimaler Anwendung von statistischen Methoden entwickelt. Danach wurde in Kapitel 5 für das System BAZ 1 und in Kapitel 6 für die Klasse der Produktionssysteme die bestmögliche Güte von Vorhersagemodellen mittels der selbstentwickelten Methodik aus Kapitel 4 bestimmt und in verschiedenen Szenarien untersucht. Die Szenarien stellen dabei drei mögliche Ausgangssituationen dar:

- **Szenario 1** definiert sich durch eine dauerhafte Datenerfassung und -analyse vom Betriebsbeginn $t_{data} = 0$ an.
- **Szenario 2** grenzt sich von Szenario 1 durch einen späteren Beginn der Datenerfassung $t_{data} > 0$ und damit auch der Datenanalyse ab.
- **Szenario 3** untersucht basierend auf Szenario 1 die Nutzung von Daten aus baugleichen Systemen.

Die Ergebnisse der möglichen Güte von Vorhersagemodellen mit idealisiertem Klassifikator sind in allen drei Szenarien sowohl für das System BAZ 1 als auch für die umfassende Klasse der Produktionssysteme (genannt Stichprobe 1) vielversprechend, womit ein Einsatz von Predictive Maintenance als durchaus sinnvoll zu interpretieren ist.

Im **Szenario 1** kann mit den getroffenen Annahmen für die Stichprobe 1 (siehe Abschnitt 6.3.2) eine gewinnbringende Lernrate LPR von 100 % nach 15 Jahren der Betriebszeit erreicht werden kann. Die damit einhergehenden Gütemaße liegen für die Trefferrate TPR bei $86,7 \pm 7,5$ % und komplementär bei der Nicht-Alarmrate FNR bei $23,9 \pm 5,4$ % zum Betriebsende, also nach 15 Jahren.

Durch einen späteren Beginn der Datenerfassung (z.B. Beginn ab 4 Jahren) würde sich bei **Szenario 2** der für die Erfassung der Ausfalldaten bzw. Trainingsdaten sinnvoll nutzbare Zeitraum verkürzen. Trotzdem hat das Szenario gezeigt, dass immer noch für jede Komponente der Ausfall erlernt wird. Die Ergebnisse für Stichprobe 1 im Szenario 2 mit einem gewählten Beginn der Datenerfassung ab 4 Jahre (Trefferate TPR bei $65,6 \pm 4,9$ %, Nicht-Alarmrate FNR bei $39,5 \pm 3,3$ % aus Anhang A.2) lassen dennoch die Ableitung zu, dass der Einsatz von Predictive Maintenance auch zu einem späteren Betriebszeitpunkt des Systems wertvoll ist.

Die Ergebnisse für **Szenario 3** sind jedoch sehr viel besser, denn durch den Übertrag des auf den erlernten Ausfällen eines anderen, baugleichen Systems basierenden Vorhersagemodells ist idealisiert betrachtet eine Steigerung der Trefferrate TPR auf 100 % und damit eine Reduktion der Nicht-Alarmrate FNR auf 0 % vorstellbar. Dies wurde einmalig anhand des Systems BAZ 1 im Kapitel 5.4 demonstriert.

Zusammengefasst konnten für alle drei Szenarien die Auswirkungen der Datenbasis auf die Güte des Vorhersagemodells bestimmt werden. Durch die Einführung des Gütemaßes Lernrate LPR erfolgte die Definition einer ausreichenden Datenbasis.

Diese Arbeit hat jedoch nicht nur eine Vorgehensweise zur Bestimmung der bestmöglichen Datenbasis für Systeme in der Motoren-Komponenten-Produktion entwickelt, sondern auch die Gütwerte für beliebige Zeitpunkte in der Betriebsphase berechnet.

Durch die Verwendung einer breiten Stichprobe an Produktionssystemen (Stichprobe 1) zur Bestimmung der generalisierbaren Ergebnisse und der Berechnung der zugehörigen statistischen Unsicherheit für die Ergebnisse, wie in Kapitel 6 geschehen, ist es ab jetzt möglich vielfältige Abschätzungen auch für andere Produktionsstätten vorzunehmen. In diesem Kapitel wurde zusätzlich mit Hilfe des unabhängigen Systems BAZ 1 validiert, dass die Ergebnisse allgemein gültig für die Klasse der Produktionssysteme sind und die Ergebnisse als Abschätzung auf andere Produktionssysteme einfach übertragbar sind. Eine darauf basierende, umfangreiche Abschätzung der (quantitativen) Mehrwerte von Predictive Maintenance hat bereits in Kapitel 7 stattgefunden. Laut dieser konnte die technische Verfügbarkeit im Falle der Stichprobe 1 und einer ursprünglichen technischen Verfügbarkeit von 90 % auf 98,7 %, also abhängig vom Zeitpunkt um bis zu 8,7 % gesteigert werden, sofern statt einer reaktiven eine voraussagende Instandhaltungsstrategie eingesetzt werden würde. Damit ergibt sich für ein fiktives, realitätsnahes Beispiel eine denkbare Kostenreduktion um bis zu 71 % und eine daraus abgeleitete Gewinnsteigerung von bis zu 8 %. Damit hat die abschließende Begutachtung der Mehrwerte, den Beitrag von Predictive Maintenance zur Wettbewerbssteigerung in Form von Steigerung der technischen Verfügbarkeit und Reduzierung der Kosten bzw. Erhöhung des Gewinns deutlich bekräftigt.

8.2 Wissenschaftlicher Beitrag dieser Arbeit

Diese Arbeit liefert nicht nur einen Beitrag für die Wissenschaft, sondern treibt auch den Stand der Technik voran. Nachfolgend wird dies deutlich bei einer abschließenden Zusammenfassung des wissenschaftlichen Beitrags.

Das Thema der Arbeit ist der Einsatz von Predictive Maintenance in der diskreten Fertigung. Dazu liefert die Arbeit zunächst eine Motivation zur Nutzung einer voraussagenden Instandhaltungsstrategie in der Produktion. Diese führt zu einer Steigerung der Wettbewerbsfähigkeit durch Erhöhung der technischen Verfügbarkeit. Eine tiefere Analyse des Standes von Wissenschaft und Technik, in Form einer umfangreichen Marktanalyse unter Berücksichtigung der umsatzstärksten Maschinenbauunternehmen, zeigte eine Lücke an vollumfänglichen, verfügbaren Produkten für den Einsatz von Predictive Maintenance auf.

Dies ist nicht verwunderlich, denn der Einsatz der voraussagenden Instandhaltung bringt einige Herausforderungen mit sich. Für die grundlegende Festlegung dieser Herausforderungen wurde zunächst eine Studie zu den Herausforderungen bei der Implementierung von Predictive Maintenance aus dem Jahr 2017 herangezogen und

mit dem aktuellen Stand der Technik abgeglichen. Durch die Ausführung von Implementierungsmöglichkeiten wurde die Hürde für einen erfolgreichen Einsatz der Technologie – Predictive Maintenance – gesenkt. Allerdings konnte für die Herausforderung – Anwendung statistischer Methoden – keine pauschale Antwort gegeben werden. In der Arbeit wurde daher in Kapitel 3 eine umfassende Betrachtung der Anwendung von statistischen Methoden im Rahmen von Predictive Maintenance vorgenommen. Der sich daraus ergebende Systementwurf für die voraussagende Instandhaltung lässt sich auf verschiedenste Branchen übertragen und umfasst eine komplette Datenanalyse von Datenaufbereitung bis hin zur Validierung und Anwendung eines Modells. Auch eine Möglichkeit zur Erfassung von Daten in der Fertigung wurde aufgezeigt. Zur Erstellung eines Vorhersagemodells wurden die bislang gebräuchlichen einfachen statistischen Methoden durch komplexere analytische Methoden des Machine Learning erweitert. Neben dem herkömmlichen Vorgehen aus Daten eines Systems zu lernen, hat die Arbeit einen neuen Ansatz des übertragenden Lernens (Transfer-Learning) für das Thema Predictive Maintenance eingeführt und erstmalig den Mehrwert dieses Ansatzes untersucht und bewertet.

Eine weit größere Herausforderung als die Anwendung statistischer Methoden stellt die Bewertung einer ausreichenden Datenbasis dar. Diese Arbeit hat erstmals Lösungen zur Definition der Datenbasis vorgestellt. Neben einer methodischen Vorgehensweise zur Bewertung der Datenbasis wurde ein Software-Framework geschaffen, welches die Auswertung verschiedenster Lebensdauerdaten für die Allgemeinheit ermöglicht. Dieses lässt nicht nur die Festlegung des Zeitpunkts zu, bei dem die bestmögliche Datenbasis erreicht wird, sondern ermöglicht auch die Bestimmung der maximal möglichen Güte von Vorhersagemodellen bzw. der oberen Schranke für das jeweilige Gütemaß zu jedem Zeitpunkt. Durch die Darstellung des zeitlichen Verlaufs ist auch eine weitere objektive Abschätzung des Mehrwerts in Form einer gesteigerten technischen Verfügbarkeit möglich. Dieser Schritt ist erstmalig in der Wissenschaft durchgeführt worden und kann der Technik zum Einsatz in der Produktion verhelfen. Durch die erarbeitete Vorgehensweise ist es nämlich möglich, die Wirtschaftlichkeit des Einsatzes von Predictive Maintenance je System oder für eine Vielzahl an Systemen zu berechnen. Erwähnenswert ist hierbei noch, dass wenn eine Abschätzung für eine Klasse an Systemen bereits wie hier im Falle der Produktionssysteme mit der Methode vorgenommen wurde, sich diese Abschätzung (wie ebenfalls beispielhaft demonstriert) einfach auf andere Systeme der betreffenden Klasse übertragen lässt. Eine erste Wirtschaftlichkeitsbewertung wurde in Kapitel 7.3 durchgeführt und kann als Schablone für zukünftige Wirtschaftlichkeitsbetrachtungen dienen.

Diese Arbeit wäre nicht möglich gewesen, wenn die aktuellen Methoden nicht erweitert oder übertragen worden wären. Insbesondere die Konfusionsmatrix hatte bisher nur im Machine Learning und somit in der Bewertung von Algorithmen Anwendung gefunden. Durch diese Arbeit können nun die Leistungswerte, also die sich aus der Konfusionsmatrix ergebenden Gütemaße, in Zusammenhang mit der betreffenden Datenbasis gebracht werden. Dazu wurden in Kapitel 4 Annahmen getroffen

und die Gütemaße in die Produktion übertragen sowie entsprechend erweitert. Damit die Methode übertragbar und vielfältig anwendbar ist, wurde durch zwei Schritte eine Generalisierbarkeit der Ergebnisse geschaffen. Als Erstes wurde eine Stichprobe gewählt, welche die Grundgesamtheit bestmöglich abbildet. Als zweites wurde für die Ergebnisse eine Näherungsfunktion bestimmt. Für die Bestimmung der Näherungsfunktion wurde, ein Polynom n -ten Grades bestimmt, wobei die allgemein bekannte Methode der kleinsten Quadrate passend adaptiert wurde. Diese angepasste Vorgehensweise lässt sich sicherlich nicht nur im Rahmen dieser Arbeit, sondern auch für die Anpassung von Verteilungen in anderen Bereichen anwenden. Durch die Berechnung der Standardabweichung konnte die Streuung der Gütewerte über den zeitlichen Verlauf aufgezeigt werden. Diese Vorgehensweise lässt sich künftig auch auf andere Verteilungen einer anders gearteten Stichprobe anwenden. Mittels der in Kapitel 5 bestimmten Gütewerte für das System BAZ 1 wurde sogar bereits eine Validierung der Gütewerte für die Stichprobe 1 vorgenommen und somit der Wissenschaft eine langfristig verwendbare Kennzahl für die voraussagende Instandhaltung von Produktionssystemen gegeben.

In der bisherigen Literatur wurde bislang nicht in dieser Art der Einsatz und schon gar nicht der unterschiedliche Einsatz von Predictive Maintenance thematisiert. Diese Arbeit hat erstmalig sogar drei verschiedene Umsetzungsszenarien aufgestellt sowie miteinander verglichen und damit deutlich gemacht, dass voraussagende Instandhaltung und deren Mehrwert nicht pauschalisiert werden kann und darf.

8.3 Ausblick

Abschließend für diese Arbeit werden im Folgenden mögliche weitere Schritte zur Bearbeitung des Themas „Einsatz von Predictive Maintenance“ benannt.

Die bisherige Arbeit hat bereits allgemeingültige Ergebnisse geliefert, allerdings würde eine Erhebung **weiterer Lebensdauerdaten** und erneute Anwendung der Methode und somit Bestimmung der Gütemaße die Generalisierbarkeit der Ergebnisse weiter untermauern. Auch wäre es denkbar, wenn auch zeitaufwendig und fehleranfällig, die Auswertung nochmals mit Lebensdauerdaten von real ausgefallenen Komponenten durchzuführen und abzugleichen, wodurch die Methode und die damit bislang erhaltenen Ergebnisse weitergehend unter Beweis gestellt werden könnten.

Eine weitere Möglichkeit den Realitätsgrad der Methode zu steigern, wäre die **Beschreibung der Ausfallwahrscheinlichkeiten durch komplexere Verteilungen**. Allerdings ist dies sehr beschwerlich, da für jeden Ausfall die Wahrscheinlichkeitsverteilung individuell gewählt werden müsste. Die bisherige Annahme in Abschnitt 4.2 legt daher eine allgemeingültige Beschreibung der Verteilung für die Ausfallwahrscheinlichkeit durch eine Delta-Funktion nahe.

Möglich wäre noch eine Erweiterung der entwickelten Methode um zwei Punkte:

- Die Häufigkeit, mit der ein Ausfall einer Komponente auftritt, bevor dieser erlernt und vermeidbar ist, liegt aktuell bei eins ($s_K = 1$). Klassifikationsmodelle bevorzugen jedoch eine größere Anzahl an Daten eines Ausfalls, um die Klassifikation korrekt durchführen zu können. Daher ist es denkbar, die Beobachtungshäufigkeit für einen Komponentenausfall und damit die **Häufigkeit zum Erlernen des Ausfalls**, zu verändern und damit neue Abschätzungen zu tätigen. Die Methode gibt dies bereits her. Bislang wurde aber darauf verzichtet, weil damit nur die obere Schranke für die Güte eines Vorhersagemodells sinken würde.
- Außerdem wurde in dieser Arbeit angenommen, dass ein Ausfall einer Komponente immer mit den gleichen Symptomen auftritt. Dies könnte jedoch nicht immer der Fall sein. Daher erscheint es hilfreich, die Methode zu ergänzen und **unterschiedliche Ausfalltypen je Komponente** zu berücksichtigen. Diese Ergänzung ist bereits durch die Anpassung der Stichprobe bzw. deren Komponenten-Lebensdauern⁵¹ auf triviale Weise möglich.

Neben der Erweiterung der Methode könnte ein weiterer, definitiv möglicher und durchaus sinnvoller Schritt die **Übertragung der Methodik auf andere Branchen** (Fahrzeug, Flugzeug, Pharmaindustrie, ...) sein. Denn auch hier besteht der Bedarf an voraussagender Instandhaltung und es steht ebenfalls die Frage im Raum, inwiefern deren Einsatz wirtschaftlich ist. Damit die Abschätzung des Aufwandes und des Mehrwertes einer voraussagenden Instandhaltung allerdings anwenderfreundlich erfolgen kann, sollte das im Rahmen dieser Arbeit entwickelte Framework hin zu einer **professionellen Applikation** weiterentwickelt werden. Das bisherige Software-Framework ermöglicht aber bereits jetzt die Auswertung der Gütemaße und die Bestimmung des Mehrwertes für andere Klassen an Systemen bzw. Industriesparten, allerdings müssen Grundinformationen wie die Lebensdauerdaten und die aktuelle technische Verfügbarkeit manuell in den Code eingetragen werden. Die professionelle Applikation hingegen würde die Grundinformationen in einer benutzerfreundlichen Oberfläche entgegennehmen und könnte „auf einen Mausklick“ alle Ergebnisse berechnen und ausgeben.

Neben der Bewertung von Predictive Maintenance sollte in der zukünftigen wissenschaftlichen Betrachtung aber auch die ausgedehntere Untersuchung **eventuell weiterer Herausforderungen beim Einsatz der Technologie** nicht zu kurz kommen. Denn zusätzlich zu den in Abschnitt 0 benannten Herausforderungen sind aus meiner Sicht noch weitere Herausforderungen bislang noch nicht hinreichend wissenschaftlich betrachtet worden:

- **Ganzheitliche Datenerfassung:** Heutzutage ist es noch nicht bei jedem Gerät (z.B. SPS) möglich Daten zu erfassen. Zusätzliche Geräte zur Erfassung von Daten (IoT-Gateways) werden in den nächsten Jahren sicherlich Abhilfe

⁵¹ Sollte eine Komponente A zwei Ausfalltypen haben, so sollte diese in Komponente A1 und Komponente A2 aufgeteilt und entsprechend die Lebensdauer des Ausfallstyps angenommen werden.

schaffen, jedoch führen weitere Geräte zu neuen Herausforderungen (z.B. Betrieb, Stromverbrauch, Platzverbrauch, Wartung, ...). Daher sollten sich künftige Forschungen mit multifunktional nutzbaren Geräten befassen.

- **Standardisierte Datenformate:** IoT-Daten werden bisher meist unstrukturiert abgespeichert und sind deswegen für Predictive Maintenance nur schwer verwendbar. Daher ist es empfehlenswert ein standardisiertes Datenformat für diese festzulegen.
- **Dauer der Datenspeicherung:** Häufig wird angenommen, dass Daten für immer in der Cloud gespeichert werden könnten. Allerdings gibt es auch hier physikalische Grenzen (vgl. Data Growth Explosion). Daher ist zielführend zu analysieren, wie lange Daten für den Einsatz von voraussagender Instandhaltung gespeichert werden müssen.
- **Echtzeitdaten in der Cloud:** Heute gibt es Möglichkeiten die Daten in Echtzeit in die Cloud zu bringen und somit vorgesehene Analysen dort durchzuführen. Allerdings führt eine schnellere Verfügbarkeit der Daten in der Cloud oft zu höheren Kosten. Es sollte daher untersucht werden, mit welcher Geschwindigkeit Daten für Analysen in der Cloud verfügbar sein müssen, damit ein Vorhersagemodell auch nutzbringend eingesetzt werden kann und zum Beispiel noch rechtzeitig eine Wartung vor dem eigentlichen Ausfall eingeplant werden kann.
- **Auditfähigkeit:** Früher wurden Protokolle (analoge Daten) auditiert und heutzutage sollte mit digitalen Daten ähnlich verfahren werden. Dies unterscheidet sich erst einmal vom Prozess nicht, jedoch müssen diese vor Fremdzugriff gesichert und eine Veränderung an den Daten nachweisbar gemacht werden. Es ist für die Zukunft sicherlich relevant, einen Prozess für die Auditfähigkeit von Predictive Maintenance einzuführen.
- **Akzeptanz:** Schon bei der Einführung des ersten Webstuhls Ende des 18. Jahrhunderts gab es Kritiker dieser damals neuen Technologie und auch die weitere Revolution durch Fließbänder und Automation wurde immer wieder kritisch von der Gesellschaft hinterfragt. Schlussendlich hat uns jedoch jede dieser Technologien zu unserem heutigen Industrialisierungsgrad gebracht. Es ist daher notwendig, die Akzeptanz bei den Menschen für die voraussagende Instandhaltung und verwandte Technologien zu stärken.
- **Geschäftsmodell:** Eine weitere organisatorische Herausforderung besteht in den neuen Einkaufsprozessen. Möglichkeiten wie Pay-per-Use oder jährliche Lizenzmodelle stellen Unternehmen vor neue Prozesse und hinterfragen immer wieder den Mehrwert einer Technologie. Eine Beschreibung dieser neuen Geschäftsmodelle sollte für Predictive Maintenance veröffentlicht werden.

Zusammenfassend kann im Ausblick festgehalten werden, dass die Herausforderungen und Mehrwerte beim Einsatz von Predictive Maintenance unter Verwendung von Machine Learning in der diskreten Fertigung größtenteils lösbar oder zumindest benennbar sind.

A Bestimmung der Vorhersagegüte

A.1 Mittlere Lebensdauern je Komponente zugeordnet zum jeweiligen System

Nachfolgend werden für das jeweilige betrachtete System die mittleren Lebensdauern je Komponente L_K angegeben.

Automation 1 = {10.000, 42.000, 60.000, 60.000, 60.000, 66.000, 80.000, 80.000, 90.000, 120.000}

Automation 2 = {5.000, 20.000, 20.000, 20.500, 20.500, 20.500, 30.000, 30.000, 32.500, 32.500, 32.500, 105.000, 120.000}

Bandofen 1 = {1.000, 3.000, 6.000, 12.000, 12.000, 20.000, 24.000, 30.000, 36.000, 36.000, 48.000}

BAZ 1 = {10.000, 10.000, 10.000, 15.000, 18.000, 20.000, 20.000, 20.000, 20.000, 20.000, 25.000, 30.000, 30.000, 30.000, 100.000}

BAZ 2 = {6.000, 6.000, 6.000, 6.000, 12.000, 12.000, 15.000, 18.000, 30.000, 30.000, 30.000, 30.000, 36.000, 50.000, 50.000, 50.000, 60.000}

BAZ 3 = {6.000, 15.000, 15.000, 15.000, 20.000, 20.000, 20.000, 20.000, 20.000, 30.000, 30.000, 30.000, 30.000, 30.000, 30.000, 30.000, 30.000, 35.000, 60.000, 60.000, 60.000}

BAZ 4 = {5.000, 18.000, 18.000, 18.000, 24.000, 30.000, 30.000, 30.000, 30.000, 30.000, 36.000, 36.000, 60.000, 60.000, 60.000, 60.000, 60.000, 60.000, 60.000, 60.000, 60.000}

BAZ 5 = {40.000, 40.000, 40.000, 40.000, 40.000, 40.000, 45.000, 45.000, 45.000, 50.000, 50.000, 50.000, 50.000, 50.000, 50.000, 50.000, 50.000, 50.000, 50.000, 50.000, 50.000, 70.000}

BAZ 6 = {3.103, 6.923, 8.182, 13.500, 15.000, 18.000, 22.500, 22.500, 30.000, 30.000, 36.000, 45.000, 45.000, 45.000, 45.000, 90.000, 90.000, 90.000}

BAZ 7 = {12.000, 18.000, 18.000, 18.000, 20.280, 21.000, 21.000, 21.600, 23.000, 23.000, 30.000, 36.000, 36.000}

Biegeanlage 1 = {3.176, 4.909, 6.000, 6.000, 13.500, 13.500, 18.000, 28.000, 28.000, 36.000, 36.000, 54.000, 54.000, 54.000}

Dichtheitsprüfanlage 1 = {4.700, 18.800, 18.800, 23.500, 23.500, 23.500, 23.500, 23.500}

Drehmaschine 1 = {4.209, 9.400, 9.400, 12.000, 14.286, 15.000, 16.667, 16.667, 16.667, 17.241, 18.000, 18.868, 19.048, 19.929, 19.929, 25.000, 27.000, 30.000, 30.000, 30.000, 40.000, 40.000, 41.667}

Fügesystem 1 = {3.200, 3.200, 3.200, 20.000, 27.000, 27.100, 27.100, 30.000, 30.000, 30.000, 30.000, 30.000}

Fügesystem 2 = {1.000, 3.250, 3.430, 8.000, 8.310, 9.000, 9.000, 9.260, 10.000, 15.000, 15.000, 15.000, 20.600, 21.600, 21.600, 22.500, 22.500, 23.000, 25.000, 25.000, 25.000, 27.000, 30.000, 30.000, 30.000, 30.000, 30.000, 30.400, 32.500, 32.800, 32.800, 32.800, 37.500, 38.000, 40.000, 40.000, 50.000, 50.000, 50.000, 50.000, 70.000, 80.000, 80.000, 80.000, 80.000, 90.000, 95.000, 110.130}

Handlingsanlage 1 = {20.000, 20.500, 23.500, 30.000, 32.500, 32.500, 105.000, 120.000}

Handlingsanlage 2 = {6.500, 6.500, 23.500, 32.500, 32.500, 32.500, 50.000, 50.000, 50.000, 50.000, 50.000}

Härtemaschine 1 = {1.200, 30.000, 30.000, 30.000, 80.000}

Härtemaschine 2 = {500, 1.000, 5.000, 7.500, 23.500, 25.000, 25.000, 25.000, 50.000}

Heftsystem 1 = {5.000, 10.000, 10.000, 10.000, 10.000, 10.000, 15.000, 15.000, 15.000, 15.000, 15.000, 15.000}

Messsystem 1 = {10.000, 50.000, 50.000, 50.000, 100.000, 100.000}

Messsystem 2 = {6.250, 7.143, 10.000, 10.000, 10.000, 10.000, 12.000, 12.500, 12.500, 18.750, 20.000, 25.000, 25.000, 25.000, 50.000, 50.000, 50.000}

Messsystem 3 = {18.000, 30.000, 36.000}

Montagesystem 1 = {5.000, 5.000, 10.000, 10.000, 10.000, 10.000, 10.000, 20.000, 20.000, 20.000, 20.000, 20.000}

Montagesystem 2 = {20.000, 20.000, 30.000, 40.000, 40.000, 40.000, 40.000, 40.000, 40.000, 40.000, 40.000, 40.000, 40.000, 40.000, 40.000, 40.000, 40.000, 40.000, 40.000, 40.000, 40.000, 50.000, 50.000, 80.000, 80.000, 80.000, 80.000, 80.000, 80.000, 80.000, 80.000}

Montagesystem 3 = {20.000, 20.000, 30.000, 30.000, 30.000, 30.000, 30.000, 30.000, 30.000, 30.000}

Montagesystem 4 = {24.000, 27.000, 27.000, 54.000, 54.000, 54.000, 54.000, 54.000, 54.000, 54.000, 54.000, 54.000, 108.000, 108.000}

Reinigungssystem 1 = {5.000, 12.500, 12.500, 12.500, 25.000, 27.000, 29.000, 33.000, 35.000, 42.000, 50.000, 58.000, 58.000, 58.000, 58.000, 58.000, 58.000, 58.000, 58.000, 58.000, 58.000}

Richtmaschine 1 = {9.000, 9.000, 10.000, 35.000, 35.000, 40.000, 40.000, 40.000, 45.000, 45.000, 45.000, 45.000, 50.000, 50.000, 50.000, 50.000, 100.000, 100.000}

Richtmaschine 2 = {8.500, 10.000, 10.000, 10.000, 10.000, 10.000, 10.000, 25.000, 25.000, 25.000, 25.000, 25.000, 25.000, 25.000, 25.000, 30.000, 50.000}

Schleifsystem 1 = {3.000, 12.000, 12.000, 12.000, 20.000, 20.000, 20.000, 20.000, 20.000, 20.000, 25.000, 25.000, 30.000, 30.000, 30.000, 30.000, 60.000, 60.000}

Schleifsystem 2 = {13.000, 15.000, 16.000, 16.000, 16.000, 16.000, 16.000, 16.000, 20.000, 24.000, 24.000, 30.000, 30.000, 30.000, 30.000, 30.000, 45.000, 60.000, 60.000, 60.000, 60.000, 60.000, 60.000, 60.000, 60.000, 60.000, 60.000, 60.000}

Wälzfräsmaschine 1 = {20.000, 20.000, 20.000, 20.000, 20.000, 25.000, 25.000, 25.000, 25.000, 30.000}

Werkzeugmaschine 1 = {12.000, 12.000, 12.000, 30.000, 30.000, 30.000, 34.000, 34.000, 34.000, 34.000, 34.000, 34.000, 34.000, 36.000, 36.000}

Werkzeugmaschine 2 = {7.000, 18.800, 23.500, 23.500, 23.500, 23.500, 23.500, 23.500, 23.500, 28.200, 28.200, 37.600, 37.600, 42.300, 47.000, 47.000, 47.000, 70.500, 94.000}

A.2 Generalisierung der Vorhersagegüte für Stichprobe 1 im Szenario 2

Dieser Anhang stellt die Gütewerte für die allgemeingültige Klasse von Produktionssystemen vor, bei der die Datenerfassung erst ab 35.040 Stunden bzw. 4 Jahren beginnt.

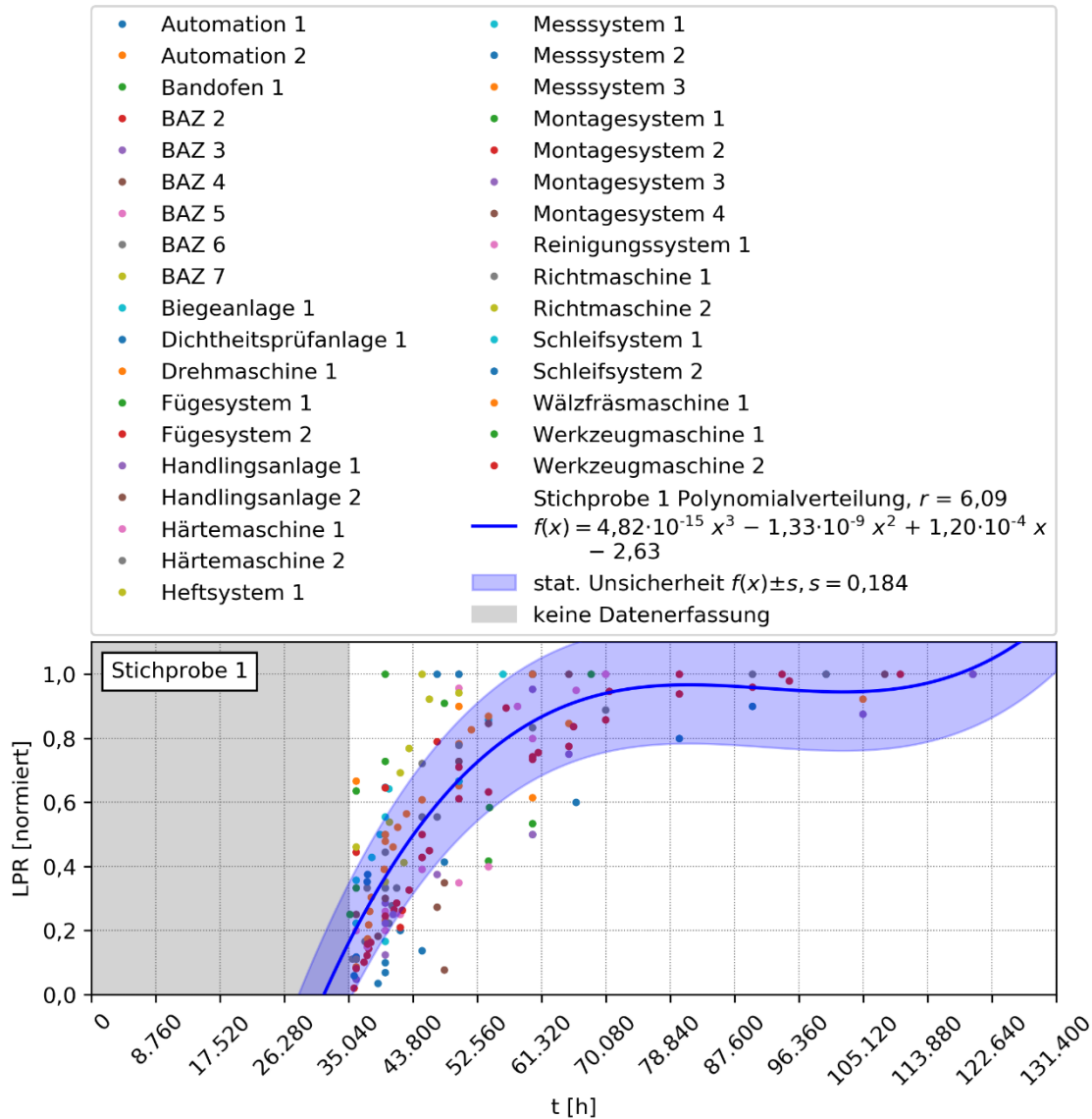


Bild A.1: Normierte Datenpunkte aller einzelnen Systeme und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Lernrate LPR für erlernbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb: Dargestellt sind die Datenpunkte der Systeme (in verschiedenen Farben) sowie die Näherungsfunktionen der Stichprobe 1 (in Blau) und deren statistische Unsicherheit s (Hüllkurve in Hellblau) für die Lernrate LPR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Komponenten des jeweiligen Systems $|K|$ normiert. Für die Näherungsfunktion der Stichprobe 1 wurde ein Polynom 3. Grades gewählt. Der ausgegraute Bereich stellt den Zeitraum $t = [0, 35.040]$ in Stunden h dar, in dem noch keine Datenerfassung erfolgt ist.

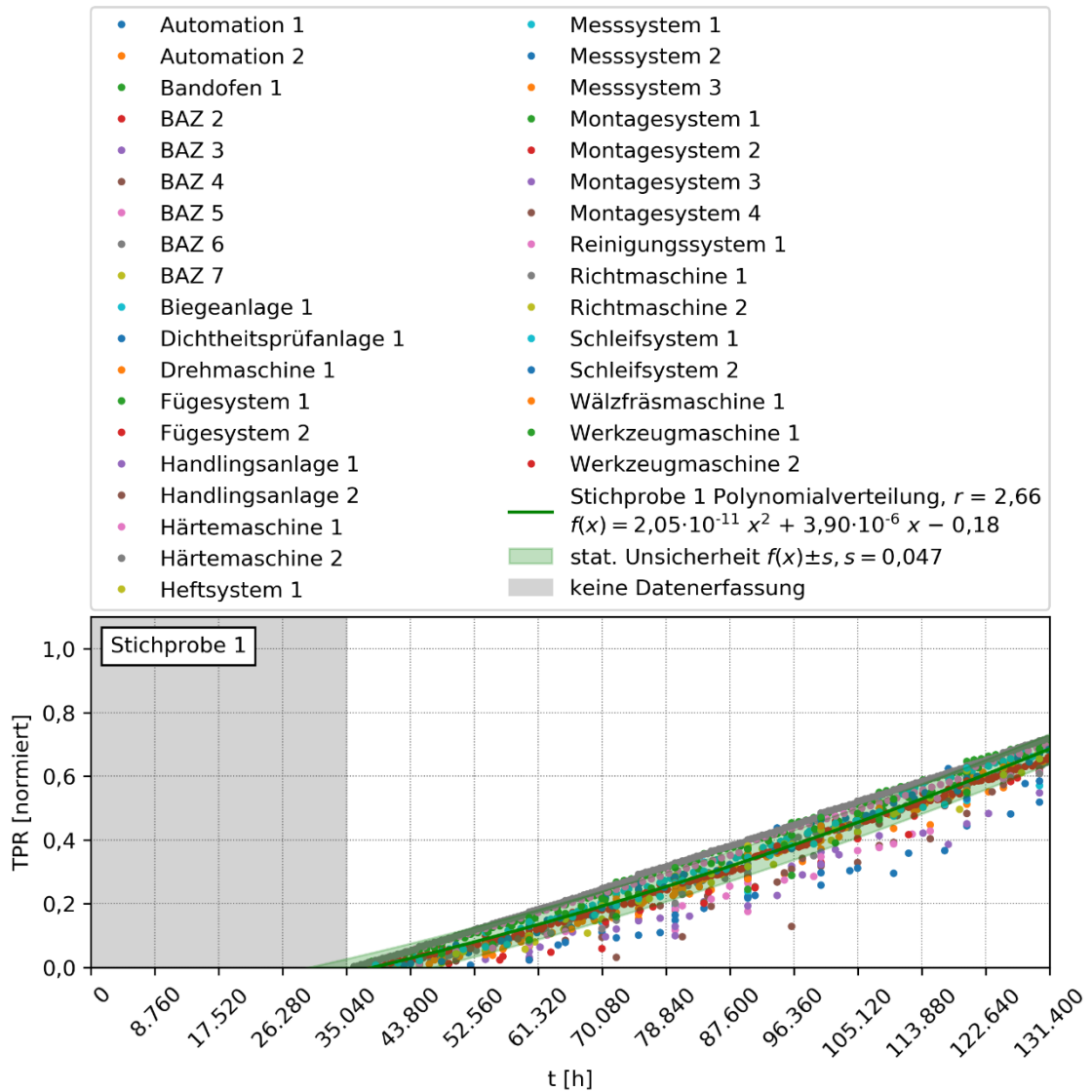


Bild A.2: Normierte Datenpunkte aller einzelnen Systeme und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Trefferrate TPR für vermeidbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb: Dargestellt sind die Datenpunkte der Systeme (in verschiedenen Farben) sowie die Näherungsfunktionen der Stichprobe 1 (in Grün) und deren statistische Unsicherheit s (Hüllkurve in Hellgrün) für die Trefferrate TPR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Ausfälle des jeweiligen Systems $P(0 \leq t < B)$ normiert. Für die Näherungsfunktion der Stichprobe 1 wurde ein Polynom 2. Grades gewählt. Der ausgegraute Bereich stellt den Zeitraum $t = [0, 35.040[$ in Stunden h dar, in dem noch keine Datenerfassung erfolgt ist.

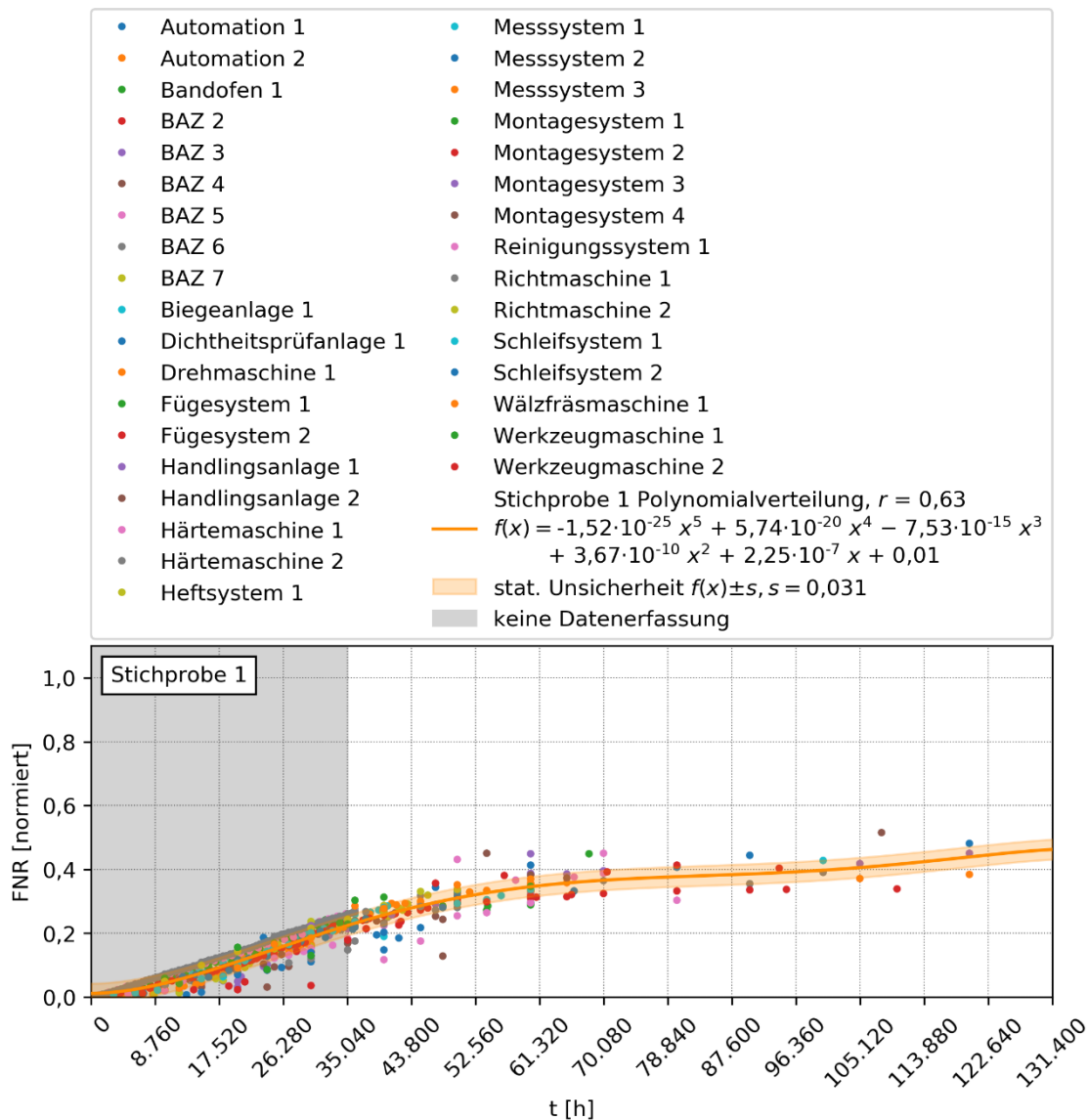


Bild A.3: Normierte Datenpunkte aller einzelnen Systeme und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Nicht-Alarmrate FNR für verbleibende Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb: Dargestellt sind die Datenpunkte der Systeme (in verschiedenen Farben) sowie die Näherungsfunktionen der Stichprobe 1 (in Orange) und deren statistische Unsicherheit s (Hüllkurve in Hellorange) für die Nicht-Alarmrate FNR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Ausfälle des jeweiligen Systems $P(0 \leq t < B)$ normiert. Für die Näherungsfunktion der Stichprobe 1 wurde ein Polynom 5. Grades gewählt. Der ausgegraute Bereich stellt den Zeitraum $t = [0, 35.040[$ in Stunden h dar, in dem noch keine Datenerfassung erfolgt ist.

A.3 Validierung der Stichprobe 1 durch BAZ 1 im Szenario 2

In diesem Anhang wird die Generalisierbarkeit der Güterwerte für die Klasse von Produktionssystemen anhand eines Vergleiches mit den Güterwerten von dem einzelnen und unabhängigen System BAZ 1 überprüft. Wie erwartet, liegt die Majorität der Datenpunkte von System BAZ 1 innerhalb des für die Klasse der Produktionssysteme ermittelten Unsicherheitsbandes. Daraus lässt sich ableiten, dass die für die Klasse von Produktionssystemen bestimmten Güterwerte von allgemeiner Natur und damit generell anwendbar sind. Die Datenerfassung in diesem Szenario 2 beginnt erst ab 35.040 Stunden bzw. 4 Jahren.

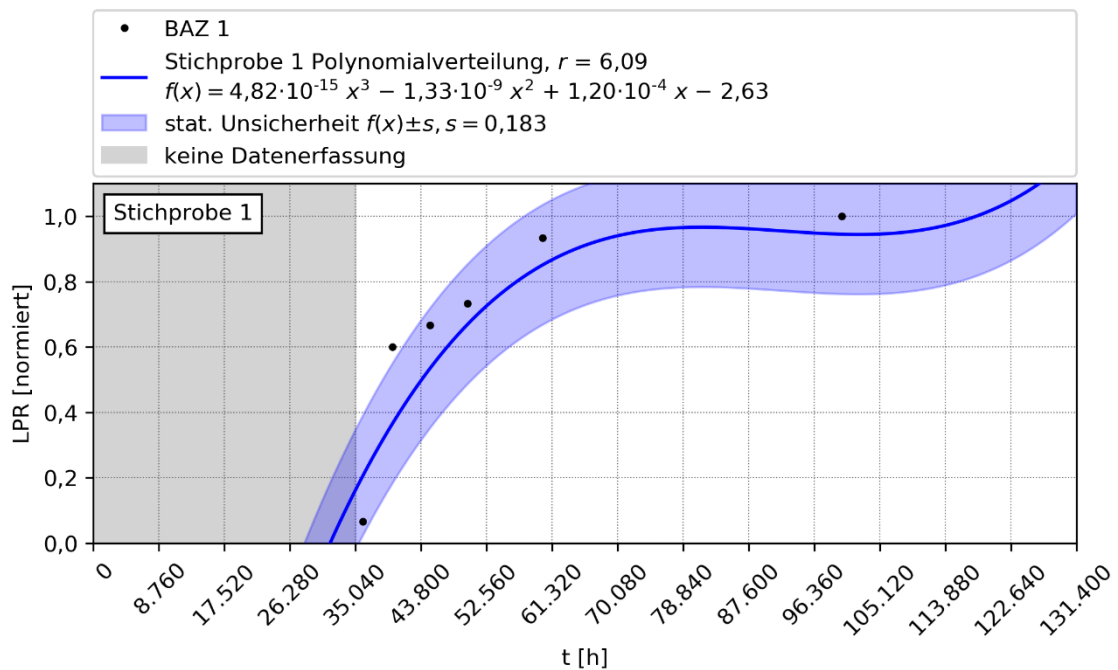


Bild A.4: Normierte Datenpunkte des Systems BAZ 1 und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Lernrate LPR für erlernbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb: Dargestellt sind die Datenpunkte des Systems BAZ 1 (in Schwarz) sowie die Näherungsfunktionen der Stichprobe 1 (in Blau) und deren statistische Unsicherheit s (Hüllkurve in Hellblau) für die Lernrate LPR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Komponenten für BAZ 1 $|K| = 15$ normiert. Für die Näherungsfunktion der Stichprobe 1 wurde ein Polynom 3. Grades gewählt. Der ausgegraute Bereich stellt den Zeitraum $t = [0, 35.040[$ in Stunden h dar, in dem noch keine Datenerfassung erfolgt ist.

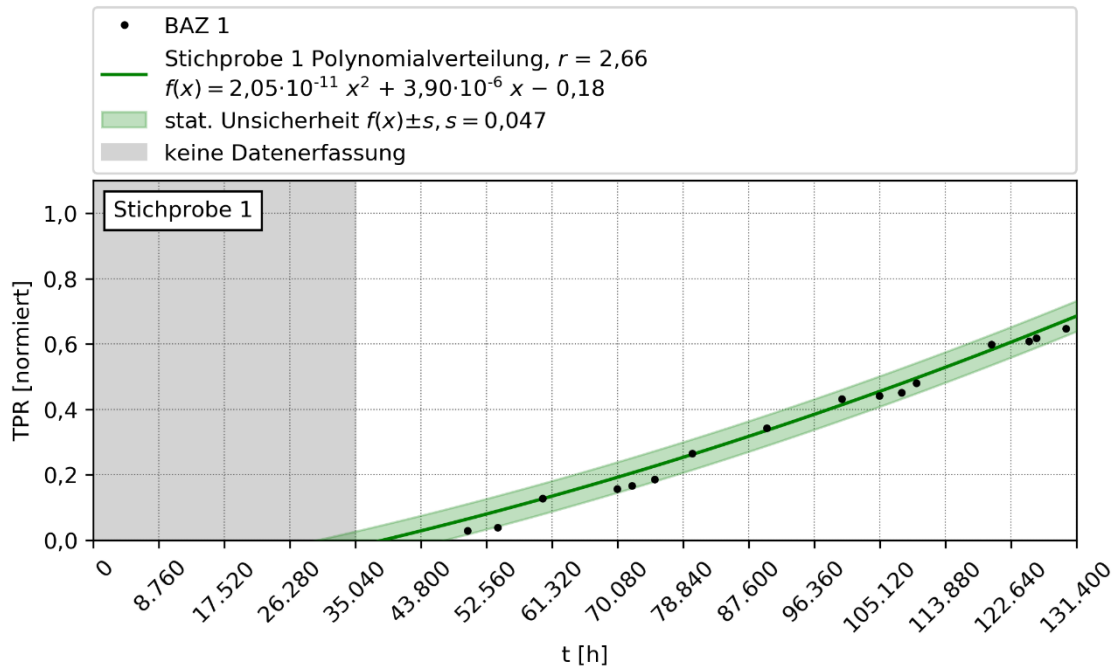


Bild A.5: Normierte Datenpunkte des Systems BAZ 1 und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Trefferrate TPR für vermeidbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb: Dargestellt sind die Datenpunkte des Systems BAZ 1 (in Schwarz) sowie die Näherungsfunktionen der Stichprobe 1 (in Grün) und deren statistische Unsicherheit s (Hüllkurve in Hellgrün) für die Trefferrate TPR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Ausfälle für BAZ 1 $P_{BAZ1}(0 \leq t < B) = 102$ normiert. Für die Näherungsfunktion der Stichprobe 1 wurde ein Polynom 2. Grades gewählt. Der ausgegraute Bereich stellt den Zeitraum $t = [0, 35.040[$ in Stunden h dar, in dem noch keine Datenerfassung erfolgt ist.

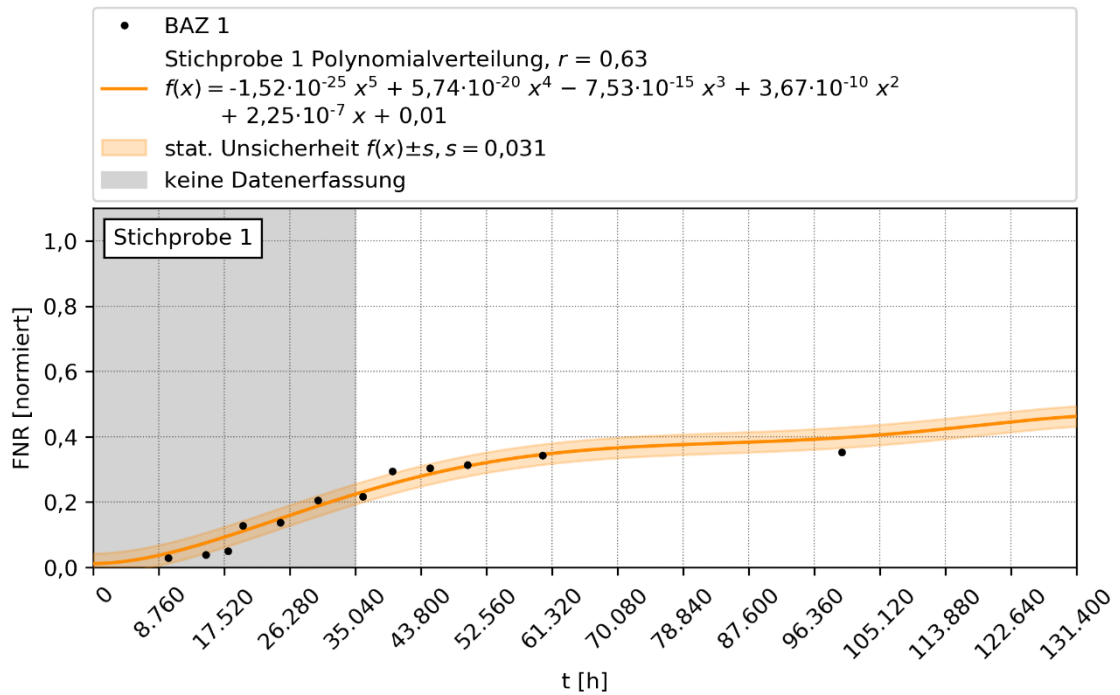


Bild A.6: Normierte Datenpunkte des Systems BAZ 1 und genäherte Verteilung der Stichprobe 1 mit statistischer Unsicherheit der Nicht-Alarmrate FNR für verbleibende Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb: Dargestellt sind die Datenpunkte des Systems BAZ 1 (in Schwarz) sowie die Näherungsfunktionen der Stichprobe 1 (in Orange) und deren statistische Unsicherheit s (Hüllkurve in Hellorange) für die Nicht-Alarmrate FNR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Ausfälle für BAZ 1 $P_{BAZ1}(0 \leq t < B) = 102$ normiert. Für die Näherungsfunktion der Stichprobe 1 wurde ein Polynom 5. Grades gewählt. Der ausgegraute Bereich stellt den Zeitraum $t = [0, 35.040[$ in Stunden h dar, in dem noch keine Datenerfassung erfolgt ist.

A.4 Validierung der Stichprobe 2

In diesem Anhang wurde eine weitere Validierung der Gütwerte für die Klasse von Produktionssystemen durchgeführt. Dazu wurde eine Stichprobenumfang von zufällig 24 Systemen – Stichprobe 2 – für die Generalisierung und 10 zufällig gewählten Systemen zur Validierung gewählt.

Stichprobe 2 besteht aus den Systemen Automation 1, Automation 2, Bandofen 1, BAZ 1 bis 7, Biegeanlage 1, Drehmaschine 1, Fügeseystem 2, Handlingsanlage 2, Härtemaschine 2, Messsystem 2, Montagesystem 1 bis 4, Reinigungssystem 1, Richtmaschine 1, Schleifsystem 2, Wälzfräsmaschine 1 und Werkzeugmaschine 2.

Die Testsysteme sind Dichtprüfanlage 1, Fügeseystem 1, Handlingsanlage 1, Härtemaschine 1, Heftsystem 1, Messsystem 1 und 3, Richtmaschine 2, Schleifsystem 1 und Werkzeugmaschine 1.

Wie erwartet, liegt die Majorität der Datenpunkte der 10 Testsysteme innerhalb des für die Klasse der Produktionssysteme ermittelten Unsicherheitsbandes. Es lässt sich ableiten, dass auch bei einer klassischen Kreuzvalidierung mit 70 % zum Training und 30 % zum Testen die Gütwerte zur Prognose der bestmöglichen Vorhersagegüte herangezogen werden können.

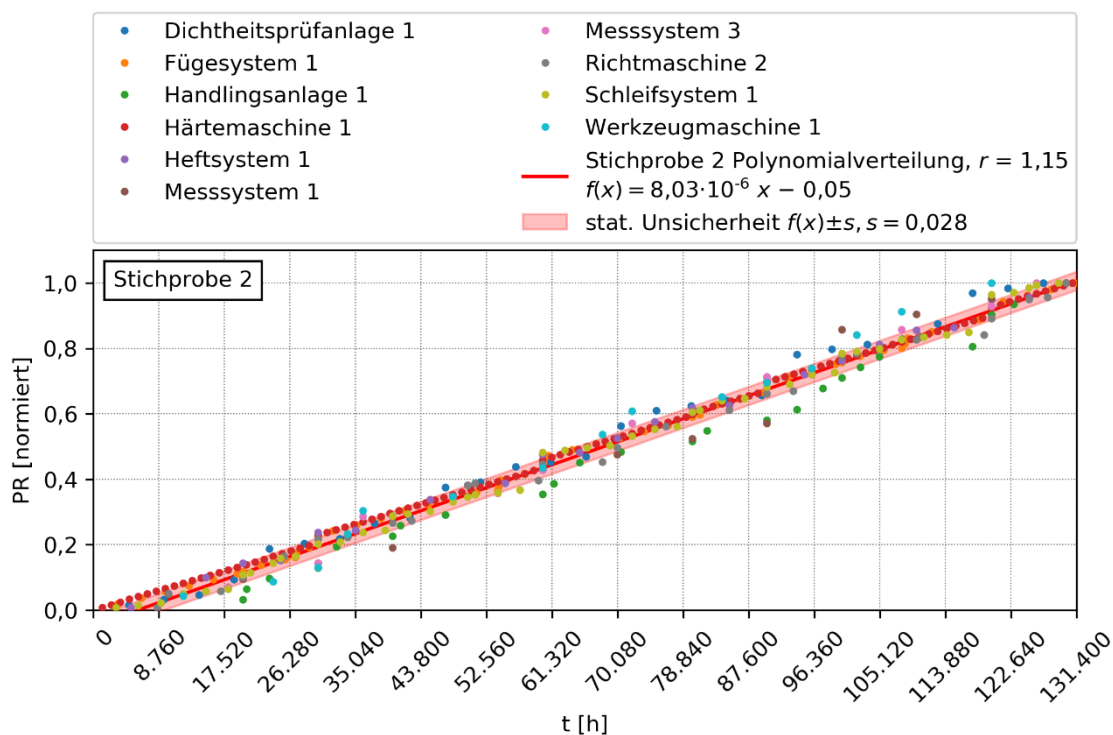


Bild A.7: Normierte Datenpunkte der Testsysteme und genäherte Verteilung der Stichprobe 2 mit statistischer Unsicherheit der Ausfallrate PR für wahre Ausfälle: Dargestellt sind die Datenpunkte der Testsysteme (in verschiedenen Farben) sowie die Näherungsfunktionen der Stichprobe 2 (in Rot) und deren statistische Unsicherheit s (Hüllkurve in Hellrot) für die Ausfallrate PR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Ausfälle des jeweiligen Systems $P(0 \leq t < B)$ normiert. Für die Näherungsfunktion der Stichprobe 2 wurde ein Polynom 1. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400[h$.

Zunächst werden die Gütwerte für das **Szenario 1** dargestellt.

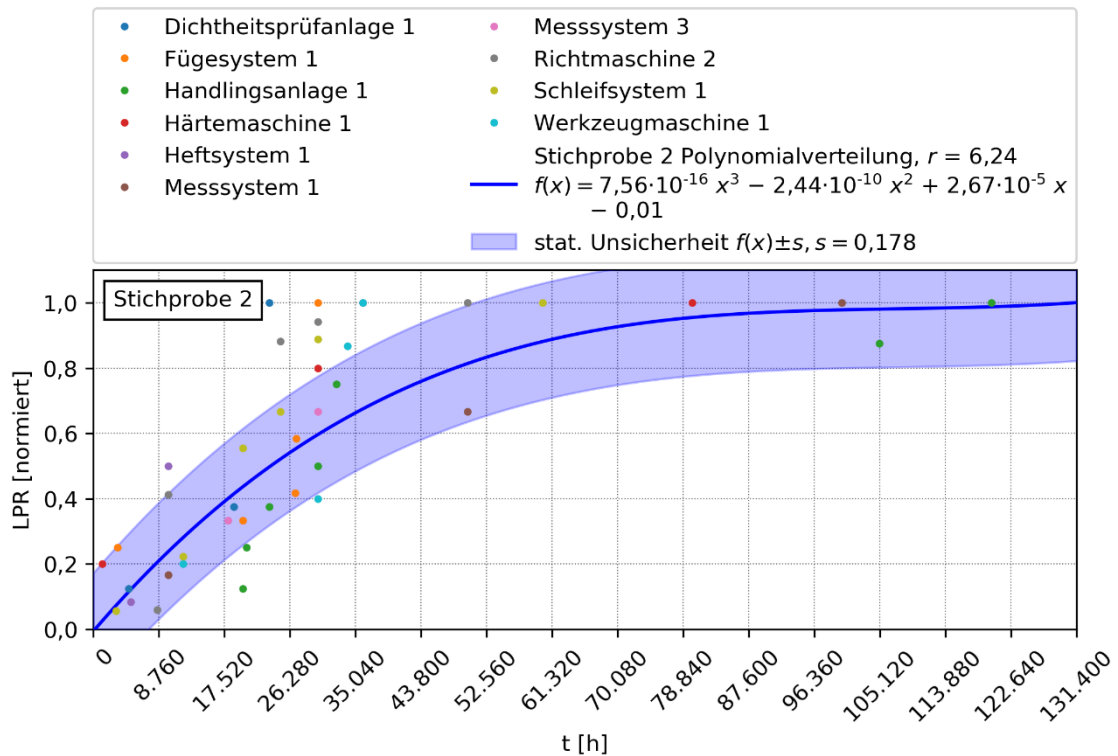


Bild A.8: Normierte Datenpunkte der Testsysteme und genäherte Verteilung der Stichprobe 2 mit statistischer Unsicherheit der Lernrate LPR für erlernbare Ausfälle entsprechend Szenario 1: Dargestellt sind die Datenpunkte der Testsysteme (in verschiedenen Farben) sowie die Näherungsfunktionen der Stichprobe 2 (in Blau) und deren statistische Unsicherheit s (Hüllkurve in Hellblau) für die Lernrate LPR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Komponenten des jeweiligen Systems $|K|$ normiert. Für die Näherungsfunktion der Stichprobe 2 wurde ein Polynom 3. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400[h$.

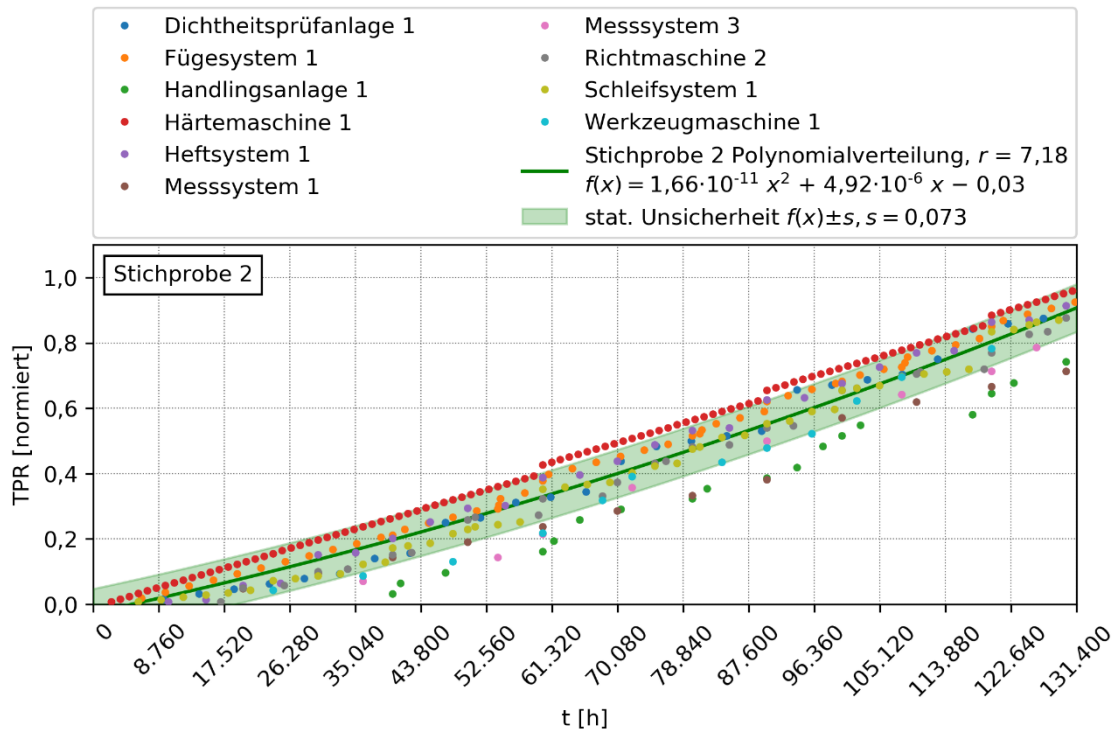


Bild A.9: Normierte Datenpunkte der Testsysteme und genäherte Verteilung der Stichprobe 2 mit statistischer Unsicherheit der Trefferrate TPR für vermeidbare Ausfälle entsprechend Szenario 1: Dargestellt sind die Datenpunkte der Testsysteme (in verschiedenen Farben) sowie die Näherungsfunktionen der Stichprobe 2 (in Grün) und deren statistische Unsicherheit s (Hüllkurve in Hellgrün) für die Trefferrate TPR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Ausfälle des jeweiligen Systems $P(0 \leq t < B)$ normiert. Für die Näherungsfunktion der Stichprobe 2 wurde ein Polynom 2. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400[h$.

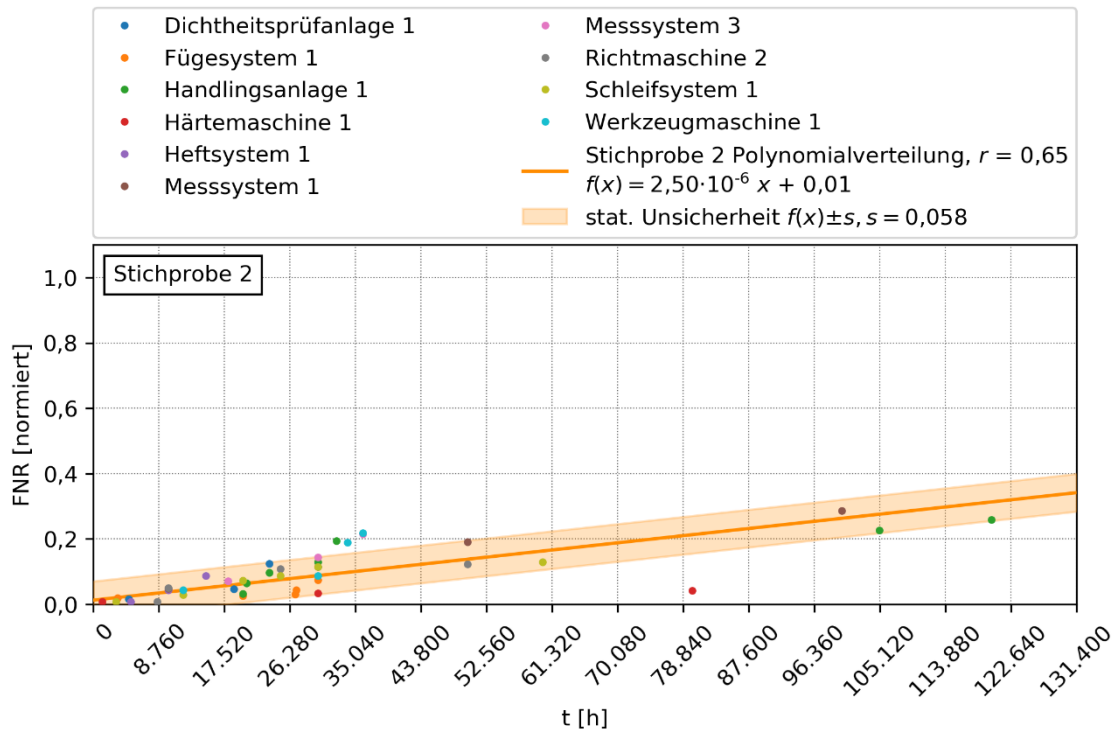


Bild A.10: Normierte Datenpunkte der Testsysteme und genäherte Verteilung der Stichprobe 2 mit statistischer Unsicherheit der Nicht-Alarmrate FNR für verbleibende Ausfälle entsprechend Szenario 1: Dargestellt sind die Datenpunkte der Testsysteme (in verschiedenen Farben) sowie die Näherungsfunktionen der Stichprobe 2 (in Orange) und deren statistische Unsicherheit s (Hüllkurve in Hellorange) für die Nicht-Alarmrate FNR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Ausfälle des jeweiligen Systems $P(0 \leq t < B)$ normiert. Für die Näherungsfunktion der Stichprobe 2 wurde ein Polynom 1. Grades gewählt. Der betrachtete Betriebszeitraum ist $t_B = [0, 131.400[h$.

Im Folgenden werden die Güterwerte für das **Szenario 2** abgebildet.

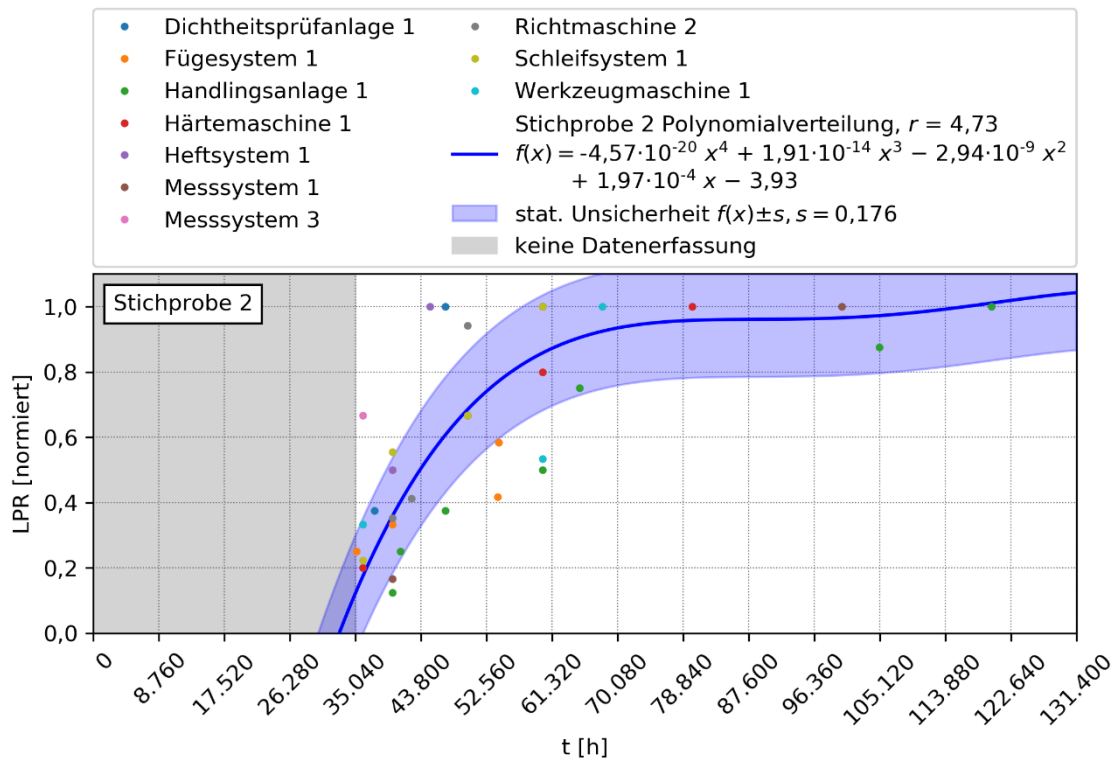


Bild A.11: Normierte Datenpunkte der Testsysteme und genäherte Verteilung der Stichprobe 2 mit statistischer Unsicherheit der Lernrate LPR für erlernbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb: Dargestellt sind die Datenpunkte der Testsysteme (in verschiedenen Farben) sowie die Näherungsfunktionen der Stichprobe 2 (in Blau) und deren statistische Unsicherheit s (Hüllkurve in Hellblau) für die Lernrate LPR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Komponenten des jeweiligen Systems $|K|$ normiert. Für die Näherungsfunktion der Stichprobe 2 wurde ein Polynom 4. Grades gewählt. Der ausgegraute Bereich stellt den Zeitraum $t = [0, 35.040[$ in Stunden h dar, in dem noch keine Datenerfassung erfolgt ist.

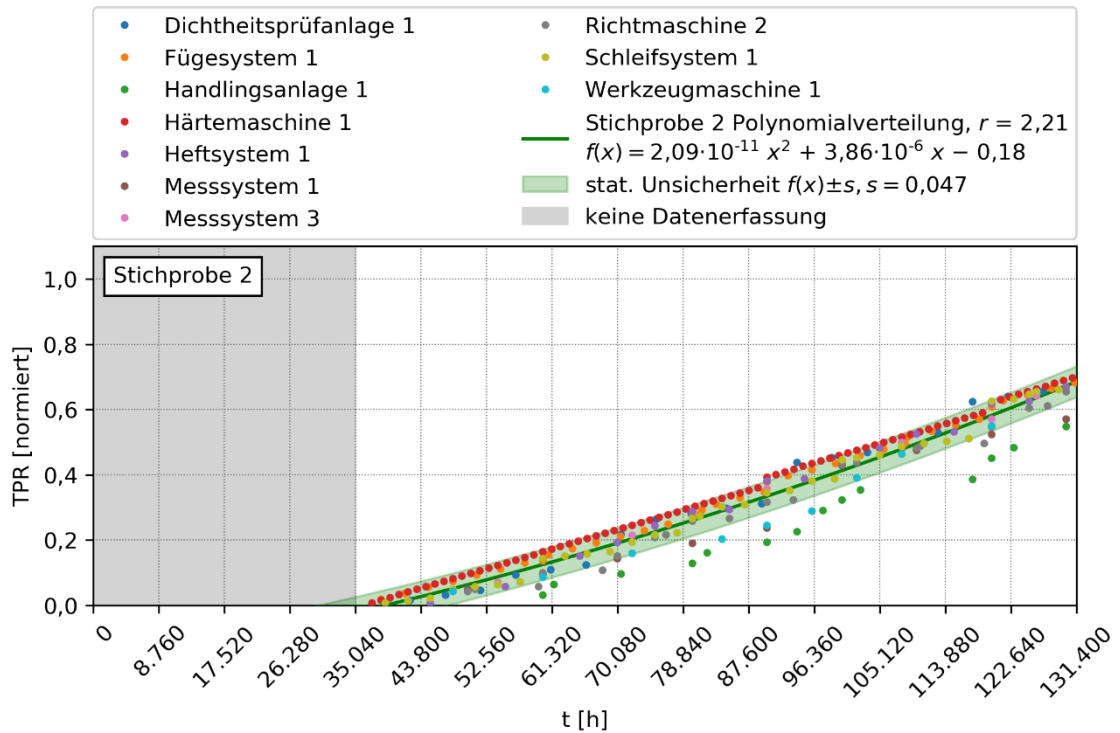


Bild A.12: Normierte Datenpunkte der Testsysteme und genäherte Verteilung der Stichprobe 2 mit statistischer Unsicherheit der Trefferrate TPR für vermeidbare Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb: Dargestellt sind die Datenpunkte der Testsysteme (in verschiedenen Farben) sowie die Näherungsfunktionen der Stichprobe 2 (in Grün) und deren statistische Unsicherheit s (Hüllkurve in Hellgrün) für die Trefferrate TPR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Ausfälle des jeweiligen Systems $P(0 \leq t < B)$ normiert. Für die Näherungsfunktion der Stichprobe 2 wurde ein Polynom 2. Grades gewählt. Der ausgegraute Bereich stellt den Zeitraum $t = [0, 35.040[$ in Stunden h dar, in dem noch keine Datenerfassung erfolgt ist.

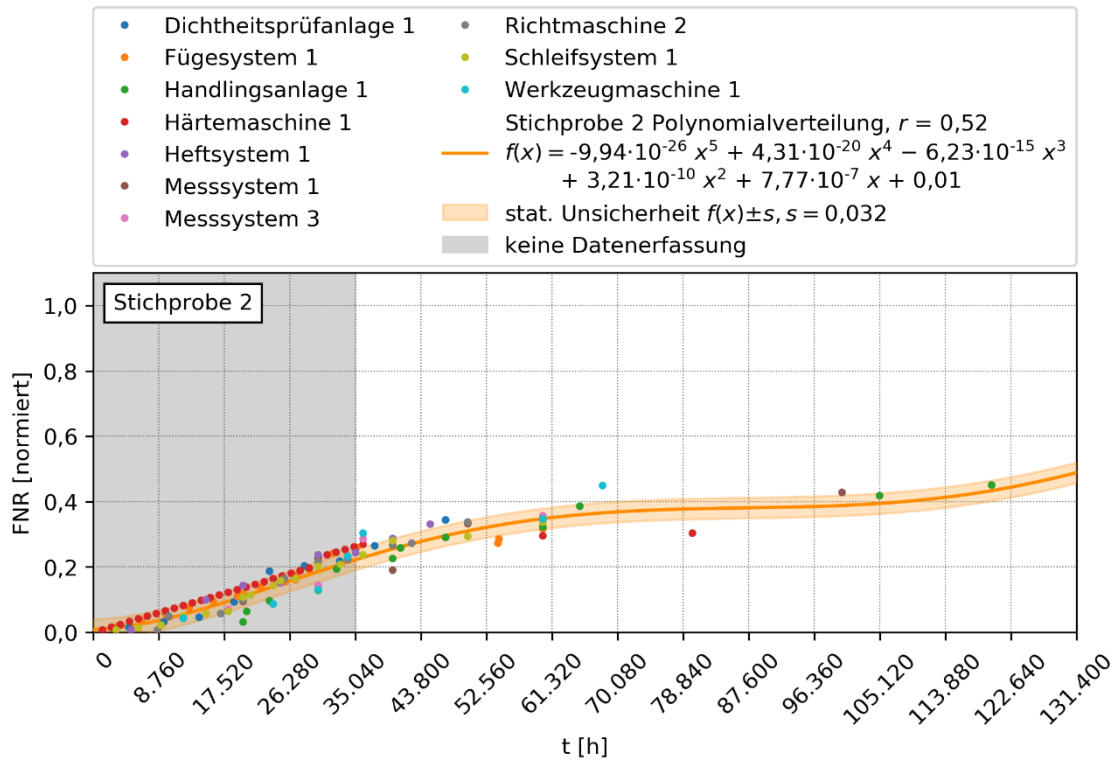


Bild A.13: Normierte Datenpunkte der Testsysteme und genäherte Verteilung der Stichprobe 2 mit statistischer Unsicherheit der Nicht-Alarmrate FNR für verbleibende Ausfälle entsprechend Szenario 2 mit Beginn der Datenerfassung nach 4 Jahren Betrieb: Dargestellt sind die Datenpunkte der Testsysteme (in verschiedenen Farben) sowie die Näherungsfunktionen der Stichprobe 2 (in Orange) und deren statistische Unsicherheit s (Hüllkurve in Hellorange) für die Nicht-Alarmrate FNR über den Betriebszeitraum t_B in Stunden h . Die Datenpunkte wurden mittels der Gesamtanzahl der Ausfälle des jeweiligen Systems $P(0 \leq t < B)$ normiert. Für die Näherungsfunktion der Stichprobe 2 wurde ein Polynom 5. Grades gewählt. Der ausgegraute Bereich stellt den Zeitraum $t = [0, 35.040[$ in Stunden h dar, in dem noch keine Datenerfassung erfolgt ist.

Literaturangaben

- [1] Daimler AG: Industrie 4.0 – Digitalisierung bei Mercedes-Benz: Die nächste Stufe der industriellen Revolution, 2015. <https://media.daimler.com/mars-MediaSite/de/instance/ko/Industrie-40--Digitalisierung-bei-Mercedes-Benz-Die-naechste-Stufe-der-industriellen-Revolution.xhtml?oid=9272047>, abgerufen am: 24.08.2018
- [2] North, K.: Wissensorientierte Unternehmensführung. Wissensmanagement gestalten. Wiesbaden: Springer Fachmedien 2016
- [3] Fisch, J., Can, A., Diedrich, C. u. Krueger, J.: Transformation der Wissenstreppe zum digitalen Fahrstuhl durch die Verwendung von Big Data-Analysen. Baden-Baden, 03. und 04. Juni. 19. Leitkongress der Mess- und Automatisierungstechnik. AUTOMATION 2018 (VDI-Berichte 2330)
- [4] Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0. Abschlussbericht des Arbeitskreises Industrie 4.0, Deutschlands Zukunft als Produktionsstandort sichern, acatec u. Forschungsunion, 2013
- [5] Manhart, K.: Was Sie über Maschinelles Lernen wissen müssen, 2018. <https://www.computerwoche.de/a/was-sie-ueber-maschinelles-lernen-wissen-muessen,3329560>, abgerufen am: 22.08.2018
- [6] Freie Presse: Wie Kleiner das neue Größer wurde. Chemnitzer Zeitung, S. 4
- [7] Ertel, W.: Grundkurs Künstliche Intelligenz. Eine praxisorientierte Einführung. Wiesbaden: Springer Vieweg 2013
- [8] Vogel-Heuser, B.: Herausforderungen und Anforderungen aus Sicht der IT und der Automatisierungstechnik. In: Bauernhansl, T., Hompel, M. t. u. Vogel-Heuser, B. (Hrsg.): Industrie 4.0 in Produktion, Automatisierung und Logistik. Anwendung, Technologien, Migration. Wiesbaden: Springer Vieweg 2014
- [9] ISO/IEC 29161:2016-08. Information technology - Data structure - Unique identification for the Internet of Things
- [10] Tutanch: Was ist das Industrial Internet of Things (IIoT)? Definition. <https://www.bigdata-insider.de/was-ist-das-industrial-internet-of-things-iiot-a-654986/>, abgerufen am: 08.04.2018
- [11] Fisch, J., Haeussler, B., Can, A. u. Diedrich, C.: Aufbau einer Schnittstelle für die dienstorientierte Automatisierung mittels Analytik-Entscheidungen. Parametrierung von Steuerungen im Cloud-Zeitalter. atp edition 2018, 4-5, S. 26–29
- [12] ISO/IEC TR 20547-5:2018. Information technology - Big data reference architecture - Part 5: Standard roadmap
- [13] Gartner: IT Glossary. www.gartner.com/it-glossary, abgerufen am: 09.08.2020
- [14] Luber, S.: BigData Insider. www.bigdata-insider.de, abgerufen am: 09.08.2020
- [15] Engel, P.: Entwicklung neuer Diagnoseverfahren auf der Basis multivariater Methoden und deren Anwendung in einem Condition Monitoring System, Universität Duisburg-Essen Diss. 2012
- [16] BITKOM: Welche Bedeutung hat Industrie 4.0 für Ihr Unternehmen? Industrie 4.0 - Umfrage zur Bedeutung in Deutschland 2018, Industrie 4.0 - Wo steht Deutschland?, Seite 3 2018. <https://de.statista.com/statistik/daten/studie/830769/umfrage/bedeutung-von-industrie-40-in-deutschland/>

- [17] Gandomi, A. u. Haider, M.: Beyond the hype. Big data concepts, methods, and analytics. *International Journal of Information Management* 35 (2015) 2, S. 137–144
- [18] VDI 3633:2013-12. Simulation von Logistik-, Materialfluss- und Produktionssystemen. Begriffe
- [19] DIN 25424-1:1981-09. Fehlerbaumanalyse. Methode und Bildzeichen
- [20] Schenk, M.: Instandhaltung technischer Systeme. Methoden und Werkzeuge zur Gewährleistung eines sicheren und wirtschaftlichen Anlagenbetriebs. Berlin, Heidelberg: Springer-Verlag 2010
- [21] Linß, G.: Qualitätssicherung. Technische Zuverlässigkeit: Lehr- und Arbeitsbuch. München: Carl Hanser Verlag 2016
- [22] VDI 3423:2011-08. Verfügbarkeit von Maschinen und Anlagen. Begriffe, Zeiterfassung und Berechnung
- [23] DIN 13306:2010-12. Instandhaltung - Begriffe der Instandhaltung
- [24] DIN 40041:1990-12. Zuverlässigkeit. Begriffe
- [25] DIN 60300-3-11:2010-05. Zuverlässigkeitsmanagement – Teil 3-11: Anwendungsleitfaden – Auf die Funktionsfähigkeit bezogene Instandhaltung
- [26] Huf, A.: Kumulative Lastermittlung aus Antriebsdaten zur Zustandsbewertung von Werkzeugmaschinenkomponenten, Diss. 2012
- [27] Daimler AG: Lastenheft Powertrain Teil I. Allgemeine Ausführungen. 2015
- [28] Neugebauer, R.: Werkzeugmaschinen. Aufbau, Funktion und Anwendung von spanenden und abtragenden Werkzeugmaschinen. Berlin, Heidelberg: Springer Vieweg 2012
- [29] Bauernhansl, T., Hompel, M. t. u. Vogel-Heuser, B. (Hrsg.): Industrie 4.0 in Produktion, Automatisierung und Logistik. Anwendung, Technologien, Migration. Wiesbaden: Springer Vieweg 2014
- [30] Fisch, J. u. Diedrich, C.: Methodische Untersuchung des Komplexitätsanstiegs von Produktionssystemen. *at - Automatisierungstechnik* 2018, 66(6), S. 449–455
- [31] Weck, M. u. Brecher, C.: Werkzeugmaschinen 4. Werkzeugmaschinen, Bd. 4. Berlin, Heidelberg: Springer Vieweg 2006
- [32] Weck, M. u. Brecher, C.: Werkzeugmaschinen 3. Mechatronische Systeme, Vorschubantriebe, Prozessdiagnose. Werkzeugmaschinen, Bd. 3. Berlin, Heidelberg: Springer Vieweg 2006
- [33] Lercher, B.: Konzeption und System einer Integrationsplattform zur Entwicklung von Werkzeugmaschinen. Dissertation (2008)
- [34] Favre-Bulle, B.: Automatisierung komplexer Industrieprozesse. Systeme, Verfahren und Informationsmanagement. Wien: Springer-Verlag 2004
- [35] DIN 31051:2012-09. Grundlagen der Instandhaltung
- [36] Eisenbach, D.: Künstliche Neuronale Netze zur Prognose von Zeitreihen
- [37] Berger, R.: Worin liegt Ihrer Einschätzung nach der Kundennutzen von "Predictive Maintenance"? *Predictive Maintenance - Kundennutzen in Deutschland 2017*, Predictive Maintenance, S. 7 Stand: 2017. <https://de.statista.com/statistik/daten/studie/755099/umfrage/kundennutzen-von-predictive-maintenance-in-deutschland/>

- [38] Fink, O.: Data-Driven Intelligent Predictive Maintenance of Industrial Assets. In: Smith, A. E. (Hrsg.): Women in Industrial and Systems Engineering. Key Advances and Perspectives on Emerging Topics, S. 589–605
- [39] Frenus u. T-Systems: Stimmen Sie den nachfolgend genannten Aussagen zu oder nicht zu? Predictive Maintenance - Zukunftsfähigkeit in Europa 2017, Customers' Voice: Predictive Maintenance in Manufacturing, Western Europe, S. 7 2017. <https://de.statista.com/statistik/daten/studie/819148/umfrage/zukunftsfahigkeit-von-predictive-maintenance-in-europa/>
- [40] Berger, R.: Inwiefern besteht in Ihrem Unternehmen bereits ein Produkt-/Leistungsangebot im Bereich "Predictive Maintenance"? Predictive Maintenance - Produkt-/Leistungsangebot in Deutschland 2017, Deutschland Stand: 2017. <https://sta.cir-mcs.e.corpintra.net/statistik/daten/studie/755090/umfrage/produkt-leistungsangebot-im-bereich-predictive-maintenance-in-deutschland/>
- [41] Das sind die größten Werkzeugmaschinen-Hersteller Europas. Exklusives Umsatzranking. <https://www.produktion.de/wirtschaft/das-sind-die-groessten-werkzeugmaschinen-hersteller-europas-387.html>
- [42] DMG Mori AG. <http://www.dmgmori.com/>, abgerufen am: 25.11.2019
- [43] GROB. <https://www.grobgroup.com/service-schulungen/maschinen-ueberpruefung/condition-monitoring/>, abgerufen am: 09.08.2020
- [44] United Grinding Digital Solutions. <https://www.grinding.ch/de/customer-care/dienstleistungen/digital-solutions/>, abgerufen am: 09.08.2020
- [45] Networking@EMO: Interview mit Dr. Andreas Kühne von der ANACISION, 2019. <https://www.emag.com/blog/interview-mit-dr-andreas-kuehne/>, abgerufen am: 03.11.2019
- [46] EMAG: Industrie 4.0 Kooperation: EMAG erwirbt Anteile am Datenanalyse-Spezialisten anacision, 2018. <https://www.emag.com/de/presse/neuigkeiten/detailansicht/article/industrie-4-0-kooperation-emag-erwirbt-anteile-am-datenanalyse-spezialisten-anacision.html>, abgerufen am: 03.11.2019
- [47] emag.com: Industry 4.0. <https://www.emag.com/industry-40.html>, abgerufen am: 03.11.2019
- [48] iXpanel - i4.0 ready. Das Cockpit für die einfache Integration der Maschine in Ihre Betriebsorganisation. <https://www.index-werke.de/de/produkte/software-steuerung/index-ixpanel/>, abgerufen am: 03.11.2019
- [49] Siemens: Das ist MindSphere. <https://new.siemens.com/global/de/produkte/software/mindsphere.html>
- [50] Heller: HELLER4Industry. Fit für Industrie 4.0: Durch Digitalisierung die Maschinenproduktivität steigern. <https://www.heller.biz/maschinen-und-loesungen/heller4industry/>, abgerufen am: 03.11.2019
- [51] CHIRON Werke GmbH & Co. KG: CHIRON DATAline. Die innovative Plattform für integrierte Maschinen- und Prozessdiagnose. <https://chiron.de/loesungen/smartline/dataline>, abgerufen am: 09.08.2020
- [52] Hermle Gruppe: HIMS - Das zentrale Überwachungstool. https://www.hermle.de/de/bearbeitungszentren/digitale_bausteine, abgerufen am: 09.08.2020
- [53] Schwäbische Werkzeugmaschinen GmbH: Technology. Machines. Systems. www.sw-machines.de, abgerufen am: 09.08.2020

- [54] Starrag Group: Website. www.starrag.com, abgerufen am: 25.11.2019
- [55] GF Machining Solutions: Website. www.gfms.com, abgerufen am: 26.11.2019
- [56] SUSHINSKAYA, S.: Mikron Group. Medien- und Analystenkonferenz 2018
- [57] DVS Technology Group: Website. <http://www.dvs-gruppe.com/>
- [58] Danobat Group. <https://www.danobatgroup.com/de/neuigkeiten/danobat-prasentiert-die-aktuellsten-entwicklungen-auf-der-amb>, abgerufen am: 26.11.2019
- [59] MARPOSS Monitoring Solutions GmbH: Website. www.artis.de, abgerufen am: 03.11.2019
- [60] GfM Gesellschaft für Maschinendiagnose mbH: Werkzeugmaschinen. <https://maschinendiagnose.de/condition-monitoring-werkzeugmaschinen.htm>, abgerufen am: 03.11.2019
- [61] Value-driven Intelligence with Voith OnCumulus. <http://voith.com/corpen/digital-solutions/oncumulus.html>
- [62] Predictive Maintenance. Service der Zukunft – und wo er wirklich steht, Feldmann, S., Lässig, R. u. Herweg, O., 2017
- [63] BearingPoint: Wie groß schätzen Sie folgende technische Herausforderungen für die erfolgreiche Umsetzung von "Predictive Maintenance" in Ihrem Unternehmen ein? Predictive Maintenance - Herausforderungen bei der Umsetzung im DACH-Gebiet 2017, Chancen und Herausforderungen von Predictive Maintenance in der Industrie, Seite 8 2017. <https://de.statista.com/statistik/daten/studie/801057/umfrage/herausforderungen-bei-der-umsetzung-von-predictive-maintenance-im-dach-gebiet/>
- [64] Siemens AG: MindSphere security model. Enabling customers to confidently operate in a secure cloud environment. Version 1.0 (2018)
- [65] Gierow, H.: Honda stoppt Produktion wegen Wanna Cry. <https://www.golem.de/news/ransomware-honda-stoppt-produktion-wegen-wanna-cry-1706-128491.html>, abgerufen am: 23.10.2020
- [66] Daimler: eXtollo Unsere Big Data Plattform. <https://www.daimler.com/karriere/ueber-uns/artificial-intelligence/community-und-tools/extollo.html>, abgerufen am: 23.10.2020
- [67] Mohamed, K. S.: Neuromorphic Computing and Beyond: Parallel, Approximation, Near Memory, and Quantum. Springer 2020
- [68] Veltjens, K.: Big Data Analytics erfordert eine verlässliche Datenbasis, 2015. <https://ap-verlag.de/big-data-analytics-erfordert-eine-verlaessliche-datenbasis/16192/>, abgerufen am: 27.11.2019
- [69] Daily, J. u. Peterson, J.: Predictive Maintenance: How Big Data Analysis Can Improve Maintenance. In: Richter, K. u. Walther, J. (Hrsg.): Supply chain integration challenges in commercial aerospace. A comprehensive perspective on the aviation value chain. Cham: Springer 2017
- [70] Gluchowski, P., Schieder, C., Gmeiner, A. u. Trenz, S.: Automatisierung von Geschäftsprozessen im Maschinen- und Anlagenbau – Fallstudie zu Predictive Maintenance. HMD Praxis der Wirtschaftsinformatik 56 (2019) 5, S. 986–999
- [71] APA-OTS: Von Fraunhofer Austria mitentwickelte Instandhaltungs-App wacht über Industrieanlagen. https://www.ots.at/presseaussendung/OTS_20201006_OTS0125/von-fraunhofer-austria-mitentwickelte-instandhaltungs-app-wacht-ueber-industrieanlagen

- [72] Verein Deutscher Werkzeugmaschinenfabriken e.V.: umati. <https://vdw.de/technik-und-normung/umati/>, abgerufen am: 09.08.2020
- [73] Morrison, A. u. Rao, A.: Machine learning overview (infographic). <https://www.digitalpulse.pwc.com.au/infographic-machine-learning/>, abgerufen am: 09.08.2020
- [74] Hall, P., Dean, J., Kabul, I. K. u. Silva, J.: An Overview of Machine Learning with SAS® Enterprise Miner™. <https://support.sas.com/resources/papers/proceedings14/SAS313-2014.pdf>, abgerufen am: 22.08.2018
- [75] Heißmeyer, S., Renken, V., Overmeyer, L., Ströbel, G. u. Goch, G.: Voraussagende Instandhaltung mit der digitalen Maschinenakte - Maschine lernt mit. QZ Qualität und Zuverlässigkeit, Heft 01/2011, S. 54-55
- [76] Datenanalyse. Vorausschauende - Wartung optimieren mit KI. SCOPE - Industriemagazin für Produktion Heft 03/2020, S. 40
- [77] Instandhaltung als Teil eines ganzheitlichen Fertigungsleitsystems Planen auf Knopfdruck. Industrieanzeiger (2020) Heft 6, S. 64
- [78] Hönig, M. u. Zeiler, P., Prof. Dr.-Ing.: Predictive Maintenance: Big Data im Einsatz, 2019. <https://automationspraxis.industrie.de/industrie-4-0/predictive-maintenance-big-data-im-einsatz/>, abgerufen am: 09.08.2020
- [79] Viel Spielraum für KI-Anwendungen. Vorausschauende Wartung unter Aspekten künstlicher Intelligenz (KI), Teil 2. KEM Konstruktion Elektronik Maschinenbau (2019) Heft 11, S. 62
- [80] Wilkens, U., Lins, D., Prinz, C. u. Kuhlenkötter, B.: Lernen und Kompetenzentwicklung in KI-basierten adaptiven Systemen. Industrie Management (2020) Nr. 6, S. 30–34
- [81] Busse, A., Lauer, J. u. Metternich, J.: Nutzenorientierter Einsatz digitaler Systeme zur Prozessüberwachung. Implementierungsansatz zur Umsetzung „digitaler Services“ in der Instandhaltung mit Hilfe maschinellen Lernens. Productivity Management (2018) 1, S. 28–31
- [82] Kempe, S.: Häufige Muster in zeitbezogenen Daten, Univ., Fak. für Informatik Diss. Magdeburg, Magdeburg 2008
- [83] Bink, R. u. Zschech, P.: Predictive Maintenance in der industriellen Praxis. HMD Praxis der Wirtschaftsinformatik 55 (2018) 3, S. 552–565
- [84] Stefan Boschert u. Roland Rosen: Digitaler Zwilling - Der Simulationsaspekt. In: Hehenberger, P. u. Bradley, D. (Hrsg.): Mechatronic Futures. Challenges and Solutions for Mechatronic Systems and their Designers. Cham: Springer International Publishing 2016, S. 59-74
- [85] Herrmann, J.: Maschinelles Lernen. In: Herrmann, J. (Hrsg.): Maschinelles Lernen und wissensbasierte Systeme. Systematische Einführung mit praxisorientierten Fallstudien. Berlin, New York: Springer 1997, S. 15–21
- [86] Senatsverwaltung für Stadtentwicklung und Wohnen Berlin: Big Data und Crowd Data für die Berliner Stadtentwicklungsplanung
- [87] ISO/IEC 20922. Information technology — Message Queuing Telemetry Transport (MQTT) v3.1.1
- [88] Siemens: MQTT-Publisher (unverschlüsselt) für die S7-1500/S7-1200 und S7-300. <https://support.industry.siemens.com/cs/ww/de/view/109748872>, abgerufen am: 03.04.2018

- [89] Schmitt, T. L. u. Andres, M.: Methoden zur Modellbildung und Simulation mechatronischer Systeme. Wiesbaden: Springer Fachmedien Wiesbaden 2019
- [90] Cleff, T.: Deskriptive Statistik und Explorative Datenanalyse. Eine computergestützte Einführung mit Excel, SPSS und STATA. Wiesbaden: Gabler Verlag 2015
- [91] van der Aalst, Wil M.P.: Process Mining. Discovery, Conformance and Enhancement of Business Processes. Springer-Verlag Berlin Heidelberg 2011
- [92] Dinov, I. D.: Data Science and Predictive Analytics. Biomedical and Health Applications using R. Cham: Springer International Publishing 2018
- [93] Runkler, T. A.: Data Mining. Wiesbaden: Springer Fachmedien 2015
- [94] Schneider, B.: Neuronale Netze für betriebliche Anwendungen: Anwendungspotentiale und existierende Systeme. Arbeitsbericht
- [95] Understanding Transfer learning: For deep learning, 2018. <https://www.houseofbots.com/news-detail/4337-1-understanding-transfer-learning-for-deep-learning>, abgerufen am: 12.01.2019
- [96] Becker, R.: Transfer Learning - so können neuronale Netze voneinander Lernen, 12.02.18. <https://jaai.de/transfer-learning-1739/>, abgerufen am: 12.01.2019
- [97] VG 95287-3:2013-07. Entwicklung prüfbarer Produkte - Anwendung, Ziele und Leitfaden für die Erarbeitung der Leistungsbeschreibung, Verfahrensnorm. Beiblatt 3: Anwendungsrichtlinie für die Ermittlung von Vorgabewerten für MTBF und Ausfallraten von Systemen und Komponenten
- [98] Bertsche, B. u. Lechner, G.: Zuverlässigkeit im Fahrzeug- und Maschinenbau. Ermittlung von Bauteil- und System-Zuverlässigkeiten. VDI-Buch. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg New York 2004
- [99] Glöckler, M.: Simulation mechatronischer Systeme. Grundlagen und Beispiele für MATLAB und Simulink. Lehrbuch. Wiesbaden: Springer Vieweg 2018
- [100] Willink, R.: Guide to the Expression of Uncertainty in Measurement. In: Willink, R. (Hrsg.): Measurement uncertainty and probability. Cambridge: Cambridge University Press 2013, S. 237–244

Verzeichnis der betreuten studentischen Arbeiten

- Auris, F. J.: Analyse kausaler Datenzusammenhänge in einer Werkzeugmaschine, Bachelorarbeit 2015
- Bingöl, C.: Bestimmung der Komplexität von Produktionsanlagen im Anlagenentstehungsprozess, Bachelorarbeit 2017
- Boronenko, K.: Entwicklung einer Methode zur Ermittlung der Soll-Referenz aus Produktionsdaten, Masterarbeit 2016
- Esselun, D.: Analyse des hypothetischen Komplexitätsanstiegs von Produktionsanlagen im Zuge der Digitalisierung, Bachelorarbeit 2017
- Hille, B.: Entwicklung einer Methode zur automatischen Überprüfung der Taktzeitmessungen von Produktionsanlagen, Masterarbeit 2017
- Kontradenko, O.: Identifikation von ähnlichen Systemzuständen eines Mehrgrößensystems durch maschinelles Lernen am Beispiel eines Fertigungssystems, Masterarbeit 2019
- Krupschtat, J.: Analyse von SPS-Projekten für die zukünftige Standardisierung, Bachelorarbeit 2014
- Lex, P.: Eignungsanalyse von Visualisierungselementen in grafischen Benutzeroberflächen durch Kennzahlengruppierung, Bachelorarbeit 2018
- Remlein, A.: Methode zur Gruppierung von Produktionsprozessen und Komponenten, Bachelorarbeit 2016
- Stephan, P.: Konzeption eines Dienstes zur automatisierten Parametrierung und Steuerung der Produktion über eine IoT-Plattform, Masterarbeit 2019
- Strigli, M.: Implementierung eines anhand von Useranforderungen konzeptionierten Cloud-Services für Industrieroboter auf einer Analyseplattform zur Steigerung der Gesamtanlageneffektivität in der Motorenkomponentenfertigung, Masterarbeit 2019
- Werner, N. G.: Erfassung von Prozessdaten an einer Werkzeugmaschine und Analyse mittels statistischer Methoden, Bachelorarbeit 2015
- Wolf, T.: Analyse des energetischen Optimierungspotenzials von Industrierobotern anhand des KUKA „Energy Efficiency“-Konzeptes, Bachelorarbeit 2015

Ich danke allen Studenten, deren Abschlussarbeit ich betreuen durfte, und weiterhin möchte ich mich auch bei meinen ehemaligen Praktikanten und Werkstudenten Alexandra Arkhipova, Sarah Balzuweit, Andreas Daubenschütz, Feyza Güney, Tuan Hoang, Hannes Kunstmann, Daniel Maelzer, Yann-Paul Marsch, Alona Reshytko, Rene Stieger, Hüseyin Türkyilmaz, Christian Wegner, Leo Wilms und Li Zeng für die gute Zusammenarbeit bedanken.

Ehrenerklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die Hilfe eines kommerziellen Promotionsberaters habe ich nicht in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Verwendete fremde und eigene Quellen sind als solche kenntlich gemacht.

Ich habe insbesondere nicht wissentlich:

- Ergebnisse erfunden oder widersprüchliche Ergebnisse verschwiegen,
- statistische Verfahren absichtlich missbraucht, um Daten in ungerechtfertigter Weise zu interpretieren,
- fremde Ergebnisse oder Veröffentlichungen plagiiert,
- fremde Forschungsergebnisse verzerrt wiedergegeben.

Mir ist bekannt, dass Verstöße gegen das Urheberrecht Unterlassungs- und Schadensersatzansprüche des Urhebers sowie eine strafrechtliche Ahndung durch die Strafverfolgungsbehörden begründen kann.

Ich erkläre mich damit einverstanden, dass die Dissertation ggf. mit Mitteln der elektronischen Datenverarbeitung auf Plagiate überprüft werden kann.

Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form als Dissertation eingereicht und ist als Ganzes auch noch nicht veröffentlicht.

Berlin, 23.02.2022

Jessica Fisch