

28. VDI-Fachtagung

Technische Zuverlässigkeit 2017

Entwicklung und Betrieb
zuverlässiger Produkte



VDI-Berichte 2307

VDI-BERICHTE

Herausgeber: VDI Wissensforum GmbH

28. VDI-Fachtagung

Technische Zuverlässigkeit 2017

Entwicklung und Betrieb
zuverlässiger Produkte

Leonberg bei Stuttgart, 17. und 18. Mai 2017



VDI-Berichte 2307

Bibliographische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind im Internet unter <http://dnb.ddb.de> abrufbar.

Bibliographic information published by the Deutsche Nationalbibliothek
(German National Library)

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliographie (German National Bibliography); detailed bibliographic data is available via Internet at <http://dnb.ddb.de>.

© VDI Verlag GmbH · Düsseldorf 2017

Alle Rechte vorbehalten, auch das des Nachdruckes, der Wiedergabe (Photokopie, Mikrokopie), der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, auszugsweise oder vollständig.

Der VDI-Bericht, der die Vorträge der Tagung enthält, erscheint als nichtredigierter Manuskriptdruck. Die einzelnen Beiträge geben die auf persönlichen Erkenntnissen beruhenden Ansichten und Erfahrungen der jeweiligen Vortragenden bzw. Autoren wieder.

Printed in Germany.


ISSN 0083-5560

ISBN 978-3-18-092307-9





Vorwort

1


Plenarvortrag



<i>F. Pasch,</i> <i>W. Hartewieg,</i> <i>D. Kirschmann</i>		Das Ende des zufälligen Zufalls – Warum scheinbar zufällige Feldbeanstandungen systematisch sein können	3
--	---	---	---

Prognostics & Health Management



<i>W. Sextro,</i> <i>J. K. Kimotho,</i> <i>T. Kaul,</i> <i>T. Meyer,</i>		Entwicklung verlässlicher, intelligenter Systeme – Herausforderungen und Chancen	17
<i>J. Fisch,</i> <i>M. Rossdeutscher,</i> <i>C. Diedrich</i>		Anwendung datenbasierter Methoden auf Werkzeugmaschinenendaten zur Abweichungserkennung vom Normalbetrieb	31
<i>T. Rieker,</i> <i>M. Bartholdt,</i> <i>B. Bertsche,</i> <i>P. Zeiler,</i> <i>A. Jacobi</i>		Zuverlässigkeitsprognose in der Anlaufphase für „x-Months-in-Service“: Präzisierung und Validierung der Beanstandungsquote	43
<i>C. Döbel</i>		Einsatz textiler Sensoren für die Lebensdauer-vorhersage von Maschinen	55

Modellierung und Simulation von Zuverlässigkeit



<i>F. Müller,</i> <i>P. Zeiler,</i> <i>B. Bertsche</i>		Bootstrap-Monte-Carlo-Simulation von Zuverlässigkeit und Aussagewahrscheinlichkeit bei periodischer Instandhaltung	69
<i>C. Hayer,</i> <i>S. Fiebig,</i> <i>T. Vietor,</i> <i>J. Sellschopp</i>		Robustheitsoptimierung innerhalb des Entwicklungsprozesses durch Integration von Fertigungstoleranzen in die Simulation	83

<i>S. Kemmler, M. Dazer, T. Lepold, B. Bertsche</i>		Lebensdaueranalyse auf Basis von multidimensionaler Zuverlässigkeits- und Robust Design Simulation – Integrale Betrachtung der Robusten Zuverlässigkeit	97
<i>J. Heinrich, F. Plinke, J. Hauschild</i>		Zustandsbasierte Sicherheits- und Verfügbarkeitsanalyse unterschiedlich automatisierter Fahrfunktionen mittels Monte-Carlo-Simulation	109




Zuverlässigkeit und Big Data

<i>D. Ortmann, C. Lisiecki</i>		Vorhersage von Verfügbarkeitszahlen in der Entwicklungsphase neuer Produktgenerationen	125
<i>P. Dobry, A. Jacobi, B. Bertsche</i>		Anwendung von ausgewählten multivariaten Analysemethoden des Data Mining zur Identifizierung von ausfallverbundenen Fahrverhaltensmustern auf Basis von Kundenbelastungskollektiven	139






Zuverlässigkeit komplexer Systeme

<i>S. Ochs, E. M. Slomski, T. Melz</i>		Stochastische Sensitivitätsmaße für smarte Systeme	151
<i>F. Long, P. Zeiler, B. Bertsche</i>		Prognose der Verfügbarkeit von flexiblen Produktionssystemen in der Industrie 4.0	163

Qualität und Zuverlässigkeit und funktionale Sicherheit

<i>O. Bielefeld, H. Dransfeld, N. Schlüter</i>		Entwicklung eines Vorgehenskonzeptes zur Analyse von Fehlerfolgeketten in komplexen Systemen	177
<i>A. Jacobi, C. Jordan, E. Kaganova</i>		Zuverlässigkeitsmanagement für den Anlauf von Neuproduktprojekten in der Automobilindustrie – Zuverlässigkeit bestimmen und messen	189
<i>O. El Adlouni, H. Schäbe</i>		Nachweis der funktionalen Sicherheit für eine Abschalt- und Erdungsautomatik (AEA) mittels Bahnspezifischer Normen EN 50126, -28, -29.	201

Zuverlässigkeitstests und besonders beschleunigte Verfahren

<i>M. Bartholdt, M. Bollmann, P. Zeiler, B. Bertsche</i>		Absicherung der Gesamtzuverlässigkeit eines Getriebes unter Berücksichtigung von Vorkenntnissen – Anwendung und Validierung	217
<i>A. Romer, J.-M. Veith</i>		Berücksichtigung sich ändernder Steuergeräteparameter auf entstehende Beanspruchungen für Zuverlässigkeitstests durch Monte-Carlo Simulation	229
<i>T. Buschhaus, P. Glöckner, M. Metzele</i>		Berücksichtigung inhomogener Anforderungen in der Zuverlässigkeitsabsicherung am Beispiel eines Fahrwerksystems eines 1st-Tier Zulieferers der Nutzfahrzeugindustrie	241
<i>A. Müller, M. Hinz, S. Bracke</i>		Auslegung von Zeitraffertests auf Basis numerischer Simulationen im Rahmen der Dentalimplantat-Erprobung	253
<i>M. Stohrer, K. Lucan, B. Bertsche</i>		Planung eines Zuverlässigkeits-DoE im Spannungsfeld zwischen Test-Design und Erfolgswahrscheinlichkeit	269

Vorwort

Die Sicherstellung einer von Kundenseite gewünschten hohen Zuverlässigkeit erfordert im Zeitalter der Industrie 4.0 und der zunehmenden Komplexität der Systeme eine ständige Weiterentwicklung bekannter Ansätze. Letztendlich muss die Zuverlässigkeit eines Produktes über den gesamten Produktlebenslauf betrachtet werden und beherrschbar sein.

Derzeitige Bemühungen betrachten sowohl die Weiterentwicklung von analytischen Zuverlässigkeitsmethoden als auch umfassendere Datenanalysen und optimierte Zuverlässigkeitstests.

Die VDI Wissensforum GmbH und die VDI-Gesellschaft Produkt- und Prozessgestaltung (GPP) führt bereits **zum 28. Mal die Tagung Technische Zuverlässigkeit (TTZ)** durch.

Der vorliegende Tagungsband enthält die Manuskripte der Referenten, soweit sie zum Zeitpunkt der Drucklegung vorlagen. Vorher fand ein Review des Programmausschusses statt.

Der Tagungsleiter und die Mitglieder des Programmausschusses danken allen, die beim Gelingen der Veranstaltung mitgewirkt haben.

Wir freuen uns, Sie auf der 28. Fachtagung „Technische Zuverlässigkeit“ am 17. bis 18. Mai 2017 in Leonberg begrüßen zu dürfen.

Im Namen des Programmausschusses

Tagungsleiter



Prof. Dr.-Ing. Bernd Bertsche

Institutsleiter, Institut für Maschinenelementen (IMA)

Universität Stuttgart

Fachlicher Träger

Die VDI-Gesellschaft Produkt- und Prozessgestaltung bietet mit ihren Fachbereichen für alle Branchen abgesichertes Wissen zur Gestaltung von Produkten und Prozessen sowie deren Optimierung bezüglich Qualität, Zeit und Kosten-Nutzenverhältnis.

Fachbereich Sicherheit und Zuverlässigkeit

Der Fachbereich Sicherheit und Zuverlässigkeit widmet sich der Weiterentwicklung der Methoden zur Analyse, Bewertung und Prognose der Zuverlässigkeit. Handlungsempfehlungen in Form von VDI-Richtlinien, Erfahrungsaustausch und Veranstaltungen sind Ergebnisse der Aktivitäten.

www.vdi.de/zuverlaessigkeit

Das Ende des zufälligen Zufalls

Warum scheinbar zufällige Feldbeanstandungen systematisch sein können



Dipl.-Ing. **Florian Pasch**,

TU Berlin, Fachgebiet Qualitätswissenschaft, Berlin;

Dipl.-Ing. **Wilfried Hartewieg**,

Volkswagen AG, Zentrale Konzern-Qualitätssicherung, Wolfsburg;

Dr.-Ing. **Daniel Kirschmann**,

Robert Bosch GmbH, Engineering Design Power Electronics, Reutlingen

Kurzfassung

Die Zuverlässigkeit von Produkten im Feldeinsatz ist ein Erfolgsfaktor für Unternehmen. Abweichungen von der geplanten und erprobten Zuverlässigkeit der Produkte bzw. Systeme im Feld und damit deren Bauteile müssen schnell und sicher identifiziert werden. In der dazu gehörenden Feldbeobachtung werden Ausfalldaten ausgewertet. Erfolgen diese Auswertungen nur auf Systemebene aber nicht auf Bauteilebene, mischen sich die Fehlermechanismen der Bauteile und ergeben bei einer Systembetrachtung ein Ausfallverhalten, das unter Umständen als „zufällig“ erscheinen kann. Folgen hiervon sind zum Beispiel, dass Alterungsmechanismen (z.B. Verschleiß) verspätet erkannt und Abstellmaßnahmen nicht rechtzeitig eingeleitet werden können.

Dieser Sachverhalt wird in diesem Beitrag mittels einer Simulationsstudie diskutiert und erläutert. Es wird aufgezeigt, unter welchen Umständen sich eine Schar von Fehlermechanismen bei bestimmten Anteilen von Früh-, Zufalls- sowie Alterungsausfällen innerhalb einer Grundgesamtheit in Abhängigkeit von bestimmten Auswertungszeitpunkten (Einsatzmonate) zu einem scheinbar zufälligen Ausfallverhalten (Weibull-Parameter $\beta \approx 1$) ergeben.

Als Ergebnis zeigt sich, dass ein $\beta \approx 1$ auf Systemebene tatsächlich zustande kommt, wenn mehrere Systeme betrachtet werden und Anteile dieser Grundgesamtheit sowohl von Früh- als auch von Alterungsausfällen betroffen sind. Dabei dürfen diese Anteile nicht zu dominant sein. Entsprechend verbergen sich unterschiedliche Fehlermechanismen auf Bauteilebene in den Systemen, die bei einer Auswertung auf Systemebene ein scheinbar „zufälliges“ Ausfallverhalten ergeben. Eine Auswertung sollte also immer auf Bauteilebene und nicht auf Systemebene erfolgen, um das „wahre“ Ausfallverhalten eines Systems beschreiben zu können.

1. Ausgangssituation

Die Sicherstellung der geforderten Zuverlässigkeit für alle Anwendungsfälle von komplexen Produkten ist im Rahmen der Entwicklung nur eingeschränkt möglich, da nicht alle späteren Beanspruchungen und Belastungen in der Nutzungsphase vollständig bekannt sind. In der Entwicklungsphase werden Annahmen zu Belastungen in der späteren Nutzungsphase (z.B. Gebrauchs- als auch Umweltbedingungen) sowie Annahmen zur erzielbaren Fertigungsqualität (Beanspruchbarkeit) getroffen [1]. Entsprechend werden in der Entwicklung zum Nachweis der Zuverlässigkeit System- (Fahrzeug-), Subsystem- und Komponententests unter diesen Annahmen durchgeführt sowie mit verschiedenen Methoden der Nachweis einer definierten Ziel-Zuverlässigkeit unter gegebenen Randbedingungen geführt (Methoden dazu sind zu finden in [1-3]). So werden beispielsweise Fahrzeug-Komponenten unter der Annahme von wirkenden Belastungen durch unterschiedliche Nutzerbedingungen (Mischkollektiven) mit einer definierten Bauteilqualität (Beanspruchbarkeit) entwickelt und getestet. Allerdings sind die Tests oft nur gerafft mit einer begrenzten Stichprobengröße im Verhältnis zur Serienproduktion möglich, so dass nicht alle möglichen Aspekte (insbesondere Alterung) in ihrem vollen Ausmaß erkannt werden können. Bilden sich zusätzlich in der Nutzung der Fahrzeuge andere spezifische Nutzungsbedingungen aus als angenommen (Verschiebung der Anteile im Mischkollektiv) oder zeigt sich eine unzureichende Beanspruchbarkeit der Komponenten (z.B. Produktionsschwankungen), kann dies zu Fehlerarten bei Systemen führen, die in der Entwicklung nicht bekannt waren. Die Anzahl der tatsächlich auftretenden Fehlerarten sowie deren Häufigkeiten sind entsprechend abhängig von den tatsächlichen Nutzungsbedingungen als Mischbeanspruchung sowie der realen Bauteil-Qualität, die oft chargenabhängig ist.

2. Analyse von Feldbeanstandungen

Zur Systemverbesserung sind die Informationen aus der Nutzung durch statistische Analysen von Felddaten zeitnah nach Produkteinführung auszuwerten. Oft dominieren in dieser frühen Phase Früh- und Zufallsausfälle, aber auch erste Alterungsausfälle treten an Systemen auf, die extrem beansprucht oder Belastungen ausgesetzt waren, die in der Erprobung nicht berücksichtigt werden konnten. Prinzipiell sollten die unterschiedlichen Fehlermechanismen einzeln (homogene Grundgesamtheit) unter Verwendung eines unimodalen Ausfallmodells ausgewertet werden, um das gesamte Ausfallverhalten mit Hilfe einer statistischen Verteilungsfunktion approximieren zu können [4]. Im industriellen Umfeld findet die Weibullverteilung am häufigsten Verwendung zur Beschreibung dieses Ausfallverhaltens [1]. Sie bietet die Besonderheit, dass sich anhand dessen Formparameter β die Art des Fehlerme-

chanismus charakterisieren lässt. So können Fehlermechanismen mit einem $\beta < 1$ als Früh-ausfall, mit einem $\beta = 1$ als Zufallsausfall und mit einem $\beta > 1$ als Alterungsausfälle (z.B. Verschleiß) klassifiziert werden [5]. Leider ist es oft nicht möglich oder praktikabel, einzelne Ausfälle aus dem Feld spezifischen Fehlerbildern (Komponente mit dessen Fehlermecha-nismus) zuzuordnen. Dies führt dazu, dass Ausfalldaten von verschiedenen Fehlermecha-nismen bewusst oder unbewusst zur Bestimmung des Ausfallverhaltens genutzt werden [6]. Findet nun eine Approximation dieser Ausfalldaten durch eine einzelne Verteilung wie der Weibullverteilung statt, kann dies sehr unpräzise sein. Es zeigen sich häufig hohe Abwei-chungen zwischen der angenommenen Verteilung und den Ausfalldaten [7]. Oft ist grafisch schon direkt erkennbar (durch z.B. Steigungswechsel in der doppel-logarithmischen Darstel-lung), dass es sich um eine Mischverteilung handelt. In diesem Fall sollten die Daten nach Fehlerbildern klassifiziert werden, um mit unimodalen Ausfallmodellen eine sinnvolle Ab-schätzung des Ausfallverhaltens zu ermöglichen. Ist dies nicht möglich, können spezifische Modelle oder Verfahren zur Auswertung von Mischverteilungen genutzt werden.

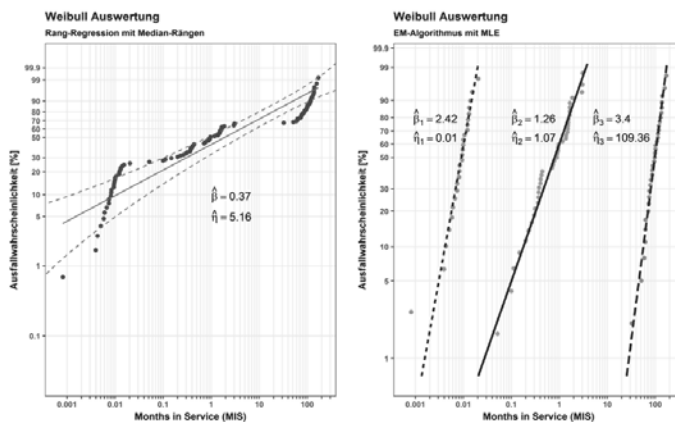


Bild 1: Unimodale Weibullauswertung aller Ausfalldaten von verschiedenen Fehler-mechanismen (links) und drei Weibullauswertung nach Fehlermechanismen getrennt (rechts).

Die Abbildung 1 zeigt eine Auswertung mit simulierten Ausfalldaten von verschiedenen Fehlermechanismen sowie eine nach Fehlermechanismen getrennte Auswertung dieser Daten anhand einer automatisierten Klassifizierung der Daten durch einen Expectation-

Maximization-Algorithmus [8]. Bei dieser Auswertung ist die Mischverteilung durch drei unterschiedliche Fehlerarten grafisch leicht ersichtlich. Problematisch wird es jedoch, wenn sich viele Fehlerbilder innerhalb einer Auswertung vermischen. Diese Ausfalldaten ergeben dann oft eine Weibullverteilung mit einem $\beta \approx 1$, was einem „mehr oder weniger“ zufallsbedingtes Ausfallverhalten entspräche. Wird dieses bei der Interpretation des analysierten Ausfallverhaltens nicht kritisch hinterfragt, können systematische Fehlermechanismen wie Alterungsausfälle in der Menge der unterschiedlichen Fehlerbilder übersehen werden. Dabei ist gerade die Erkennung von alterungsbedingten Ausfällen in frühen Serienphasen dringend notwendig, um geeignete Maßnahmen zur Reduzierung der Fehlerfolgen frühzeitig ergreifen zu können.

In der Fachliteratur wird dieser Sachverhalt selten diskutiert sowie in der praktischen Arbeit häufig nicht berücksichtigt. Insbesondere [5] geht kurz auf diesen Sachverhalt ein und schildert, dass die Weibullverteilung für viele gemischte Fehlermechanismen sich zu einem $\beta \approx 1$ ergeben und entsprechend äquivalent zur Exponentialverteilung ist. Allerdings führt die Anwendung der Exponentialverteilung oft zu Fehlinterpretationen. So kritisiert [5], dass Standards wie das MIL-HDBK-217 [9] für das Beschreiben des Ausfallverhaltens von elektronischen Bauteilen die Exponentialverteilung nutzen, obwohl Felddaten für Früh- oder Verschleißausfälle sprechen [5]. Allerdings geht [5] nicht genauer darauf ein, unter welchen Umständen sich ein $\beta \approx 1$ ergibt. Um diesen Sachverhalt zu untersuchen wird nachfolgend ein Beispiel vorgestellt, das die Problematik der gemeinsamen Auswertung von vielen Fehlerbildern verdeutlicht. Daran anschließend wird durch eine Simulationsstudie aufgezeigt, mit welchen Anteilen an Früh-, Zufalls- sowie Alterungsausfällen einzelner Bauteile in der Gesamtpopulation von Systemen mit wie vielen unterschiedlichen Fehlerbildern und zu welchen bestimmten Auswertungszeitpunkten sich ein solch scheinbar zufälliges Ausfallverhalten (Weibull-Parameter $\beta \approx 1$) ergibt.

3. Beispiel für scheinbar zufällige Feldbeanstandungen $\beta \approx 1$

Als Beispiel werden hier 5.000 Systeme betrachtet, die je 100 Bauteile besitzen. Diese Bauteile können unabhängig voneinander mit unterschiedlichen Fehlermechanismen behaftet ausfallen. Das Ausfallverhalten der Systeme wird nach einem Einsatzzeitraum von 6, 24 und 60 Monaten analysiert. Es werden nur Erstaussfälle betrachtet. Bei einem Defekt des Systems wird auf eine Reparatur verzichtet.

Die erste Auswertung der Feldbeanstandungen findet nach sechs Monaten (6 MIS) der Systeme im Feldeinsatz statt. Die Weibullauswertung (Abbildung 2 linke Grafik) ergibt einen Anteil von beanstandeten Systemen von etwas mehr als 1% mit einem β -Wert von 0,8, was auf Frühausfälle hinweist. Eine genauere Fehlerbetrachtung bzw. Analyse der Fehlermechanismen zeigt, dass fast 50% aller Beanstandungen zu diesem Zeitpunkt auf 2 von bisher 16 beanstandeten Bauteilen zurückzuführen sind. Die Abbildung 2 (rechte Grafik) zeigt einen Pareto-Plot, der die Häufigkeit von Beanstandungen zum Zeitpunkt 6 MIS für die einzelnen Bauteile darstellt. Dieser zeigt deutlich, dass Frühausfälle die Beanstandungen dominieren, jedoch auch schon erste verschleißbedingte Ausfälle bei weiteren Bauteilen sichtbar werden.

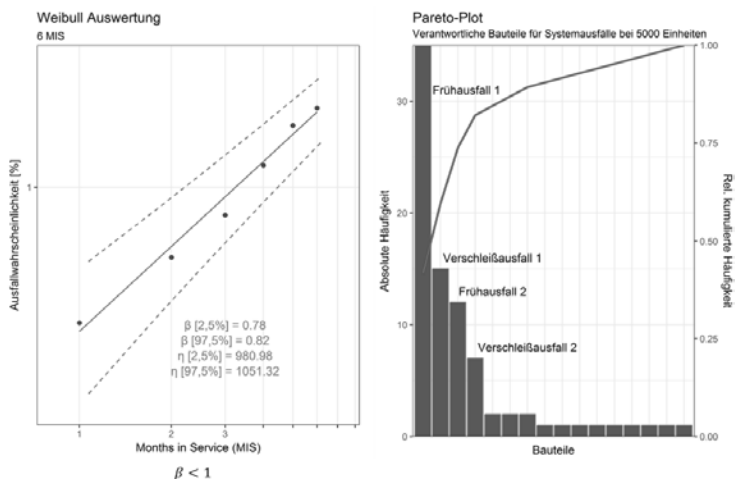


Bild 2: Die Weibullauswertung (links) sowie der Pareto-Plot (rechts) für die Feldbeanstandungen des Systems nach 6 MIS.

Eine weitere Weibullauswertung wird nach 24 MIS für das Gesamtsystem (Abbildung 3 linke Grafik) durchgeführt. Der Anteil der ausgefallenen Systeme liegt bei etwas mehr als 5% mit einem β -Wert nicht signifikant unterschiedlich von 1. Der Pareto-Plot zeigt, dass nach 24 MIS vor allem zwei verschleißbedingte Ausfallbilder dominieren, die in der ersten Auswertung durch die Dominanz der Frühausfälle nicht besonders auffällig waren.

Beginnt man erst zu diesem Zeitpunkt (24 MIS) die Analyse der verschleißbedingten Ausfallbilder verliert man wertvolle Fehlerabstellzeit. Problematisch wird dieser Sachverhalt dann im Zeitraum bis 60 MIS. Hier steigt die Häufigkeit der Beanstandungen deutlich an und macht einen Sprung von etwas über 5% nach 24 MIS auf etwa 17% nach 60 MIS. Auch hier liefert eine Weibullanalyse auf Systemebene über 60 MIS keine Anhaltspunkte für Probleme auf Bauteilebene, die abgestellt werden müssten. Es ergibt sich zwar eine höhere Ausfallhäufigkeit, allerdings mit einem β -Wert nicht signifikant unterschiedlich von 1. Eine detailliertere Analyse der einzelnen Fehlermechanismen ergibt jedoch, dass inzwischen vor allem verschleißbedingtes Ausfallverhalten von verschiedenen Bauteilen für die hohe Beanstandungshäufigkeit verantwortlich ist. Dabei sind zwei Bauteile besonders stark von Verschleiß betroffen.

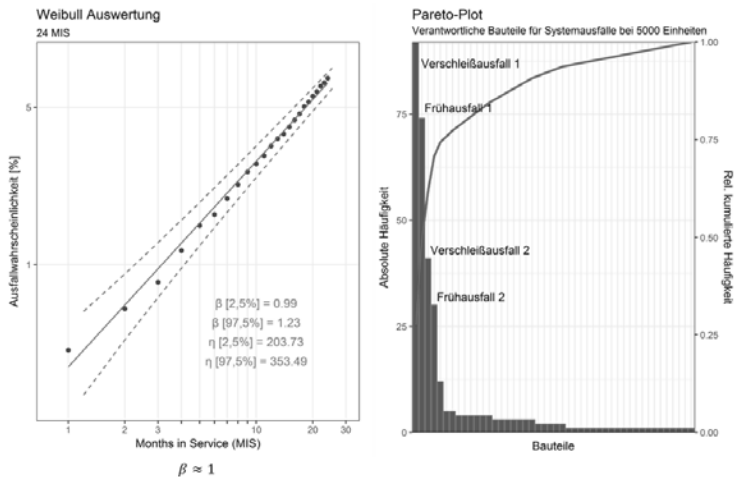


Bild 3: Die Weibullauswertung (links) sowie der Pareto-Plot (rechts) für die Feldbeanstandungen des Systems nach 24 MIS.

Die Alterungsmechanismen wären bei detaillierter Fehleranalyse auch schon nach 6 MIS erkennbar gewesen, werden aber durch die dominanten Frühausfälle in der Gesamtauswertung überdeckt. Die Abbildung 4 zeigt die Weibullanalyse nach 60 MIS sowie die Entwicklung der zwei relevanten Ausfallbilder Frühausfall sowie Verschleiß über MIS. Anhand der rechten Grafik in Abbildung 4 ist deutlich erkennbar, dass der Anteil der Beanstandungen aufgrund

von Verschleiß über der Zeit deutlich zunimmt wohingegen der Anteil der Beanstandungen bedingt durch die Frühausfälle über der Zeit ein stabiles Niveau erreicht.

Dieses Beispiel macht deutlich, dass eine Auswertung von Ausfalldaten über verschiedene Fehlerbilder bei der Betrachtung auf Systemebene schnell zu falschen Schlussfolgerungen führen kann bzw. keinen Hinweis auf die tatsächlich wirkenden Fehlermechanismen auf Bauteilebene in dem System liefert, da ein β -Wert der Weibullverteilung nahezu 1 erscheint.

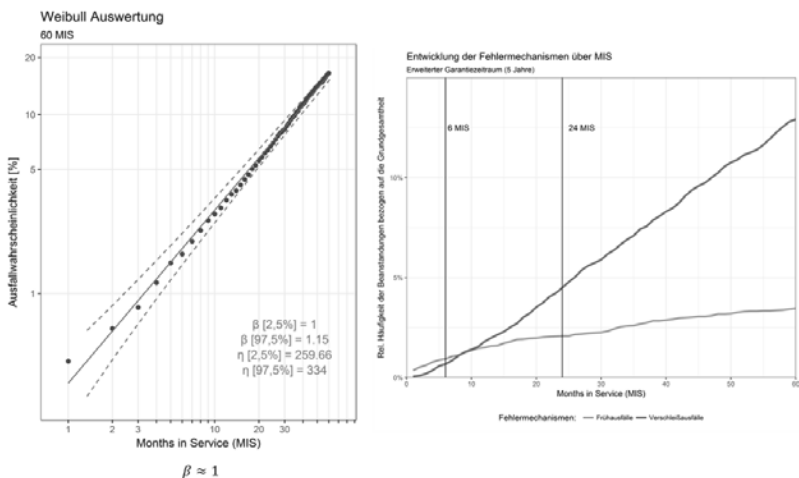


Bild 4: Die Weibullanalyse (links) für das System nach 60 MIS sowie die Entwicklung der Früh- sowie Verschleißausfälle über MIS

Um zu zeigen, dass dieses Beispiel keinen Sonderfall darstellt, wird nachfolgend anhand einer Simulationsstudie erarbeitet, dass Weibullanalysen für komplexe Systeme mit verschiedenen Fehlermechanismen unter Umständen keinen Hinweis auf das tatsächliche Ausfallverhalten liefern und zu einem $\beta \approx 1$ tendieren.

4. Simulationsstudie zur Weibullanalyse von Systemausfällen

Zur Simulation des Ausfallverhaltens von Systemen wurde das von [4] entwickelte Ausfallmodell herangezogen. Das Modell erlaubt die Berücksichtigung von Teilpopulationen, die von einem bestimmten Fehlermechanismus betroffen sind, welcher beim Rest der Population

nicht wirkt. So können einzelne Systeme aufgrund von Fehlermechanismen ausfallen, die nur Teile der Grundgesamtheit betreffen und z.B. Chargenprobleme darstellen könnten. Die Fehlermechanismen, die dann in den einzelnen Subpopulationen auf ein System wirken, sind als konkurrierend für den Ausfall modelliert. Dies bedeutet, dass ein Ausfall potentiell aufgrund von verschiedenen Fehlermechanismen zustande kommen kann, die unabhängig sind und zu jedem Zeitpunkt einen Ausfall verursachen könnten. Jeder dieser Fehlermechanismen und dessen Ausfallzeitpunkte wird durch eine eigene Wahrscheinlichkeitsverteilung beschrieben [4].

Die Simulation wurde entsprechend diesem Modell aufgebaut. Die Abbildung 5 zeigt schematisch den Aufbau der Simulation mit beispielsweise 10 Systemen und 20 Bauteilen pro System sowie einem Fehlermechanismus pro Bauteil.

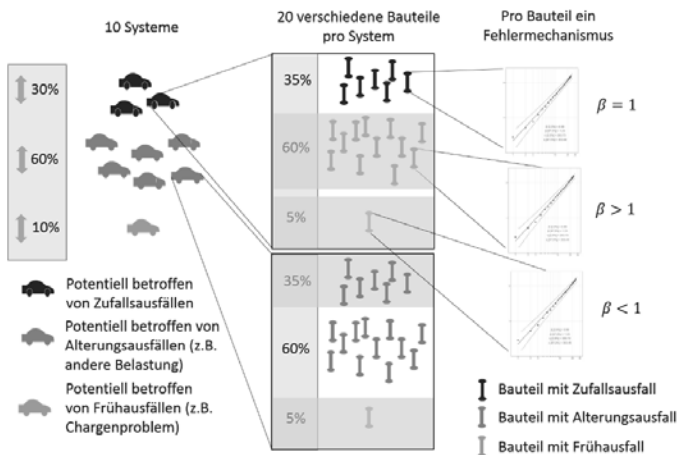


Bild 5: Prinzipieller Aufbau der Simulation mit beispielsweise 10 Systemen und 20 Bauteilen pro System

Die Anteile der betroffenen Systeme in der Grundgesamtheit wurden mit Früh-, Alterungs- und Zufallsausfällen von jeweils 0% bis 100% der betroffenen Grundgesamtheit variiert. Wenn z.B. 10% der Grundgesamtheit von Früh- oder Chargenausfällen und 60% von Alterungsausfällen betroffen waren, wurden die übrigen 30% der Grundgesamtheit mit Zufallsausfällen belegt. Wenn die Kombination von Früh- als auch Alterungsausfällen mehr als 100% der Grundge-

samtheit ausmachten (d.h. bei einem System sind sowohl Früh- als auch Alterungsausfälle potentiell möglich) wurden keine Zufallsausfälle modelliert. In der Simulation werden die Fehlermechanismen äquivalent zu Bauteilen gesehen, d.h. jedes Bauteil repräsentiert einen bestimmten Fehlermechanismus. Die Simulation wurde mit 10, 100 und 1.000 Fehlermechanismen (Bauteilen) pro System durchgeführt. Zur Modellierung der Wahrscheinlichkeitsverteilungen der einzelnen Fehlermechanismen wurde die 2-parametrische Weibullverteilung verwendet mit der Ausfallwahrscheinlichkeit $F(t)$ bis zum Zeitpunkt t einen Ausfall zu generieren:

$$F(t) = 1 - e^{-\left(\frac{t}{\eta}\right)^\beta} \quad (1)$$

Der Parameter β kann durch dessen Ausprägung wie oben beschrieben Früh-, Zufall- oder Alterungsausfälle repräsentieren. Der Parameter η ist der Lageparameter und wird auch als charakteristische Lebensdauer bezeichnet. Für die Simulation wurden sowohl die β - als auch die η -Werte für die Früh-, Zufalls- als auch Alterungsausfälle anhand einer Wertspanne zufällig simuliert.

Die nachfolgende Tabelle 1 gibt die Wertespanne für die simulierten Weibullverteilungen an.

Tabelle 1: Wertespanne für die Parameter der simulierten Weibullverteilungen

Art des Ausfalls	β	η [in Monaten]
Frühausfall	0,5 – 0,9	60 – 700
Zufallsausfall	1,0	300.000 – 370.000
Alterungsausfall	1,2 – 2,5	180 – 2.000

Aus diesen Wertespannen wurden nun 10, 100 und 1.000 Weibullverteilungen mit einem festen Anteil von 5% an Früh-, 35% an Zufalls- und 60% an Alterungsausfallverteilungen simuliert. So wurden bei einem System bestehend aus 100 Bauteilen (Fehlermechanismen), 5 Verteilungen die ein Frühausfallverhalten, 60 Verteilungen die ein Alterungsausfallverhalten und 40 Verteilungen die ein Zufallsausfallverhalten der Bauteile repräsentieren, modelliert.

Innerhalb der Teilpopulationen, wo diese Fehlermechanismen vorkommen, wirken diese konkurrierend um den Ausfallzeitpunkt.

Zur Erzeugung der Ausfallzeitpunkte wurde eine Monte-Carlo-Simulation [10] durchgeführt. Dazu wurde für jedes zu simulierende System aus den jeweils dort wirkenden Weibullvertei-

lungen ein zufälliger Ausfallzeitpunkt generiert. Der Ausfallzeitpunkt t_i des Systems i wird dann durch den minimalen Ausfallzeitpunkt von den erzeugten Ausfallzeitpunkten $(t_{i1}, t_{i2}, \dots, t_{ir})$ aus r Weibullverteilungen bestimmt:

$$t_i = \min\{t_{i1}, t_{i2}, \dots, t_{ir}\} \quad (2)$$

In der Simulation wurden insgesamt 5.000 Systeme pro Grundgesamtheit betrachtet und mit dem oben beschriebenen Vorgehen dessen Ausfallzeitpunkte erzeugt. Zur Auswertung der Ausfalldaten, wurden diese einmal nach 6, 12, 18 und 24 MIS (Months in Service) zensiert. Alle Systeme, die einen früheren Ausfallzeitpunkt besitzen als der Zensierungszeitpunkt wurden als ausgefallen klassifiziert und alle anderen Systeme als überlebend mit einer bisherigen Lebensdauer, die gleich dem Zensierungszeitpunkt ist. Die so entstandenen Lebensdauerdaten wurden durch eine Weibullverteilung approximiert. Dessen Parameter wurde durch die Maximum-Likelihood-Methode geschätzt sowie die dazugehörigen Konfidenzintervalle der Parameter unter Verwendung der Varianz-Kovarianz-Matrix über die normal approximierten Konfidenzintervalle bestimmt (Details zum Berechnungsverfahren siehe [11]). Diese Simulation wurde dann für sämtliche Kombinationen an unterschiedlichen Anteilen der Grundgesamtheit von 5.000 Systemen, die von Früh-, Zufalls- und Alterungsausfällen betroffen waren mit jeweils 10, 100 und 1000 Fehlermechanismen wiederholt und zu den Auswertungszeitpunkten 6, 12, 18 und 24 MIS analysiert. Als Ergebnis entstand eine Übersicht, bei welcher Kombination von Anteilen der betroffenen Grundgesamtheit mit Früh-, Zufalls- und Alterungsausfällen mit welchen Zensierungszeitpunkten und mit welcher Anzahl an wirkenden Fehlermechanismen bei jeweils einer Grundgesamtheit von 5000 Systemen ein signifikantes $\beta \neq 1$ zu sehen ist. Die komplette Simulation wurde mit der Skriptsprache R [12] umgesetzt. Insgesamt ergaben sich 10.000 Kombinationen pro MIS-Bereich und Anzahl von Bauteilen, was 120.000 Kombinationen insgesamt entspricht. Die Ergebnisse der Simulationsstudie sind in Abbildung 6 dargestellt.

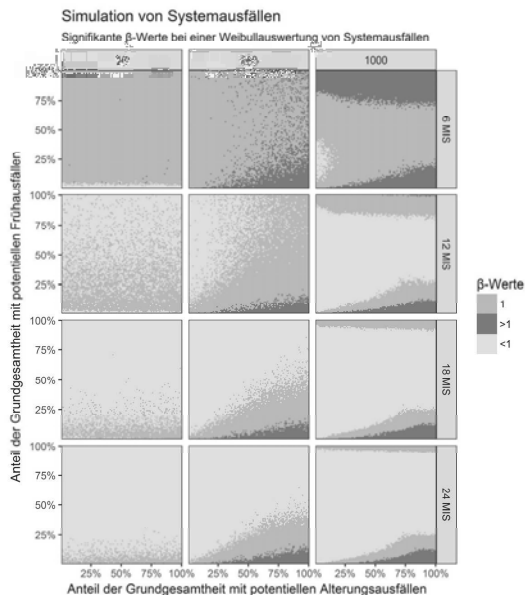


Bild 6: Simulationsergebnis für eine Weibullauswertung bei Systemen

In jedem der Diagramme in Abbildung 6 ist auf der x-Achse der Anteil der betroffenen Grundgesamtheit mit Alterungsausfällen und auf der y-Achse der Anteil der Grundgesamtheit mit Frühausfällen aufgetragen. Somit ergibt z.B. die Kombination 60% auf der x-Achse und 25% auf der y-Achse, dass in dieser Grundgesamtheit (hier immer 5000 Systeme) insgesamt 60% der System also 3000 Stück von Alterungs- und 1250 von Frühausfällen potentiell betroffen sind. Der Punkt im Diagramm bei dieser Kombination spiegelt den β -Wert wieder, den man bei einer Weibullauswertung mit der entsprechenden Anzahl an Fehlermechanismen im System und zu den entsprechenden Auswertezeitpunkten in der Simulation erhalten hat. Die β -Werte wurden hier in drei Klassen eingeteilt (signifikant kleiner 1, signifikant größer 1 sowie 1) und entsprechend farblich markiert.

Als Ergebnis der Simulationsstudie lässt sich ableiten, dass eine Mischung von wenigen Systemen der Grundgesamtheit mit Frühausfallverhalten und ein mehr oder weniger beliebiger Anteil der Grundgesamtheit mit Alterungsproblemen ein $\beta \approx 1$ unter bestimmten Umstände ergeben können.

Ein höherer Anteil an Frühausfällen führt zu $\beta < 1$ unabhängig des betroffenen Anteils der Grundgesamtheit mit Alterungsmechanismen. Zu einem frühen Auswertungszeitpunkt (6 MIS) sind die Ergebnisse der Weibullauswertung noch sehr gestreut, wohingegen bei der Zunahme von Ausfalldaten (12-24 MIS) und Bauteilen pro System sich deutliche Muster ergeben. Bei einer höheren Anzahl an Fehlermechanismen im System (100,1.000) spiegelt sich die Dominanz von Alterungsmechanismen bei einem sehr geringen Anteil von Frühausfällen wieder und ergeben ein $\beta > 1$. Sobald aber der Anteil an Frühausfällen zunimmt konvergiert der β -Wert gegen 1. Bei einer zu hohen Dominanz von Frühausfällen ergibt sich ein klares $\beta < 1$. Je höher die Anzahl der Bauteile im System ist, umso schneller zeigt sich diese Dominanz.

5. Fazit

Die Weibullauswertungen des Systemausfallverhaltens zeigt bei der Zunahme von Bauteilen pro System ein immer klareres Bild, wann sich ein $\beta \approx 1$ für eine Weibullauswertung auf Systemebene ergibt. Betrachtet man dies zudem über MIS, zeigt sich hier ein ähnliches Verhalten. Erkennbar ist der Trend auch schon bei einer geringeren Anzahl von Bauteilen und bei geringeren MIS: Ein $\beta \approx 1$ auf Systemebene kommt zustande, wenn sowohl Anteile der Grundgesamtheit von Früh- als auch von Alterungsausfällen betroffen sind und diese jeweils nicht zu dominant sind. Dabei ist das konkrete Verhältnis der Anteile untereinander abhängig von der Anzahl der Bauteile pro System sowie dem Auswertungszeitpunkt. Unabhängig von dem was in der Weibullauswertung auf Systemebene ersichtlich wird, verbergen sich fast immer unterschiedliche Fehlermechanismen auf der Bauteilebene. Entsprechend muss eine Weibullauswertung nicht auf System-, sondern immer auf Bauteilebene erfolgen, um das „wahre“ Ausfallverhalten der Systeme ermitteln zu können. Dies zeigt auch das hier vorgestellte Beispiel. Als weitergehende Untersuchungen könnte noch ein komplexeres Simulationsmodell aufgebaut werden, was die Anteile der Fehlermechanismen pro Bauteil sowie deren Verhältnisse von Früh- Zufalls- und Alterungsproblemen variieren lässt. So könnte erforscht werden, ob sich auch unter diesen Bedingungen ein ähnliches Verhalten/Bild ergibt.

6. Literaturangaben

- [1] VDA Rotband 3.2: Zuverlässigkeits-Methoden und –Hilfsmittel, Frankfurt am Main: Henrich Druck + Medien GmbH 2016
- [2] Bertsche, B.; Lechner, G.: Zuverlässigkeit im Fahrzeug- und Maschinenbau, Ermittlung von Bauteil- und System-Zuverlässigkeiten, Berlin/Heidelberg: Springer-Verlag 2004
- [3] Meyna, A.; Pauli, B.: Taschenbuch der Zuverlässigkeitstechnik, Quantitative Bewertungsverfahren, München: Hanser Verlag 2010
- [4] Meyer, M.: Methoden zur Analyse von Garantiedaten für die Sicherheits- und Zuverlässigkeitsprognose von Komponenten und Baugruppen im Kraftfahrzeug, Dissertation, Bergischen Universität Wuppertal 2003
- [5] Abernethy, R.: The new Weibull handbook, Florida, USA: Abernethy, 2006
- [6] Tarum, C.: Classification and Analysis of Weibull Mixtures, SAE Technical Paper, 1999
- [7] Klein, B.: Numerische Analyse von gemischten Ausfallverteilungen in der Zuverlässigkeitstechnik, Dissertation, Stuttgart 2013
- [8] Hastie, T.; Tibshirani, R.; Friedman, J.: The elements of statistical learning, New York: Springer Verlag 2016
- [9] MIL-HDBK-217F: Military Handbook, Reliability prediction of electronic equipment, Department of Defense 1991
- [10] Robert, C.; Casella, G.: Introducing Monte Carlo Methods with R, Berlin/Heidelberg: Springer-Verlag 2010
- [11] Meeker, W.; Escobar, L.: Statistical Methods for Reliability Data, New York, USA: John Wiley & Sons 1998
- [12] Tippmann, S.: Programming tools: Adventures with R. In: Nature. Band 517, Nr. 1, 29. Dezember 2014, S. 109–110

Entwicklung verlässlicher, intelligenter Systeme

Herausforderungen und Chancen



Prof. Dr.-Ing. **W. Sextro**, **J. K. Kimotho**, M.Sc.,
T. Kaul, M.Sc., Dipl.-Ing. **T. Meyer**,
Lehrstuhl für Mechatronik und Dynamik, Universität Paderborn

Kurzfassung

Intelligente technische Systeme sind durch einen erhöhten Funktionsumfang charakterisiert, der diese dazu befähigt, autonom auf wechselnde Umgebungsbedingungen, Anforderungen und inhärente Systemzustände zu reagieren. Dies kann mit den Methoden der Selbstoptimierung erreicht werden. Hier werden mit Verfahren der Mehrzieloptimierung mögliche Betriebspunkte des Systems bestimmt zwischen denen das System im Betrieb autonom auswählt und somit eine Verhaltensadaption erwirkt. Zur Berechnung der Betriebspunkte ist es notwendig ein Modell des Systemverhaltens aufzustellen und das Verhalten hinsichtlich verschiedener, meist konfliktärer, Ziele zu quantifizieren.

Bei der Modellierung des Systemverhaltens und der Formulierung der Ziele stellt die Absicherung der Verlässlichkeit auf Grund der zunehmenden Systemkomplexität eine große Herausforderung dar, der im Entwicklungsprozess begegnet werden muss. Die Implementierung von Selbstoptimierung bietet darüber hinaus in Kombination mit einer Zustandsüberwachung im Betrieb die Möglichkeit einer zuverlässigkeitsbasierten Verhaltensanpassung, deren Potential zu einer Steigerung der Verlässlichkeit genutzt werden kann.

In dieser Arbeit werden die Entwicklung intelligenter technischer Systeme und die damit verbundenen notwendigen Entwicklungsschritte zur Absicherung der Verlässlichkeit anhand von selbstoptimierenden Systemen betrachtet. Dazu gehören die Formulierung verlässlichkeitsrelevanter Ziele und die Implementierung einer Zustandsüberwachung als Basis für eine zuverlässigkeitsbasierte Verhaltensanpassung. Es werden auf Grundlage einer Beschreibung der Entwicklungsschritte, Potentiale zur Steigerung der Verlässlichkeit sowie Chancen und zukünftige Herausforderungen herausgestellt und diskutiert.

Intelligente technische Systeme

Moderne maschinenbauliche Erzeugnisse sind charakterisiert durch hohe Anforderungen an ihre Leistungsfähigkeit und ihre Verlässlichkeit. Die Verlässlichkeit umfasst dabei unter anderem die Aspekte Zuverlässigkeit, Sicherheit und Verfügbarkeit (1). Um die Leistungsfähigkeit zu erhöhen, wird bereits seit geraumer Zeit die klassische mechanische Struktur durch Aktoren, Sensoren und Informationsverarbeitung ergänzt – diese Systeme werden als mechatronische Systeme bezeichnet. Durch den Einsatz leistungsfähiger Informationsverarbeitung als digitale Regler wird ein gezieltes Beeinflussen der Systemdynamik möglich. Darauf aufbauend haben sich adaptive Systeme ergeben, die sich durch eine Anpassung der Dynamik zur Laufzeit auszeichnen.

Durch die zunehmende Integration von Kommunikationsfähigkeiten, von Sensoren zur Aufnahme der Umgebung von Produkten und durch komplexe Datenauswerteverfahren wird mittlerweile über die reine Adaption des Systemverhaltens hinaus eine intelligente, selbstständige Anpassung des Systems an veränderte Anforderungen, veränderte Situationen und den Systemzustand selbst möglich. Eine Variante, diese Adaptionsintelligenz zu implementieren besteht aus der Nutzung von Selbstoptimierung.

Problematisch bei jeglichen Formen der Erweiterung der Informationsverarbeitung ist die steigende Komplexität, die nicht nur durch die Software, sondern auch und insbesondere durch Ersetzung mechanischer Funktionalitäten durch mechatronische Komponenten steigt. Ein Beispiel dafür sind diverse X-by-wire-Systeme im Kraftfahrzeug, die dazu geführt haben, dass die Zuverlässigkeit moderner KFZ maßgeblich durch die zur einwandfreien Funktion notwendige Elektronik beeinflusst wird.

Nachdem in den vergangenen Jahrzehnten einzelne Komponenten oder Systeme innerhalb der Produkte durch mechatronische Systeme ersetzt oder erweitert wurden, ist die aktuelle Herausforderung vernetzte Systeme sicher und zuverlässig interagieren zu lassen. Diese Kombination ermöglicht neuartige Ansätze: Daten können nicht mehr nur innerhalb eines Systems verarbeitet und genutzt werden, sondern werden während des Betriebs dem Nutzer, einer Prozessleitstelle oder dem Hersteller zur Verfügung gestellt. Auf Basis dieser Daten lassen sich moderne Condition Monitoring-Verfahren anwenden, um die Zuverlässigkeit und Verfügbarkeit einzelner Systeme zu steigern.

Dazu wird die Adaptionsfähigkeit nicht mehr nur zur funktionalen Adaption der Systeme genutzt, sondern auch zu einer zuverlässigkeitsbasierten Adaption. Diese ermöglicht eine Anpassung des Systems an die aktuelle Zuverlässigkeit. Die Zuverlässigkeitsanforderungen, die konfliktär zu leistungsorientierten Zielen sein können, werden durch eine Priorisierung dieser Zuverlässigkeitsziele sichergestellt. Da diese Anpassung rein auf Software-Ebene ge-

schießt und keine weiteren Aktoren im System benötigt, ist die zusätzlich entstehende Komplexität im Rahmen der Softwareentwicklung beherrschbar. Zur Entwicklung einer solchen zuverlässigkeitsbasierten Verhaltensanpassung sind neben der reinen Erkennung des Zustands weitere Schritte notwendig, die den Entwicklungsprozess erweitern.

Entwicklungsprozess selbstoptimierender technischer Systeme

Eine wesentliche Grundlage zur Entwicklung selbstoptimierender Systeme ist die Formulierung von Zielen des Systems. Mittels Mehrzieloptimierungsverfahren werden auf Basis der vorab definierten Ziele mögliche Betriebspunkte berechnet, zwischen denen das System während des Betriebs autonom wählen kann. Zur Nutzung von Mehrzieloptimierung ist allerdings neben der Formulierung der Ziele auch eine exakte Quantifizierung der Zielerreichung notwendig, die nur auf Basis eines Modells des Systemverhaltens erfolgen kann. Das Modell muss dabei den Zusammenhang zwischen allen zur Laufzeit veränderlichen Parametern und allen relevanten Zielen abdecken, was einen ganzheitlichen Modellierungsansatz erfordert. Es reicht nicht mehr, Teilmodelle einzelner Domänen losgelöst voneinander zu betrachten, sondern das Systemmodell muss alle Domänen und eventuell auch sekundäre Teilaspekte beinhalten, die bei einem rein zum Reglerentwurf oder zur Auslegung der mechanischen Struktur aufgestellten Modells nicht notwendig wären. Dies führt zu einer starken Durchmischung der einzelnen Domänen Maschinenbau, Elektrotechnik, Regelungstechnik und Informatik innerhalb des Entwicklungsprozesses. Dabei ergeben sich insbesondere durch die zwangsläufig notwendigen Abstimmungen zwischen den Mitarbeitern der einzelnen Domänen große Risiken in Bezug auf die Modellkonsistenz, die in allererster Linie die Verlässlichkeit des gesamten Systems gefährden (1–3).

Dieser Prozess erfordert die Unterstützung durch geeignete Werkzeuge zur Sicherstellung synchroner Modellierungen innerhalb der einzelnen Domänen. Ein für das Systemverhalten letztendlich maßgebliches Bindeglied stellt dabei das innerhalb der Domäne Regelungstechnik aufgestellte Modell des Systemverhaltens dar, das die dynamischen Eigenschaften mechanischer Komponenten (Elastizitäten, Massen,...), elektronische Teilsysteme (Energieversorgung, Filter, Sensoren,...) mit den für das Verhalten entscheidenden Regelgesetzen und ihren Parametern verknüpft, die dann zur Erzeugung von Steuergerätecode dienen. Durch die Festlegung dieser Reglerparameter wird auch über die Lebensdauer physischer Komponenten des Systems entschieden. Denn hohe Reglerdynamik führt meist zu hohen Geschwindigkeiten, Stellkräften und Strömen, zu niedriger Reglerdynamik kann dagegen ungewolltes Systemverhalten wie starke Überschwinger begünstigen. Somit kann eine nicht optimale Reglerdynamik die zugehörigen Komponenten stark belasten und zu einer kürzeren

Lebensdauer führen. Die genutzte Reglerdynamik ist dadurch eine maßgebliche Stellgröße, die zwischen dynamischen Systemverhalten und seiner Zuverlässigkeit ausgewählt werden kann. Während der Entwicklung selbstoptimierender Systeme ist daher eine selbstständige Auswahl von Reglerparameter während des Betriebs auf Basis der vorab formulierten Ziele der Schlüssel zur Implementierung von Verhaltensadaption. Die zur Berechnung möglicher Betriebspunkte notwendige Quantifizierung der Zielwerte bewertet dabei das (dynamische) Systemverhalten, muss aber zugleich auch die gegensätzliche Zuverlässigkeit bewerten. Für die Abbildung beider Aspekte ist eine Integration der Zuverlässigkeitsmodellierung in das Verhaltensmodell geeignet.

Integrierte Modellierung zur Absicherung von Verlässlichkeits- und Verhaltensanforderungen

Die Absicherung von Verlässlichkeits- und Verhaltensanforderungen in selbstoptimierenden Systemen erfordert eine geeignete Quantifizierung der gewählten Ziele. Zur Abbildung der Verhaltensanforderungen (Effizienz, Betriebskosten, Performance u. a.) eignen sich beispielsweise Kostenfunktionale, die das Verhalten direkt bezüglich der Ziele bewerten. Die Abbildung der Verlässlichkeit erfordert dem gegenüber eine eigenständige Modellierung, die nicht kongruent zu dem Modell des Systemverhaltens ist. Die Herausforderung ist die Modellierung des Zusammenhangs zwischen Systemverhalten, Degradation der Systemelemente und Zuverlässigkeit des Gesamtsystems. Diese Aspekte, Verhalten und Zuverlässigkeit von Komponenten und des Gesamtsystems, weisen eine starke Abhängigkeit auf. So verändert der aktuell gewählte Betriebspunkt das (dynamische) Systemverhalten und damit auch die wirkenden Lasten im System selbst. Die Degradation der Systemkomponenten hängt im hohen Maße von diesen Belastungen ab und ist maßgeblich für die Zuverlässigkeit des gesamten Systems.

Das bisherige Vorgehen zur Modellierung der Verlässlichkeit ist eine manuelle, auf Expertenwissen basierende Aufgabe, die eine allgemeine geschlossene Betrachtung des Systemverhaltens und der Verlässlichkeit im Rahmen einer Mehrzieloptimierung ausschließt (vgl. Bild 1 oben). In jeder Optimierungsiteration ist eine Synchronisation zwischen Verhaltens- und Verlässlichkeitsmodell erforderlich, die manuell nicht umsetzbar ist. Es ist daher eine automatisierte Ableitung eines Verlässlichkeitsmodells aus dem Modell des Systemverhaltens umzusetzen, um eine geschlossene Modellierung zu erreichen. So ist es möglich, Verlässlichkeits- und Verhaltensanforderungen gleichermaßen in der Mehrzieloptimierung zu betrachten (vgl. Bild 1 unten).

Um die Auswirkung der sich in Abhängigkeit des gewählten Betriebspunkts ändernden Belastungen auf die Degradation der einzelnen Komponenten abzubilden, werden Lebensdauerschätzer an relevante Komponenten annotiert und bilden zusammen mit dem Verhaltensmodell das vollständige Systemmodell (Bild 2). Dieses Systemmodell umfasst Verhaltens- sowie Verlässlichkeitsaspekte, hier die Zuverlässigkeit, und wird jeweils für die aktuell gewählten Optimierungsparameter simuliert. Als Eingangsgrößen müssen Soll-Trajektorien für das System gewählt werden, die charakteristische Betriebszustände des Systems abdecken. Dies ist notwendig, um das System für den Betrieb und die zu erwartenden Umgebungs- und Einsatzbedingungen optimieren zu können.

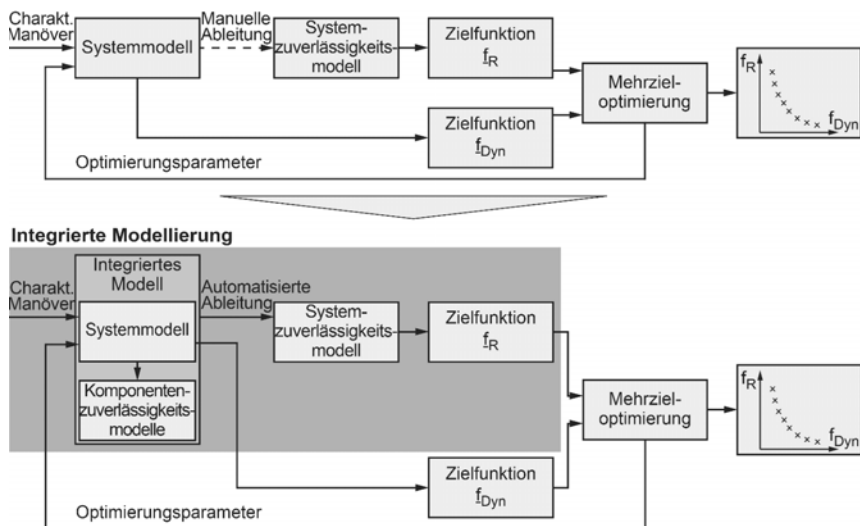


Bild 1: Vergleich der Vorgehen zur Berücksichtigung verlässlichkeitsrelevanter Ziele für die Mehrzieloptimierung (nach (4)). Oben: bisheriges Vorgehen. Unten: Integrierte Modellierung.

Eine automatisierte Ableitung eines Verlässlichkeitsmodells aus dem Systemmodell wird über eine geeignete Transformation ermöglicht (4, 5). Auf Basis des Verlässlichkeitsmodells ist nun die Formulierung verlässlichkeitsrelevanter Zielfunktionen, die Quantifizierung des Systemverhaltens hinsichtlich geforderter Zuverlässigkeit, Verfügbarkeit und Sicherheit, möglich.

Indem eine geforderte Zuverlässigkeit zu einer bestimmten Nutzungsdauer schon frühzeitig im Entwicklungsprozess berücksichtigt wird, ist es möglich die Obsoleszenz eines technischen Systems gezielt einzustellen. Um die geforderte Nutzungsdauer auch während des Betriebs noch zu beeinflussen ist eine Anpassung des Systemverhaltens während des Betriebs notwendig.

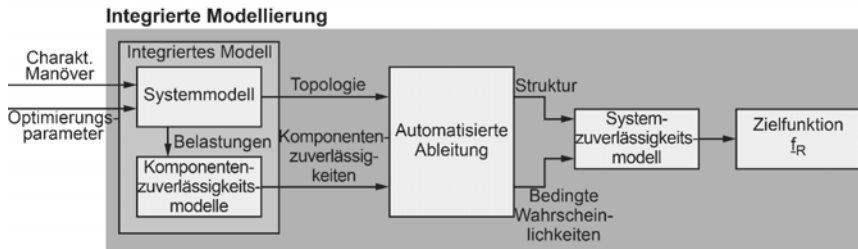


Bild 2: Modellstruktur der Integrierten Modellierung (nach (5))

Verhaltensanpassung

Um die für die Herstellung eines technischen Systems aufgewendeten Ressourcen gut auszunutzen, ist eine möglichst lange Nutzung anzustreben. Dies wird nur dann ermöglicht, wenn Anforderungen der Nutzer erfüllt sind und wenn das System lebensdauergerecht belastet wird. Die erreichbare Lebensdauer befindet sich dabei meist im Konflikt mit den Anforderungen des Nutzers an das Systemverhalten: Starke Nutzung, beispielsweise die ständige Nutzung in einem Bereich, in dem hochdynamische Regler mit großen Stellkräften notwendig sind, beansprucht das System stark und schränkt die Lebensdauer ein.

Die unterschiedliche Belastung ähnlicher Systeme mit resultierend stark unterschiedlichen Lebensdauern erfordert aufwendige Wartungsplanung. Nutzung von Condition Monitoring-Verfahren ermöglicht dabei eine Vorhersage der verbleibenden Nutzungsdauer, die als Basis einer zustandsbasierten Instandhaltung bzw. vorausschauender Wartung genutzt werden kann. Problematisch dabei ist, dass optimale Ressourcennutzung möglichst späte Wartung kurz vor dem tatsächlichen Schadenseintritt erfordert, was die Wartungsplanung zusätzlich kompliziert.

Um dies zu erleichtern, kann die Verhaltensanpassungsfähigkeit selbstoptimierender Systeme zur Sicherstellung einer vorab definierten Lebensdauer genutzt werden. Diese schränkt dabei die anderen Ziele des Systems so weit ein, dass die gewünschte Lebensdauer möglichst gut erreicht wird – zugleich aber nur so weit, dass eine bestmögliche Leistungsfähigkeit erhalten bleibt. Als Grundlage dient dabei die Formulierung von mindestens zwei Zielen: Eines muss die Zuverlässigkeit des Systems abbilden, das andere seine Leistungsfähigkeit.

Hat ein System mehrere kritische Komponenten, die eventuell sogar selbst konfliktär sind, sind weitere Ziele notwendige. Dies ist der Fall bei allen Systemen mit Redundanz, da dabei mehrere Komponenten oder Teilsysteme die gleiche Teilfunktion erbringen, aber unterschiedliche stark geschädigt werden. Ebenso muss bei mehreren Leistungsfähigkeitszielen die Menge aller Ziele erweitert werden (6).

Für den Fall je eines Zieles findet während des Betriebs eine Abwägung der beiden Ziele gegeneinander statt. Dazu wurden Regelungsverfahren entwickelt, die auf einem Health Index basieren (7). Er ist ein Maß für die Fähigkeit des Systems, weitere Schädigung zu ertragen. Bei einem neuen System ist der Health Index 1, bei einem verbrauchten System zum Zeitpunkt des Schadenseintritts 0. Der gewünschte Health Index wird dabei als zeitveränderlicher Sollwert vorgegeben. Typischerweise ist er streng monoton fallend, Ausnahmen bestehen allerdings für Systeme mit Selbstheilung (beispielsweise Brennstoffzellen, in denen sich die Membran während Ruhezeiten chemisch selbst regeneriert). Die Definition des Zeitverlaufs basiert dabei auf einer Zuverlässigkeits-Zeit, die nicht mit der tatsächlichen Zeit korrelieren muss, und beispielsweise bei Luftfahrzeugen auf Flugstunden oder bei zyklisch belasteten Komponenten auf der Anzahl Betriebszyklen basieren kann. Während des Betriebs wird der aktuelle Health Index des Systems bestimmt und als Istwert genutzt. Das selbstoptimierende System erbringt dann die Funktionalität eines Reglers, der Soll- und Istwert vergleicht und das Verhalten so vorgibt, dass der gewünschte Health Index erreicht wird (6).

Die Verwendung des Health Index ist für Ausfallarten basierend auf einem Degradationsverhalten geeignet. Um Zufallsausfällen, also Ausfallarten ohne erkennbarer Degradation, zu begegnen, wurde ein mehrstufiges Verlässlichkeitskonzept entwickelt. Hier werden online Diagnoseverfahren genutzt, um das Systemverhalten und Umfeldbedingungen zu analysieren und mögliche Komponentenausfälle zu identifizieren. Um aufgetretenen Fehlern zur Laufzeit zu begegnen, wird unter Ausnutzung der Adaptionfähigkeit das Systemverhalten, beispielsweise mit Hilfe von Rekonfiguration, angepasst (8).

Condition Monitoring zur Erkennung des aktuellen Zustands

Condition Monitoring oder die Zustandsüberwachung eines technischen Systems beinhaltet die kontinuierliche oder regelmäßige Datensammlung aus einem Netz von Sensoren oder Betriebsdaten, um den Zustand eines Systems während des Betriebs zu schätzen. Condition Monitoring kann entweder im Aktiv- oder Passiv-Modus durchgeführt werden. Im Aktiv-Modus wird das System mit einer definierten Anregung periodisch angeregt. Dieser Ansatz wird für statische, quasi-statische Systeme oder Systeme mit stochastischer Anregung verwendet, da keine definierte, kontinuierliche Systemantwort gemessen werden kann. Beispie-

le für solche Systeme sind Strukturen wie Windkraftanlagen, Brücken, Flugzeugrümpfe und Kräne. Auf der anderen Seite weisen die meisten dynamischen Systeme eine gut definierte Antwort auf, die während des Betriebs gemessen werden können. Die gewonnenen Daten werden verarbeitet, um Zustandskennzahlen, die verschiedenen Zuständen des Systems (Health States) zugeordnet sind, zu identifizieren. Diese können nachfolgend als Eingangssignal zur Wartungsentscheidung oder für zuverlässigkeitsadaptive Systeme verwendet werden. Dadurch kann die Zuverlässigkeit, Verfügbarkeit und Sicherheit des Systems erhöht werden. Typische Condition Monitoring-Ansätze auf Basis verfügbarer Daten und Systeminformationen werden im Folgenden erörtert (10).

1) Modellbasiertes Condition Monitoring

Bei dieser Methode wird ein Systemmodell genutzt, das die Degradation abbildet. Es basiert auf einer vollständigen Beschreibung des Systems und der Degradationsvorgänge. Die Variablen des Systemmodells, die mit Fehlerarten korrelieren, können dann durch Simulation des Systemmodells unter Berücksichtigung der Nutzung und der Umgebung, evaluiert werden. Somit wird eine Zustandserkennung möglich. Bild 3 zeigt die Elemente eines modellbasierten Ansatzes zur Zustandserkennung. Modellbasierte Verfahren können leicht in zuverlässigkeitsadaptive Systeme integriert werden und sollten die erste Wahl sein, wenn ein Systemmodell zur Verfügung steht. Allerdings erfordert die Modellentwicklung ein umfassendes Verständnis des Systems, wobei große, komplexe Systeme eine Herausforderung darstellen. Die Simulation komplexer Systemmodelle kann zudem sehr rechenintensiv sein, sodass die umsetzbare Modellkomplexität abhängig ist von der zur Verfügung stehenden Rechenleistung. Auf Grund einer komplexitätsbedingt hohen Simulationsdauer ist dieser Ansatz nicht für Systeme mit sehr kurzen Lebensdauern geeignet.

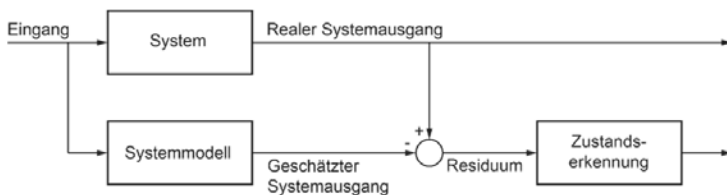


Bild 3: Modellbasierte Zustandserkennung

2) Datengetriebenes Condition Monitoring mittels maschineller Lernverfahren

Datengetriebene Methoden verwenden Maschinenlernverfahren, um Zustandsüberwachungsdaten auf eine Zustandskennzahl, die mit der Degradation des Systems korreliert, abzubilden. Diese Modelle basieren ausschließlich auf Messdaten, wodurch keine tiefgehende Kenntnis des Systems notwendig ist (9). Maschinenlernverfahren, wie neuronale Netze, Support Vector Machines (SVM), Klassifikations- und Regressionsbäume, Random Forests u. ä., können verwendet werden. Die Verfahren werden zunächst offline zum Anlernen eines Modells genutzt. Dabei wird ein Modell erstellt, das Zustandsüberwachungsdaten auf ein Ziel (Zustandskenn (11)zahl) abbildet (Bild 4). Dazu sind sowohl Zustandsüberwachungsdaten als auch Zustandskennzahl als Trainingsdaten notwendig. Das Modell kann dann online mit Zustandsüberwachungsdaten eines ähnlichen Systems verwendet werden, um den aktuellen Zustand in Echtzeit zu schätzen (12).

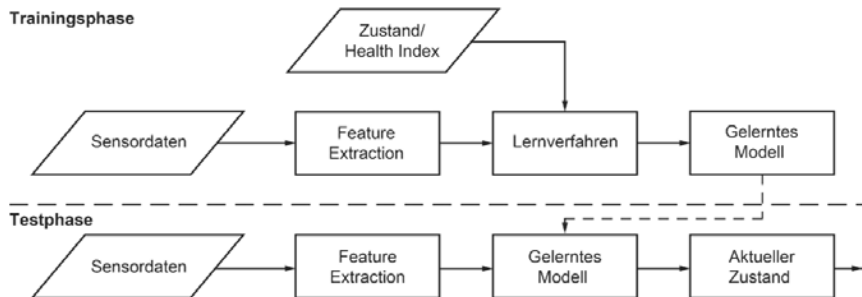


Bild 4: Allgemein Nutzung von maschinellem Lernverfahren zur Zustandsüberwachung

3) Hybrides Condition Monitoring

Um die Genauigkeit der Zustandserkennung zu erhöhen, kann ein Hybridansatz mit modellbasierten und datengetriebenen Modellen verwendet werden. Die Kombination kann durch unterschiedliche Methoden, sowie gewichtete Mehrheitsentscheidung durchgeführt werden, wobei jeweils die Vorteile der verwendeten Modelle genutzt werden (13).

Zukünftige Herausforderungen

Selbstoptimierung ist eine vielversprechende Möglichkeit zur Steigerung der Verlässlichkeit mechatronischer Systeme, die sich bereits bei Laborversuchen bewährt hat (7, 14) und die Anwendungsreife erlangt hat. Dies wurde unter anderem durch leistungsfähige Condition Monitoring-Systeme ermöglicht. Durch gezielte Softwareunterstützung des Entwicklungsprozesses können fehleranfällige manuelle Synchronisierungsprozesse reduziert werden (4, 5).

Dennoch bleibt die Komplexität selbstoptimierender Systeme hoch und im zugehörigen Entwicklungsprozess müssen Aspekte zahlreicher Domänen zugleich berücksichtigt werden.

Eine besondere Herausforderung stellt dabei die Nutzung der Mehrzieloptimierung zur Berechnung möglicher Betriebspunkte dar. Die dabei notwendigen Zielfunktionen müssen auf Basis eines Modells des Systemverhaltens quantifiziert werden. Dies bedingt einerseits eine detaillierte Modellierung des Systemverhaltens, andererseits, bedingt durch zahlreiche Auswertungen und Simulationen durch den Optimierungsalgorithmus, schnelle Rechenzeit. Dieser Konflikt wird durch die Art der genutzten Modelle erschwert: Das (dynamische) Verhalten des Systems wird typischerweise mit spezialisierten Programmen modelliert und simuliert, beispielsweise als Mehrkörperdynamikmodell. Diese Modelle können zwar in einer Mehrzieloptimierung genutzt werden, bedingen durch die eingeschränkte Verfügbarkeit des Gradienten der Zielfunktionen jedoch noch einmal mehr Auswertungen und höhere Rechenzeit während der Optimierung.

Nach dem aktuellen Stand der Technik, ist es nicht möglich, die Gültigkeit eines Condition Monitoring-Systems ohne ein bestehendes Anwendungssystem zu gewährleisten, da Trainingsdaten benötigt werden. Um diese Abhängigkeit bei der Entwicklung eines Condition Monitoring-Systems in der Entwicklungsphase eines Anwendungssystems zu beseitigen, sollte ein modellbasierter Ansatz gleichzeitig mit dem Entwurf des Anwendungssystems entwickelt werden. Mit dem Systemmodell können Trainingsdaten zur Auswertung des Condition Monitoring-Systems generiert werden. Alle möglichen Fehlerarten sollten berücksichtigt und in das Systemmodell aufgenommen werden, um die Zustandserkennung zu ermöglichen (15).

Um Verlässlichkeitssteigerung im Betrieb durch autonome Verhaltensanpassung weiter zu verbreiten ist eine weitere Herabsetzung der komplexitätsbedingten Hürden bei der Implementierung von Condition Monitoring-Systemen und insbesondere bei der Auswahl geeigneter Messgrößen und Algorithmen durch softwareunterstützte Methoden notwendig.

Chancen

Erst die Verfügbarkeit kleiner, leistungsfähiger Recheneinheiten ermöglichte den Übergang von der Mechanik zur Mechatronik. Der nächste Schritt hin zu vernetzten, adaptiven und letztlich intelligenten Systemen wurde auch erst durch die weitere Miniaturisierung ermöglicht. Mit dem Anhalten des Trends zu immer leistungsfähigeren, kleinen, energiesparenden Prozessoren mit mehreren Kernen steigt auch die Verfügbarkeit von Rechenleistung. Während vor kurzem noch ein Supercomputer zur Mehrzieloptimierung mechatronischer Systeme notwendig war, kann dies mittlerweile ein leistungsfähiger Arbeitsplatzrechner durchführen.

In absehbarer Zeit werden diese Berechnungen sogar auf im System ohnehin verfügbarer Hardware, etwa gerade nicht ausgelasteten Steuergeräten, durchgeführt werden können. Dadurch ergibt sich eine neue Stufe der Adaptionfähigkeit: Statt einer reinen Auswahl zwischen vorab berechneten Betriebspunkten ist nun auch eine Neuberechnung möglich. Damit einher geht eine Berücksichtigung neuer Ziele, neuer Umgebungsbedingungen oder die Integration neuer Komponenten.

Die Integration von echtzeitfähigen Netzwerkschnittstellen und der Austausch von Sensordaten oder Sollwerten über diese ermöglicht auch eine Erweiterung des Datenaustauschs um Eigenschaften einzelner Komponenten. So könnte bei einer Wartung eine defekte Komponente gegen eine neue getauscht werden, die ihre veränderten Eigenschaften an die anderen Komponenten weitergibt und eine Neuberechnung aller Betriebspunkte initiiert. Die Nutzung dieser Möglichkeiten würde damit eine ständige Adaption des Systems und einen jederzeit optimalen Betrieb gewährleisten.

Derzeitige Wartungsstrategien sind auf solche Systeme noch nicht vorbereitet. Der aktuelle Stand der Technik geht von einer verschleißgetriebenen Wartungsplanung aus, in der auf das Ausfallverhalten der Systeme nur reaktiv eingegangen wird. Bei klassischer zustandsbasierter Instandhaltung schränkt die Verfügbarkeit von Wartungsteams die real erreichbare Verfügbarkeit stark ein. Soll, was zur bestmöglichen Ressourcennutzung erstrebenswert ist, möglichst spät gewartet werden, steigt die Gefahr der Nicht-Verfügbarkeit von Wartungsteams und eines vermeidbaren Ausfalls vor der Wartung. Zuverlässigkeitsbasierte Verhaltensanpassung auf der anderen Seite ergänzt das Condition Monitoring um den Rückschluss zum Systemverhalten, sodass der Ausfallzeitpunkt vorgegeben werden kann und die Wartung besser planbar wird. Dies in einen Wartungsplan zu integrieren und die entstehende Verfügbarkeit zu berechnen ist ein derzeit noch ungelöstes Problem.

Die Verwendung von Betriebsdaten, in der Regel für die Leistungsüberwachung eingesetzt, kann auch für die Zustandsüberwachung mittels Big Data Analytics verwendet werden. Big Data Analytics untersucht große Mengen an Daten unterschiedlicher Arten, um Muster, Korrelationen und andere Erkenntnisse aufzudecken.

Die Weiterentwicklung und Verbreitung von Sensor- und Informationstechnik, das Internet of Things, Remote-Condition-Monitoring und die Verknüpfung mit Big Data Analytics ermöglichen detaillierteres Systemverständnis und tragen zu einem verlässlicheren Betrieb von autark agierenden technischen Systemen bei.

Zusammenfassung/Ausblick

Die Entwicklung intelligenter, technischer Systeme mit den Methoden der Selbstoptimierung führt zu einer Steigerung der Komplexität dieser Systeme und damit zu einer Gefährdung für die Verlässlichkeit. Der steigenden Komplexität kann mit geeigneten Modellierungswerkzeugen in frühen Phasen des Entwicklungsprozesses begegnet werden. Die systeminhärente Intelligenz, implementiert als Selbstoptimierung, ermöglicht eine zuverlässigkeitsbasierte Verhaltensanpassung während des Betriebs auf Basis des aktuellen Zustands des Systems. Die Umsetzbarkeit intelligenter Systeme ist durch die stark gestiegene verfügbare Rechenleistung in der Informationsverarbeitung ermöglicht worden und führt zu Systemen wie (teil-) autonomen Fahrzeugen und Produktionsanlagen in der Industrie 4.0.

Neben der gestiegenen Rechenleistung ist ein weiterer Schlüsselfaktor die Verbreitung von (vernetzten) Sensoren im Rahmen der Industrie 4.0 und dem Internet of Things, die in Kombination mit Big Data-Methoden ein tiefergehendes Systemverständnis ermöglichen. Durch die weitere Verbreitung dieser zwei Faktoren ist es möglich das Potential intelligenter Systeme stärker nutzbar zu machen während zudem ihre Verlässlichkeit gesteigert werden kann.

Quellen

1. GAUSEMEIER, J., F.-J. RAMMIG und W. SCHÄFER. *Sonderforschungsbereich 614 - Selbstoptimierende Systeme des Maschinenbaus: Verlässlichkeit selbstoptimierender Systeme -Definitionen, Anwendungen, Konzepte*: Heinz Nixdorf Institut, Universität Paderborn, 2009. 234.
2. GAUSEMEIER, J., F.-J. RAMMIG, W. SCHÄFER und W. SEXTRO. *Dependability of Self-optimizing Mechatronic Systems*: Springer-Verlag, 2014.
3. GAUSEMEIER, J., F.-J. RAMMIG und SCHÄFER W. *Design Methodology for Intelligent Technical Systems*: Springer-Verl, 2014.
4. KAUL, T., T. MEYER und W. SEXTRO. Integrierte Modellierung der Dynamik und der Verlässlichkeit komplexer mechatronischer Systeme. In: J. GAUSEMEIER, R. DUMITRESCU, F.J. RAMMIG, W. SCHÄFER und A. TRÄCHTLER, Hg. *10. Paderborner Workshop Entwurf mechatronischer*. Paderborn: Heinz Nixdorf Institut, Universität Paderborn, 2015.
5. KAUL, T., T. MEYER und W. SEXTRO. Integrated Model for Dynamics and Reliability of Intelligent Mechatronic Systems. In: PODOFILLINI ET AL, Hg. *European Safety and Reliability Conference (ESREL2015)*. London: Taylor and Francis, 2015.
6. MEYER, T. *Optimization-based reliability control of mechatronic systems*. Dissertation: Paderborn.
7. MEYER, T., C. SONDERMANN-WÖLKE, J.K. KIMOTHOTH und W. SEXTRO. Controlling the Remaining Useful Lifetime using Self-Optimization. *Chemical Engineering Transactions*, 2013, **33**, 625-630.
8. SONDERMANN-WÖLKE, C. *Entwurf und Anwendung einer erweiterten Zustandsüberwachung zur Verlässlichkeitssteigerung selbstoptimierender Systeme*, 2015.
9. GOEBEL, K., A. SAXENA, M. DAIGLE, J. CELAYA, I. ROYCHOUDHURY und S. CLEMENTS. Introduction to prognostics. In: *European PHM conference*, 2012.
10. KIMOTHOTH, J.K. *Development and Performance Evaluation of Prognostic Approaches for Technical Systems*. Dissertation.
11. BUSCHMEYER, L. *Grundlagen und Anwendungspotential elektroaktiver Polymere*: Heinz Nixdorf Institut, Paderborn, Jan 06.
12. KIMOTHOTH, J.K., C. SONDERMANN-WOELKE, T. MEYER und W. SEXTRO. Machinery Prognostic Method Based on Multi-Class Support Vector Machines and Hybrid Differential Evolution - Particle Swarm Optimization [online]. *Chemical Engineering Transactions*, 2013, **33**, 619-624. Verfügbar unter: doi:10.3303/CET1333104

13. SANKAVARAM, C., B. PATTIPATI, A. KODALI, K. PATTIPATI, M. AZAM, S. KUMAR und M. PECHT. Model-based and data-driven prognosis of automotive and electronic systems. In: *2009 IEEE International Conference on Automation Science and Engineering*, 2009, S. 96-101.
14. MEYER, M., A. UNGER, S. ALTHOFF, W. SEXTRO, M. BRÖKELMANN, M. HUNSTIG und K. GUTH. Reliable Manufacturing of Heavy Copper Wire Bonds Using Online Parameter Adaptation. In: *IEEE 66th Electronic Components and Technology Conference*, 2016, S. 622-628.
15. MEYER, T., J.K. KIMOTHO und W. SEXTRO. Anforderungen an Condition-Monitoring-Verfahren zur Nutzung im zuverlässigkeitsgeregelten Betrieb adaptiver Systeme. In: *27. Tagung Technische Zuverlässigkeit (TTZ 2015) - Entwicklung und Betrieb zuverlässiger Produkte*. Leonberg, 2015, S. 111-122.

Anwendung datenbasierter Methoden auf Werkzeugmaschinen­daten zur Abweichungserkennung vom Normalbetrieb



Dipl.-Ing. (FH) **J. Fisch**, Dr.-Ing. **M. Rossdeutscher**,
Daimler AG, Stuttgart;
Prof. Dr.-Ing. **C. Diedrich**,
Otto-von-Guericke-Universität, Magdeburg

Kurzfassung

Dieser Beitrag stellt eine Methode zur Definition der Soll-Referenz aus Produktionsdaten vor. Ziel ist es, Abweichungen vom Normalbetrieb einer Anlage zu erkennen. Dazu wird auf die generellen Schritte der Studie wie Gewinnung, Vorverarbeitung und Aufbereitung der Daten eingegangen. Anschließend wird eine Signalabstraktion, welche das Signal in Form von statistischen Kennzahlen beschreibt, erläutert. Die Methode wird auf binäre Steuerungsdaten angewendet. Auf Basis von Kennzahlen wird eine Soll-Referenz der Produktionsanlage definiert. Mit Hilfe der Soll-Referenz können die Produktionsdaten in Normal- und Abweichungstakte unterschieden werden. Die entwickelte Methode wird mit einem Anwendungsszenario von einer Werkzeugmaschine aus der Motorenkomponenten-Produktion validiert.

1. Einleitung

Der Komplexitätsgrad von Produktionsanlagen steigt im Zuge der Digitalisierung weiter an [1]. Dies äußert sich dadurch, dass Produktionsanlagen kontinuierlich große Mengen an verschiedenen Daten erzeugen, welche nur teilweise ausgewertet werden [2]. Bei dem heutigen Automatisierungsgrad in der Serienproduktion der Automobilindustrie ist das Verhalten von Maschinen daher nur schwer und mit zunehmendem Zeitaufwand von Einzelpersonen zu überblicken [3, 4].

Große Datenmengen können meist nicht mehr mit herkömmlichen Methoden analysiert werden, daher werden Big-Data-Analysen benötigt um komplexe Systeme zu beherrschen. Ein Teilgebiet von Big-Data-Analysen sind datenbasierte Methoden, welche unmittelbar auf Da-

ten angewendet werden können und dem Menschen die Informationen in verständlicher Form bereitstellen [4 bis 6].

In bisherigen Veröffentlichungen der Zustandsüberwachung von Produktionsanlagen wird die Rohdatenvisualisierung, Auswertung von wenigen Signalen und manuelle Festlegung von Alarmgrenzwerten als Funktion realisiert. Um das Verhalten von Produktionsanlagen ganzheitlich zu beschreiben, müssen alle Daten einer Maschine automatisch analysiert werden. Eine adaptierbare Lösung ist in der Literatur nicht vorhanden [6 bis 9]. Grundlage einer Datenauswertung sind statistische Kennzahlen und deren Grenzwerte, welche eine objektive Beurteilung, einen Soll-Ist-Vergleich und die Erfassung von Zusammenhängen ermöglichen [10].

Auf die grundlegenden Schritte der Datenverarbeitung und verwendeten Kennzahlen wird in Kapitel 2 eingegangen. Anschließend wird im Kapitel 3 eine Methode vorgestellt, um den Normalzustand der Maschine automatisch zu definieren und eine Abweichungserkennung von Produktionstakten durchzuführen.

2. Grundlegende Schritte der Datenverarbeitung

In den folgenden Abschnitten wird ein Konzept zur Datenverarbeitung vorgestellt, welches mit MATLAB umgesetzt wurde. Zuerst wird auf die Auswahl der Rohdaten eingegangen. Auf Grundlage der Rohdaten wird die Datenverarbeitung durchgeführt, welche die Daten einliest, aufbereitet, aussortiert und anschließend in Kennzahlen umwandelt, siehe Bild 1. Die Kennzahlen sind ein notwendiger Bestandteil des Konzepts zur Bildung von Grenzwerten, welches in Abschnitt 3.1. verdeutlicht wird.

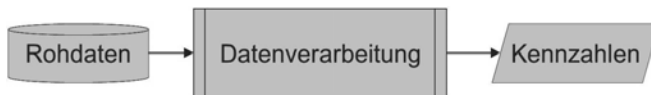


Bild 1: Schritte der Datenverarbeitung

2.1. Rohdaten

Im Rahmen von Industrie 4.0 wird vermehrt von den vorhandenen, aber nicht ausgewerteten Datenmengen gesprochen. Bereits durchgeführte Analysen in der Motorenkomponenten-Produktion ergaben eine Vielzahl von Datenerfassungssystemen. Eine anschließende Auswertung zeigte, dass keines der Systeme Daten zur Beschreibung des gesamtheitlichen

Verhaltens von einer Maschine beinhaltet. Das einzige System, welches eine begrenzte Anzahl an Informationen von Zustandsdaten der Maschine ausgibt, ist ein Betriebsdatenerfassungssystem. Dieses System erfasst Daten wie Stückzahlen, Betriebszustände, Störgründe und Taktzeiten.

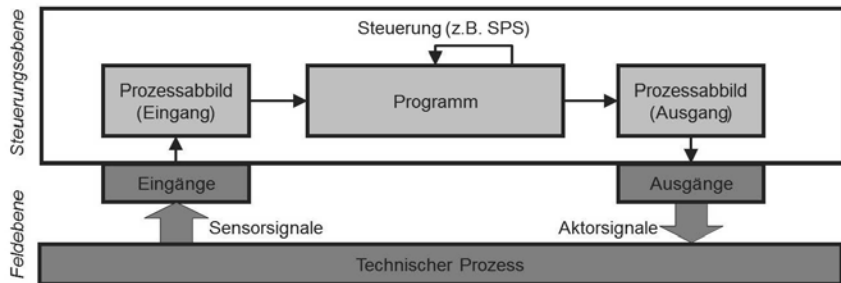


Bild 2: Prinzip des Prozessabbildes einer Steuerung [11]

Weitere Analysen zeigten, dass die Daten der Sensoren und Aktoren einer Maschine geeignet sind um das Maschinenverhalten zu beschreiben. Die in Bild 2 dargestellten Daten aus der Feldebene spiegeln sich im Prozessabbild der Steuerung und können aus der Steuerungsebene aufgezeichnet werden.

Abtastzeitpunkte	Jahr	Monat	Tag	Stunde	Minute	Sekunde	MSekunde	"E 1.4"	"E 1.3"	"E 1.0"	"E 0.7"	...
{	17	1	4	11	58	6	654	0	0	1	0	...
	17	1	4	11	58	6	744	0	0	1	0	...
	17	1	4	11	58	6	819	0	1	1	0	...
	17	1	4	11	58	6	969	0	1	1	0	...
	17	1	4	11	58	7	44	0	1	1	1	...
	17	1	4	11	58	7	119	0	1	1	1	...
	17	1	4	11	58	7	194	0	0	1	1	...
	17	1	4	11	58	7	374	0	0	1	1	...
...

Werte

Bild 3: Auszug Datenerfassung

Die Rohdaten des ausgewählten Untersuchungsobjektes bestehen aus 238 Sensoren und 146 Aktoren, welche z.B. das Bauteil im Greifer spannen, die Ladeluke öffnen oder den Kühlschmierstoffdruck überwachen. Die Daten werden mit einer Abtastrate von durchschnittlich 100 ms erfasst. Die Steuerungsdaten sind Binärdaten, es können nur die Werte „0“ oder „1“ angenommen werden. Die erfassten Daten jedes Parameters werden in einer CSV-Datei für jeden Tag gesammelt, welche rechnerisch 864.000 Abtastzeitpunkte und 330 Mio. Werte

beinhaltet. Ein schematischer Aufbau der Aufzeichnung ist in Bild 3 dargestellt. Die erzeugte Datenmenge für eine Maschine liegt bei knapp 1 GB pro Tag. Skaliert auf 150 Maschinen für ein Jahr werden 54 TB erzeugt.

2.2. Signalaufbereitung

Nach dem Einlesen der Rohdaten in die Entwicklungsumgebung werden die Signale in den folgenden Schritten aufbereitet.

1. **Produktionsdateneinteilung:** Das Konzept baut auf einer taktgebundenen Auswertung auf. Daraus resultiert eine Vergleichbarkeit von Produktionstakten. Die Einteilung der Takte erfolgt durch ein Bit in der Steuerung, das beim Ablegen des Fertigteils gesetzt wird. Die schematische Einteilung von Takten ist in Bild 4 erkennbar.

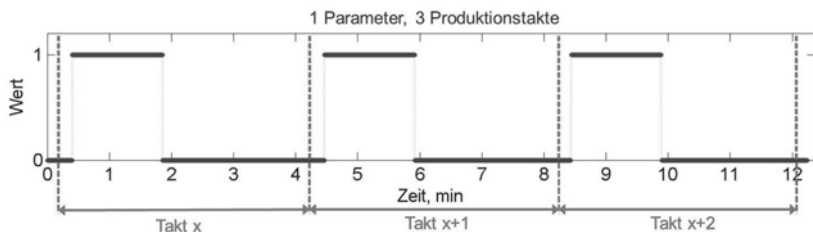


Bild 4: Schematische Produktionstakteinteilung

2. **Bewertung von Takten:** Die Takte sind auf Betriebszustand, Logistikstörung, technische Störung, Kontrollfahrt und Zeitsynchronisation zu bewerten. Für die Identifikation wird die vorhandene Betriebsdatenerfassung genutzt. Takte mit einem dieser Zustände werden in diesem Beitrag nicht verwendet. Zusätzlich werden Takte mit niedrig qualitativen Daten aussortiert. Dieses können Takte mit Datenlücken, zu großen Abständen der Abtastzeitpunkte oder nicht chronologische Zeitstempel sein.
3. **Variantezuordnung zu Takten:** Damit Takte verglichen werden können, muss jedem Takt die Bauteilvariante zugeordnet werden. Es werden nur Takte der gleichen Bauteilvarianten miteinander verglichen.
4. **Ermittlung der Datentypen je Parameter:** Auf Basis der vorhandenen Daten wird eine automatische Bestimmung des Datentyps für jeden Parameter vorgenommen. Mögliche Datentypen können Bool, Integer, Double und String sein. In diesem Beitrag wird ausschließlich auf den Datentyp Bool eingegangen.

Die Ergebnisse der Signalaufbereitung sind aufbereitete Produktionstakte. Die Schritte der Signalaufbereitung sind bei Big-Data-Analysen grundsätzlich notwendig.

2.3. Bildung von Kennzahlen

Unter einer Kennzahl wird eine Zuordnung von Werten zu Informationen verstanden, welche hauptsächlich Anwendung in betriebswirtschaftlichen Prozessen findet. Im entwickelten Konzept handelt es sich um absolute Kennzahlen, da diese quantitativ die Zeitreihe des Produktionstaktes beschreiben. Die in diesem Beitrag vorgestellten Kennzahlen sind beispielhaft für die Signalbeschreibung ausgewählt.

Für die Auswertung von binären Signalen werden drei grundlegende Kennzahlen verwendet. Der erste Kennzahltyp ist die Anzahl der Flankenwechsel, ein Flankenwechsel besteht aus der Zustandsänderung des Signals von „0“ zu „1“ oder „1“ zu „0“. Der zweite Kennzahltyp ist die Dauer eines Intervalls, wobei ein Intervall nur Werte von „0“ oder „1“ beinhaltet. Sobald ein Zustandswechsel eintritt, wird ein neues Intervall berechnet. Auf diese Weise beginnt die Berechnung beim ersten Abtastpunkt eines Wertes und endet beim letzten Punkt des gleichen Wertes. Es kann je nach Parameter eine unterschiedliche Anzahl an Intervallen geben. Der dritte Kennzahltyp ist die Aufsummierung von den Zeitdauern der Intervalle. Folgendermaßen werden alle Intervalle mit dem Wert „0“ summiert und genauso alle mit dem Wert „1“. Die Umsetzung der Kennzahlenberechnungen für einen Produktionstakt und einen Parameter ist in Bild 5 beispielhaft dargestellt.

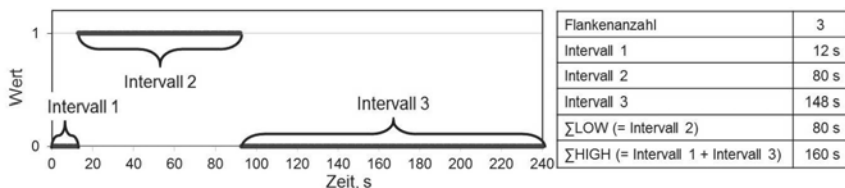


Bild 5: Beispiel eines Produktionstaktes mit Kennzahlen

Durch die Abstraktion der Zeitreihe in Kennzahlen entsteht ein statistisch vergleichbares Maß, welches die Grundlage der Abweichungserkennung ist.

3. Konzept zur Abweichungserkennung vom Normalbetrieb einer Produktionsanlage

In diesem Abschnitt wird auf Basis der Kennzahlen eine Methode zur Definition der Soll-Referenz aus Kennzahlen vorgestellt. Weiterführend wird das Verfahren der Abweichungserkennung erläutert und die Abweichung in einem Produktionstakt mittels Intervallkennzahlen der Parameter eingegrenzt. Ein Überblick über die Vorgehensweise in diesem Abschnitt gibt Bild 6. Die entwickelte Methode basiert auf den Untersuchungen von Boronenko [12].

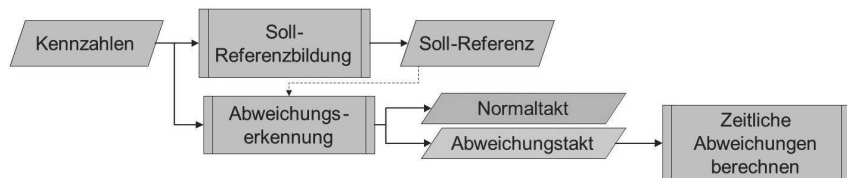


Bild 6: Schritte der Abweichungserkennung

3.1. Bildung der Soll-Referenz

Die Soll-Referenz beschreibt den ermittelten Normalzustand in Taktten von Daten einer Produktionsanlage. Diese wird während der Trainingsphase berechnet und beinhaltet die Grenzwerte der Kennzahlen aller Parameter. Für die Definition der Grenzwerte sind Kennzahlen einer repräsentativen Stichprobenmenge von Produktionstakten zu verwenden, welche die Kriterien aus Abschnitt 2.2 erfüllen.

Um die Grenzen von Kennzahlen robust zu definieren, ist bei der Auswahl einer universellen Grenzwertmethode die Verteilungsform aller Kennzahlen zu berücksichtigen. Analysen der Normalverteilung mittels des Shapiro-Wilk und Anderson-Darling Tests ergaben, dass durchschnittlich 48,17 % der Kennzahlen nicht normalverteilt sind. Die Analysen basieren auf einer Menge von 444 Kennzahlen über 225 Takte. Eine exemplarische Veranschaulichung der grafischen Analyse einer Normalverteilung mittels QQ-Plot und Histogramm für eine nicht normalverteilte Kennzahl ist in Bild 7 dargestellt.

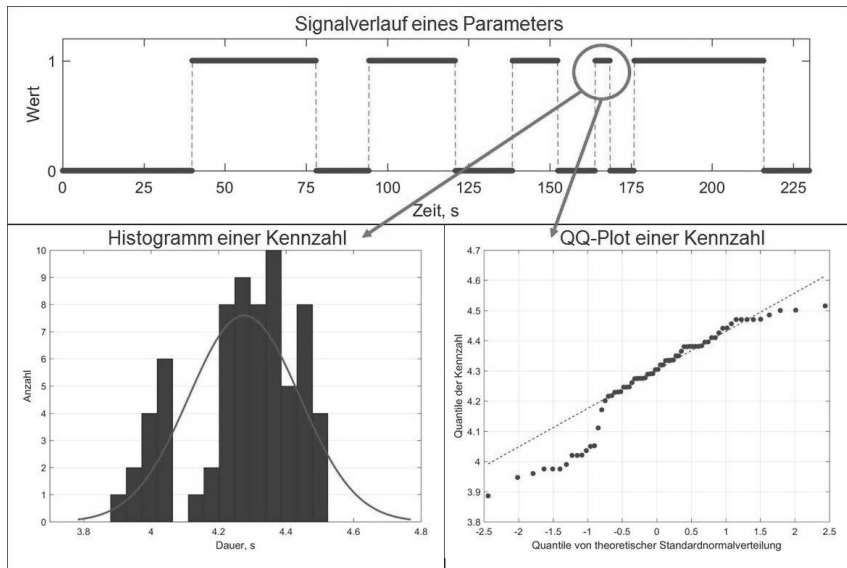


Bild 7: Grafische Untersuchung der Normalverteilung einer Kennzahl

Die Normalverteilung kann nicht für alle Kennzahlen angenommen werden, daher wurde der Interquartilsabstand als robuste Grenzwertmethode ausgewählt. Der obere Grenzwert wird nach der Formel $OG = Q_3 + k \cdot (Q_3 - Q_1)$ berechnet. Der untere Grenzwert wird nach der Formel $UG = Q_1 - k \cdot (Q_3 - Q_1)$ definiert. Die angewendete Methode basiert auf den Lageparametern, unteres (Q_1) und oberes Quartil (Q_3), einer Verteilung. Die Differenz zwischen diesen statistischen Kennzahlen wird als Interquartilsabstand (IQA) bezeichnet. Der IQA umfasst 50% der zentralen Werte der Datenreihe und charakterisiert die Streuung von Werten innerhalb des Intervalls. Der k -Faktor kann als $k = 1,5$ für explorative Studien und als $k = 3,5$ für konfirmative Studien gewählt werden [13]. Die Grenzwertbestimmung von Kennzahlen gehört zur explorativen Statistik. Auf der Grundlage von Hedderich und Sachs wurde daher ein k -Faktor von 1,5 gewählt. Als Beispiel wird in Bild 8 eine Kennzahl mit unterschiedlichen k -Faktoren, Q_1 , Q_3 , OG und UG des k -Faktors von 1,5 dargestellt.

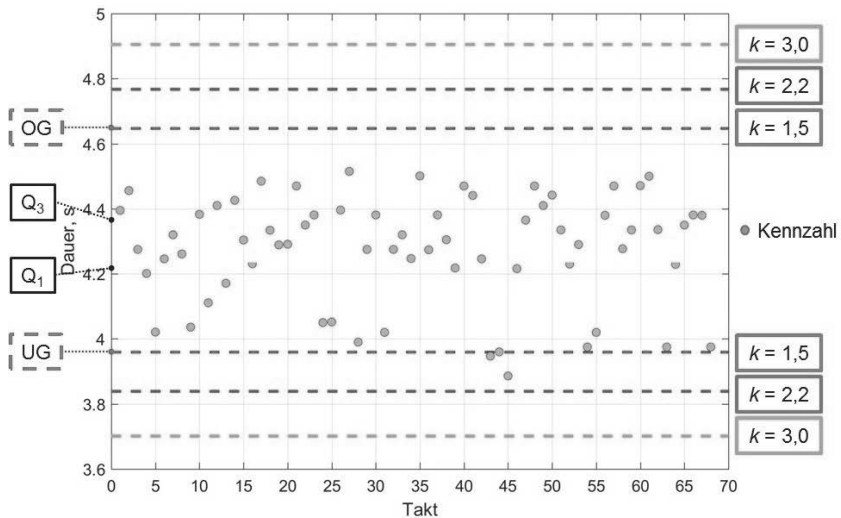


Bild 8: Exemplarische Grenzwertbildung für eine Kennzahl

Die mittels Interquartilsabstand gebildeten Grenzwerte repräsentieren dabei die am häufigsten aufgetretenen, ähnlichen Kennzahlwerte.

3.2. Abweichungserkennung

Nach erfolgreicher Trainingsphase kann anschließend die Testphase erfolgen, dabei wird ein Soll-Ist-Vergleich vorgenommen. Um eine Abweichungserkennung durchzuführen wird die Soll-Referenz für den Vergleich mit den Kennzahlen der Ist-Takte herangezogen. Das Ergebnis einer Abweichungserkennung ergibt entweder einen Normaltakt oder einen Abweichtakt. Ein Abweichtakt liegt vor, sobald eine Kennzahl des Takts außerhalb der Grenzwerte der Soll-Referenz liegt.

Wie in Bild 9 grafisch anhand von zwei Takten des Parameters a ersichtlich ist, weicht die Kennzahl des zweiten Intervalls leicht ab. Ab dem zweiten Intervall in Takt 4 entsteht eine Zeitverschiebung des Signalverlaufs. Takt 1 und Takt 4 waren beides Takte in der Trainingsphase, jedoch wurde Takt 4 als abweichend gekennzeichnet und liegt rund 6 s außerhalb des oberen Grenzwertes für den beispielhaften Parameter a von Intervall 2. Dieses zeigt, dass der Bearbeitungsprozess an der Maschine mit den Schwankungen erfolgt und die Produktionstakte nicht vollständig identisch zueinander sind.

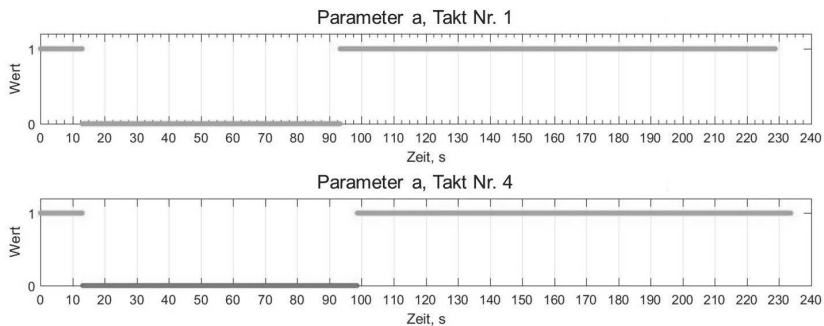


Bild 9: Vergleich von Signalverläufen des Parameter a in zwei Takten

3.3 Visualisierung von zeitlichen Abweichungen

Im nächsten Schritt wird zusätzlich zur Abweichungserkennung von Takten die zeitliche Identifizierung von Abweichungen im Takt vorgenommen. Hierbei werden die abweichenden Parameter in einem abweichenden Produktionstakt untersucht und durch die Schnittmenge wird der Abweichungszeitraum identifiziert. Ein abweichendes Intervall eines Parameters kann durch ein normales Intervall eines anderen Parameters in der zeitlichen Abweichung eingegrenzt werden. Mithilfe einer großen Anzahl an Parametern können so Abweichungszeiträume genauer als mit einzelnen Parametern lokalisiert werden. Zusätzlich werden die Abweichungsintervalle gewichtet, umso länger ein Intervall ist, desto weniger fließt dieser in die Bewertung mit ein.

Durch die übergreifende Auswertung von Parameterabweichungen lässt sich die Abweichung in einem Produktionstakt eingrenzen und genauer lokalisieren. Die grafische Übersetzung von Parameterabweichungen zu Taktabweichungen ist als Ausschnitt in Bild 10 dargestellt. Die oberen, rot gekennzeichneten vier Parameter sind in unterschiedlichen Intervallen abgewichen. Die Zusammenfassung als Takt im unteren Teil des Bildes zeigt zwei Abweichungen in den Bereichen von 94 bis 122 s und von 180 bis 205 s. Eine Erwartung bei der Beurteilung von Bild 10 des einzelnen Parameters 4 ist eine Abweichung in dem Zeitraum von 41 bis 129 s des Taktes. Mittels der Kombination von Parameter 4 und 2 kann hingegen eine Abweichung in den ersten 95 s ausgeschlossen werden, da Intervall 2 und 3 von Parameter 2 einem Normalverlauf entsprechen und somit die Abweichung für diesen Zeitbereich aufheben.

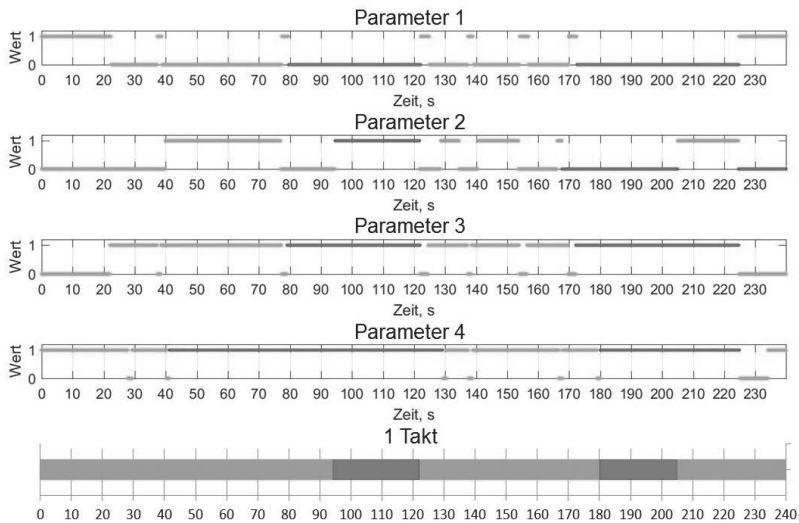


Bild 10: Grafische Darstellung von abweichendem Zeitraum in einem Takt

4. Umsetzungsbeispiel für die Abweichungserkennung

Als Beispiel für die Umsetzung des in Abschnitt 3 erklärten Konzeptes wird die Änderung des Bearbeitungsprogrammes aus zwei Produktionstagen verwendet. Das Untersuchungsobjekt ist eine Werkzeugmaschine aus der Motorenkomponenten-Produktion mit zwölf Achsen, welche den Rohling für eine Nockenwellenbuchse bearbeiten. Die erfassten Daten dieser Maschine wurden bereits in Abschnitt 2.1 erläutert und bestehen aus binären Steuerungsdaten. Nach erfolgreicher Signalaufbereitung ist es möglich aus den Kennzahlen eine Soll-Referenz zu bilden und mit neuen Daten aus anderen Produktionstakten zu vergleichen.

Die Soll-Referenz wurde mit Hilfe von 68 Takten trainiert. Von den Trainingsdaten wurden sieben Takte als Abweichungstakte mit geringer Abweichung in einzelnen Parametern identifiziert. Ein Takt hatte eine stärkere Abweichung, die Taktzeit wurde um 25% überschritten. Ursache dessen war eine Reduzierung der Bearbeitungsgeschwindigkeit durch eine vorhergehende Logistikstörung. Die Testdaten des zweiten Tages, welcher 44 Tage später aufgezeichnet wurde, bestanden aus 30 Takten. Bei der Abweichungserkennung wurden alle Takte als stark abweichend identifiziert. Mit Hilfe der in Abschnitt 3.3. erklärten Methode konnten die Abweichungen den entsprechenden Zeiträumen zugeordnet werden. Die Ursache der

Abweichung war eine Änderung des Bearbeitungsprogrammes. Dabei haben sich die Kennzahlen der Intervalllänge verändert, jedoch ist die Anzahl der Flankenwechsel gleichgeblieben. Daraus resultiert, dass die Steuerungsdaten zwar die Symptome der Änderungen beinhalten, aber keinen Aufschluss über eine Musteränderung geben. Somit ist die Ursache kein Teil der erfassten Daten.

5. Zusammenfassung und Ausblick

In diesem Beitrag wurde ein Konzept zur Abweichungserkennung vom Normalbetrieb einer Produktionsanlage vorgestellt. Zusätzlich wurde eine Methode dargestellt, welche die abweichenden Zeiträume je Takt identifiziert. Anschließend wurde auf einen Anwendungsfall aus der Motorenkomponenten-Produktion eingegangen und die Methode überprüft. Das Konzept ermöglicht eine benutzerunabhängige Auswertung von großen Mengen an Produktionsdaten im Rahmen von Big-Data-Analysen.

In zukünftigen Arbeiten ist das Konzept auf weitere Daten und Datentypen, wie Antriebsdaten einer CNC-Maschine, zu übertragen. Dazu müssen zunächst entsprechende Schritte für die Signalaufbereitung und -abstraktion von analogen Signalen in Form von Kennzahlen herausgefunden werden. Außerdem ist es notwendig die gewählten Kennzahlen auf deren Aussagekraft zu validieren. Mittels weiterer Szenarien sind die Sensibilität und das ganzheitliche Potential der Methode festzustellen.

Literaturangaben

- [1] Steinhilper, R., Westermann, H.-H., Butzer, S., Haumann, M. u. Seifert, S.: Komplexität messbar machen. Eine Methodik zur Quantifizierung von Komplexitätstreibern und -wirkungen am Beispiel der Refabrikation. ZWF - Zeitschrift für wirtschaftlichen Fabrikbetrieb 107 (2012), S. 360–365
- [2] Industrie 4.0 - Chancen und Herausforderungen der vierten industriellen Revolution, Studie, PWC u. Strategy&, 2014
- [3] Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0. Abschlussbericht des Arbeitskreises Industrie 4.0, Deutschlands Zukunft als Produktionsstandort sichern, Forschungsunion u. acatec, 2013
- [4] Baron, P.: Big Data für IT-Entscheider. Riesige Datenmengen und moderne Technologien gewinnbringend nutzen. München: Carl Hanser Verlag 2013

- [5] Diesner, M.: Massendaten im Fertigungsumfeld. *visIT Big Data* 17 (2016), S. 6–7
- [6] Engel, P.: Entwicklung neuer Diagnoseverfahren auf der Basis multivariater Methoden und deren Anwendung in einem Condition Monitoring System, Universität Duisburg-Essen Diss. 2012
- [7] Jermy, G. T., Castle, R. C., Gimblett, C. F. u. Chakrabarti, R.: Monitoring out of normal conditions in repetitive cycle production machinery. *Fifth International Conference on Factory 2000 - The Technology Exploitation Process*. 1997, S. 29–33
- [8] Emonts-Holley, R.: Smart Maintenance: Industrie 4.0 in der Instandhaltung. Integrative Softwarelösungen für ein intelligentes, bedarfsorientiertes Instandhaltungsmanagement in komplexen Produktionsumgebungen. *UdZ - Unternehmen der Zukunft* 16 (2015) 1, S. 32–34
- [9] Schwenken, U.: Eine Methode zur Fehlerbewertung und zur adaptiven Motorleistungsbegrenzung auf der Basis einer modellbasierten Diagnose am Beispiel eines PKW-Kühlsystems, Ruhr-Universität Bochum Diss. 2006
- [10] VDMA 66412-1; 2009-10. *Manufacturing Execution Systems (MES) - Kennzahlen*
- [11] Favre-Bulle, B.: Automatisierung komplexer Industrieprozesse. Systeme, Verfahren und Informationsmanagement. Wien: Springer-Verlag 2004
- [12] Boronenko, K.: Entwicklung einer Methode zur Ermittlung der Soll-Referenz aus Produktionsdaten, Otto-von-Guericke-Universität Magdeburg Masterthesis 2016
- [13] Hedderich, J. u. Sachs, L.: Angewandte Statistik. Methodensammlung mit R. Heidelberg: Springer Spektrum 2016

Zuverlässigkeitsprognose in der Anlaufphase für „x-Months-in-Service“: Präzisierung und Validierung der Beanstandungsquote



x-Months-in-Service Reliability Prediction During Ramp-Up: Specification and Validation of Complaints Quota

Dipl.-Ing. **Timo Rieker**, Dipl.-Ing. **Michael Bartholdt**,
Prof. Dr.-Ing. **Bernd Bertsche**, Institut für Maschinenelemente (IMA),
Universität Stuttgart;
Prof. Dr.-Ing. **Peter Zeiler**, Hochschule Esslingen,
Fakultät Mechatronik und Elektrotechnik;
Dr.-Ing. **Andreas Jacobi**, Daimler AG, Sindelfingen

Kurzfassung

Die Beanstandungsquote cph („complaints per hundred“) nach x-monatiger Betriebszeit (z.B. 12MiS, Months in Service) wurde basierend auf [1] weiterentwickelt, vgl. [2]. Die Zuverlässigkeitsprognose für 12MiS wird in der Anlaufphase einer Fahrzeugproduktion auf eine Baureihe nach 3, 6 und 9 Monaten nach Markteinführung und ein Motorprojekt retrospektiv angewandt und mit der realen Beanstandungsquote nach 12MiS verglichen, um das Vorgehen zu validieren.

Abstract

The complaints quota cph (complaints per hundred) after x months in service (MiS) is further specified based on [1] and [2]. The reliability prediction for 12MiS is applied retrospectively to a vehicle project during ramp-up after 3, 6 and 9 months after market launch as well as to an engine project. The obtained predicted values are compared to the actually observed quotas after 12MiS in order to validate the approach.

1. Einleitung und Motivation

Die Prognose der Zuverlässigkeit hilft, Ausfälle bzw. Beanstandungen nach Art und Häufigkeit vorauszusagen, woraufhin sich geeignete Korrekturmaßnahmen ableiten lassen. Sie ist umso wertvoller, je früher sie so präzise wie möglich erfolgt. Dann sind Maßnahmen zur Fehlerabstellung und Steigerung der Kundenzufriedenheit effizienter und effektiver. Zuverlässigkeitsprognosen während der Entwicklung basieren auf Testkombinationen verschiedener Prototypen auf beliebigen Systemebenen und bei ggf. unterschiedlichen Reifegraden. Zuverlässigkeitsprognosen während des anschließenden Produktionsanlaufs sollten sich auf das Feldverhalten stützen, das in dieser ersten Nutzungsphase beobachtet wird. Der Einbezug von Bean-

standungen statt Ausfällen in die Prognose spiegelt noch besser die Kundenzufriedenheit wieder, an der die Unternehmensprozesse eines Premium-Automobilherstellers ausgerichtet sind. Die Prognose des Beanstandungsgeschehens über z.B. den Garantie- und Kulanzzeitraum in Verbindung mit der Kenntnis der anfallenden Behebungskosten lässt die Abschätzung der sich so kumulierenden Kosten zu. Im Rahmen dieser Arbeit wird eine kundennahe und insbesondere in Anlaufphasen anwendbare Möglichkeit der Zuverlässigkeitsprognose präzisiert und validiert.

2. Die 12MiS-Beanstandungsquote

Die Management-Kennzahl cph („complaints per hundred“) basiert auf Felddaten und lässt sich bereits auf Daten aus der frühen Anlaufphase eines Neuproduktprojekts anwenden. Sie berücksichtigt daher den höheren Reifegrad des Serienproduktes gegenüber einem Prototyp. Der Prognosehorizont, d.h. „x“ in „x-Months-in-Service“ ist dabei variabel. Unter „Months-in-Service“ (MiS) wird die Zeit verstanden, in der das Fahrzeug in Kundenhand ist. Dieser Zeitraum beginnt mit dem Tag der Erstzulassung. Im Rahmen dieser Arbeit wird auf die 12MiS-Prognose eingegangen, da sie als Referenzwert im Reliability-Growth-Monitoring und im Fehlerabstellprozess der Daimler AG [2] dient.

Zugrunde gelegt wird das Verhältnis der Anzahl Ausfälle bezogen auf die Anzahl aller möglichen Ausfälle (d.h. aller Fahrzeuge) und dem Grad des Erreichens der erwarteten Gesamtleistung (kumulativ erreichte Laufleistung zum Prognosezeitpunkt bezogen auf die erwartete, kumulative Gesamtleistung zum prognostizierten Zeitpunkt), vgl. Gleichung (1).

$$\begin{aligned}
 & \frac{\frac{\text{Anzahl der Ausfälle}}{\text{Anzahl aller Fzg}}}{\frac{\text{Kumulativ erreichte Laufl.}}{\text{Kumulativ erwartete Gesamtleistung}}} = \frac{\frac{\text{Anzahl der Ausfälle}}{\text{Kumulativ erreichte Laufl.}}}{\frac{\text{Anzahl aller Fzg}}{\text{Kumulativ erwartete Gesamtleistung}}} \quad (1) \\
 & = \frac{\frac{x}{\sum_i LL_i(t)}}{n \cdot LL_{12-MiS}} = \frac{x}{\sum_i LL_i(t)} \cdot LL_{12-MiS}
 \end{aligned}$$

Schätzt man die Anzahl der Ausfälle x über den Chi-Quadrat-Schätzer und die charakteristische Lebensdauer T einer Weibull-Verteilung über den Maximum Likelihood-Schätzer für T und integriert beides in die Zuverlässigkeit ausgedrückt als Weibull-Verteilung, entsteht folgender Ausdruck [3], sofern man Beanstandungen mit Ausfällen gleichsetzt:

$$R = e^{-cph} \quad (2)$$

Gleichung (3) zeigt die Berechnung der Beanstandungsquote mit den zugehörigen Eingangsgrößen.

$$cph(t, LL_{12-MiS}) = \frac{\chi^2_{2(x+1); P_A}}{2 \sum_i LL_i(t)^b} \cdot LL_{12-MiS}^b \cdot 100 \quad (3)$$

mit

cph :	„complaints per hundred“	x :	Anzahl an Beanstandungen
$\frac{\chi^2_{2(x+1);P_A}}{2}$:	Chi-Quadrat-Schätzer für Beanstandungen	LL_i :	Kumulativ erreichte Laufleistung
P_A :	Vertrauensbereich	b :	Formparameter der Weibullverteilung
LL_{12-MIS} :	Soll-Laufleistung nach 12 Monaten im Service	t :	Prognosezeitpunkt

Diese Gleichung (3) basiert auf einem Ansatz von Jordan et al. [1], vgl. [2]. Die Verwendung eines Chi-Quadrat-Schätzers für den Erwartungswert der Beanstandungen erlaubt es, Konfidenzintervalle zu den jeweils berechneten Quoten anzugeben. Sie basieren auf der Anzahl Beanstandungen. Bild 3 zeigt einen cph-Plot mit 90-%igem Vertrauensbereich.

Nachfolgend werden die Eingangsgrößen in Gleichung (3) näher untersucht und die Vorgehensweise bei deren Ermittlung vorgestellt.

2.1 Kumulativ erreichte Laufleistung LL_i

Der prinzipielle Umgang mit Feldqualitätsdaten zu beanstandeten Fahrzeugen und nicht-beanstandeten Fahrzeugen, im Folgenden als Survivor bezeichnet, ist in [4] beschrieben und dient hier als Grundlage.

Die kumulativ erreichte Laufleistung LL_i umfasst die jeweilige Laufleistung aller Fahrzeuge im Feld. Hierzu zählen beanstandete ebenso wie nicht-beanstandete Fahrzeuge. Noch nicht zugelassene Fahrzeuge werden nicht berücksichtigt.

Über eine Schnittstelle zur GuK-Datenbank stehen Informationen über die produzierten und beanstandeten Fahrzeuge zur Verfügung. Zum Prognosezeitpunkt können die Fahrzeuge wie in Bild 1 dargestellt unterteilt werden. Für die hiernach entstandenen Fahrzeuggruppen erfolgt die Datengewinnung zur Prognose auf verschiedene Weise.

Für beanstandete Fahrzeuge liegen alle zur Prognose benötigten Informationen (Erstzulassung, km-Stand und Dauer des Werkstattaufenthalts) fahrzeugindividuell und besonders während der Garantie- und Kulanzzeit (also auch während der Anlaufphase) in hoher Dichte vor. Die effektive Betriebszeit setzt sich aus der Betriebszeit vor und nach Werkstattaufenthalt zusammen. Für die Laufleistung nach Werkstattaufenthalt wird die individuelle Laufleistung pro Tag vor Werkstattaufenthalt zugrunde gelegt.

Alle produzierten Fahrzeuge gehören entweder zu den K beanstandeten oder den J nicht beanstandeten Fahrzeugen (Survivor):

$$\sum_{i=1}^n LL_i(t)^b = \sum_{j=1}^J LL_{Survivor_j}(t)^b + \sum_{k=1}^K LL_{Beanstandete_k}(t)^b \quad (4)$$

Für die Survivor liegen als für die Prognose relevante Informationen nur das Produktionsdatum, die Motorbauart und das vorgesehene Vertriebsland aus dem Produktionsbericht vor.

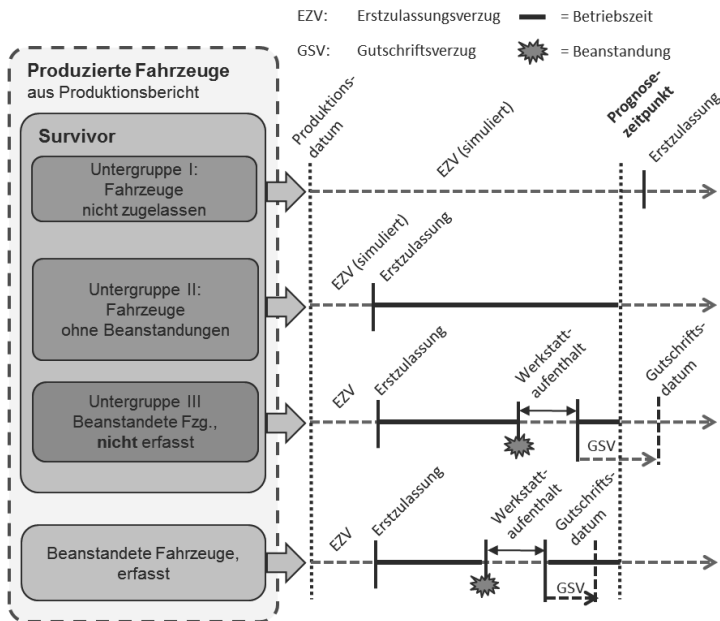


Bild 1: Übersicht über die verschiedenen Zustände der produzierten Fahrzeuge zum Prognosezeitpunkt

Diesen Fahrzeugen werden ein individuelles simuliertes Erstzulassungsdatum und eine simulierte Laufleistung pro Tag per Zufallsgröße aus repräsentativen Verteilungen zugewiesen.

Die zugrunde liegenden Verteilungen werden zuvor aus Beanstandungsdaten der Vorgängerbaureihe ermittelt (vgl. [1]), deren Datenmenge die der gerade angelaufenen aktuellen Baureihe bei weitem übersteigt. Jedoch werden im Rahmen der Validierung (vgl. Kapitel 3) Verteilungen aus Daten derselben Baureihe zugrunde gelegt. Um den unterschiedlichen Laufleistungen zwischen Fahrzeugen mit Benzin- oder Dieselmotor Rechnung zu tragen, und weil deren Verhältnis über die Produktionswochen nicht konstant ist, wurde bei den Laufleistungsverteilungen zwischen Benzin und Dieselfahrzeugen unterschieden. Die länderspezifischen Erstzulassungsverzüge weisen logistisch bedingt große Unterschiede auf. Aus Beanstandungsdaten wurden für gebildete Ländergruppen entsprechende Verteilungen der Erstzulassungsverzüge für die Zufallszahlen der aktuellen Survivor zugrunde gelegt.

Durch die Zuordnung des Erstzulassungsverzugs per Zufallsgröße ergeben sich auf der Seite der Survivor zunächst 3 Untergruppen.

Survivor-Untergruppe I: Fahrzeuge, deren zugewiesenes Erstzulassungsdatum nach dem Prognosezeitpunkt ist, generieren keine Laufleistung.

Survivor-Untergruppe II: Fahrzeuge, die sich in Folge des Erstzulassungsdatums im Feld befinden, generieren Laufleistung. Der zum Prognosezeitpunkt simulierte km-Stand ergibt sich als Produkt aus den (simulierten) Tagen im Einsatz und der zugewiesenen Laufleistung pro Tag.

Survivor-Untergruppe III: Eine dritte Untergruppe (vgl. Bild 1) der Survivor entsteht infolge des Gutschriftsverzugs und betrifft Fahrzeuge, die zum Zeitpunkt der Prognose bereits beanstandet wurden, diese Beanstandungen aber noch nicht im Beanstandungsbericht erfasst sind. Der Gutschriftsverzug (GSV) beschreibt die verstrichene Zeit zwischen Reparaturdatum und der Eintragung in das GuK-System. Diese faktisch beanstandeten Fahrzeuge würden irrtümlicherweise den nicht beanstandeten Fahrzeugen der Untergruppe II zugeordnet werden und das Prognoseergebnis verfälschen. Ein Verfahren zur Bestimmung ihres Anteils an der Survivor-Gruppe wird in Kapitel 2.2 vorgeschlagen. Dass die Beanstandungen dieser Fahrzeuge noch nicht im GuK-System erfasst sind, bedeutet nicht, dass die Behebung der Beanstandung (Reparatur) noch nicht abgeschlossen ist. Zur Berücksichtigung ihrer Zeitdauer im funktionsfähigen Zustand wird die theoretisch zur Verfügung stehende Zeit im Feld in funktionsfähigem Zustand (Prognosedatum abzüglich simuliertes Erstzulassungsdatum) um einen mittleren Werkstattaufenthalt basierend auf Vorgängerdaten reduziert. Die Laufleistung dieser dritten Untergruppe (vgl. Bild 1) zum Prognosezeitpunkt ergibt sich wieder als Produkt aus den so (simulierten) Tagen im Einsatz und der (simulierten) Laufleistung pro Tag.

2.2 Anzahl der Beanstandungen x

Die Anzahl der in Kapitel 2.1 der Survivor-Untergruppe III zugeteilten „Survivor“ (die ja eigentlich ausgefallen sind) sollte als zusätzliche Anzahl Beanstandungen in die Prognose einfließen. Die Beanstandungsquote wird sonst fälschlicherweise zu klein, d.h. zu optimistisch, berechnet (vgl. Gleichung (1)).

Wie in Kapitel 2.1 erwähnt, spielt der Gutschriftsverzug in diesem Zusammenhang eine entscheidende Rolle. Bild 2 zeigt qualitativ eine aus den Beanstandungsdaten ermittelte Verteilung des Gutschriftsverzugs. Eingezeichnet sind zwei Prognosezeitpunkte (Daten-Abzug t_1 , Daten-Abzug t_2), der Zeitpunkt des beanstandungsverursachenden Ereignisses und 3 Bereiche.

Die Fläche unter der Dichte des Gutschriftsverzugs entspricht der Wahrscheinlichkeit, zu der die Beanstandungen nach Auftritt des erzeugenden Ereignisses in die Datenbank eingetragen werden. Der Bereich 1 der Verteilung symbolisiert den Anteil Beanstandungen die bis zum Prognosezeitpunkt t_1 erfasst sind. Diejenigen Beanstandungen zu Zeitpunkten $t > t_1$, also in den Bereichen 2 und 3, können im Abzug zum Zeitpunkt t_1 in diesem Beispiel noch nicht erfasst worden sein. Erst mit fortlaufender Zeit werden diese fehlenden Beanstandungen im G&K-System eingetragen und somit einsehbar. Das Bestreben, so früh wie möglich eine Prognose

zum Beanstandungsgeschehen abgeben zu können steht folglich im Widerspruch zum Anliegen, möglichst alle Beanstandungen zu einem Prognosezeitpunkt erfasst zu haben.

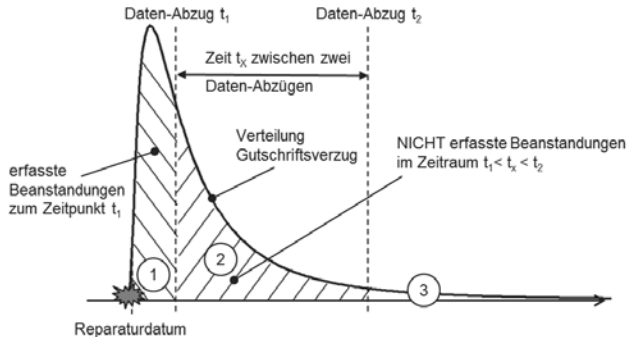


Bild 2: Einfluss des Gutschriftsverzugs auf die erfassten Beanstandungen

Die Einführung einer Quote q_x zur Berücksichtigung dieser nicht erfassten Beanstandungen schafft Abhilfe. Hierzu wird der Quotient aus Beanstandungen gebildet, die im Intervall $t_x = t_2 - t_1$ eingetragen wurden, aber Reparaturtermine vor t_1 aufweisen und vor t_1 aufgetretenen und eingetragenen Beanstandungen, vgl. Gleichung (5). Bei der Festlegung des Intervalls gilt es einen pragmatischen Kompromiss zwischen der Genauigkeit der Quote und der zeitlichen Verzögerung zu finden. Je größer das Intervall gewählt wird, desto wahrscheinlicher ist es, dass alle relevanten Beanstandungen erfasst sind. Demgegenüber steht die zeitliche Verzögerung der Prognose und damit einhergehend der Verlust der kurzfristigen Reaktionsfähigkeit auf mögliche Zuverlässigkeits- bzw. Qualitätsprobleme im Anlauf. Die Anzahl der zum Zeitpunkt t_2 infolge des Gutschriftsverzugs nicht erfassten Beanstandungen mit Eintrittszeitpunkt $< t_1$ lässt sich zum Zeitpunkt t_2 aus dem dann vorliegenden Beanstandungsbericht ermitteln. Es handelt sich um Beanstandungen mit Reparaturtermin vor t_1 , die zwischenzeitlich eingetragen sind. Die Quote wird wie folgt berechnet:

$$q_x = \frac{\sum x' \mid (t_{x'} < t_1; t_{\text{Eintrag von } x'} \in [t_1, t_2])}{\sum x'' \mid (t_{x''} < t_1; t_{\text{Eintrag von } x''} \in (0, t_1))} \quad (5)$$

Mit $x \triangleq$ Beanstandung, $t_x =$ Zeitpunkt des die Beanstandung hervorruftenden Ereignisses.

Die Gesamtanzahl an Beanstandungen x , die in die Berechnung der Beanstandungsquote („cph“) einfließt, ergibt sich letztendlich als Summe der erfassten Beanstandungen x' und der über die Quote zu erwartenden Beanstandungen x'' :

$$x = \sum x' + \sum x'' = \sum x'' (1 + q_x) \quad (6)$$

2.3 Weitere Eingangsgrößen in die 12MiS-Prognose

Soll-Laufleistung LL_{12-MiS} :

Die Soll-Laufleistung ergibt sich z.B. aus der mittleren Laufleistung einer Vorgängerbaureihe. Hierbei ist zu beachten, dass je nach Motorausstattung unterschiedliche Laufleistungsverteilungen entstehen können. So weisen Fahrzeuge mit Dieselmotoren deutliche höhere jährliche Laufleistungen auf als Fahrzeuge der gleichen Baureihe mit Benzinmotoren. Diese Unterschiede wurden im Rahmen der Validierung berücksichtigt. Eine motorenspezifische Auswertung ist damit genauer möglich.

Alternativ könnten als Soll-Laufleistung auch fest vorgegebene Werte angesetzt werden.

Weibull-Formparameter b :

Auszugehen ist vom Formparameter der Weibullverteilung des zu untersuchenden Systems. Für komplexe Systeme kann nach [3] $b = 1$ angenommen werden, wenn ≥ 3 Ausfallmechanismen verschiedener Formparameter zusammengefasst werden. Diese Annahme unterstellt exponentielles Ausfallverhalten bzw. Zufallsausfälle. Im Rahmen der Validierung der cph-Formel (vgl. Gleichung (1)) zur Prognose des Beanstandungsgeschehens wird $b = 1$ angenommen. Alternativ könnte der b -Wert des Vorgänger-Systems unterstellt werden.

2.4 Sensitivitäten der 12MiS-Prognose

Entsprechend der Lage der Eingangsgrößen in die 12MiS-Prognose, vgl. Gleichung (3), lässt sich feststellen, dass eine größere Anzahl Beanstandungen zu einem größeren cph-Wert führt. Ebenso verhält sich die cph bezüglich der kumulativ erwarteten Gesamtlaufleistung LL_{12-MiS} . Demgegenüber steht die kumulierte Laufleistung, die den Effekt durch x und LL_{12-MiS} auf die cph zu kompensieren hilft. Bild 3 (Prognosezeitpunkt 15.09.2014) stellt im Hintergrund dar, wieviel Kilometer pro Produktionswoche kumulativ zusätzlich nötig sind, um eine weitere Beanstandung, d.h. $x+1$, aufzuwiegen.

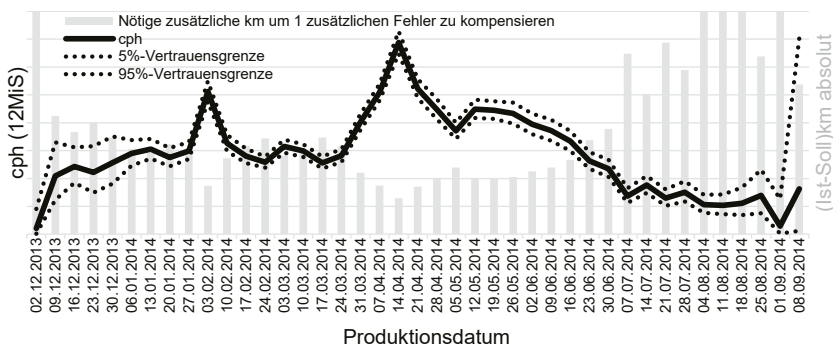


Bild 3: Benötigte zusätzliche Kilometer um einen zusätzlichen Fehler zu kompensieren

Die Anzahl der benötigten zusätzlichen Kilometer ist über die Produktionswochen nicht konstant. Sie ist insbesondere abhängig von der Anzahl der bereits beobachteten Beanstandungen x und den bereits verbuchten Flottenkilometern (LL_i). Auffällig ist der bezüglich einer Veränderung der Ausfallanzahl sehr sensitive Randbereich. Der innere Bereich ist diesbezüglich etwas stabiler. Entsprechend vorsichtig ist in den Randbereichen die Prognose zu bewerten. In [1] wird empfohlen, dass die Prognosen bezüglich Produktionsdaten der letzten 2-3 Monate ab Prognosezeitpunkt ausgeblendet werden sollten. Zu gering sei ihre Datengrundlage infolge zu weniger vorliegender Beanstandungen bei einer gleichzeitig noch kleinen Anzahl zugelassener Fahrzeuge (Erstzulassungsverzug) bzw. kumulierten Kilometer.

Ähnliche Effekte ergaben sich bei Analysen, die folgende Fragestellungen untersuchten:

- Welche Veränderung an Kilometern/ Beanstandungen ist nötig, um den cph-Wert der Vorwoche zu halten, also das Schwanken der cph-Werte zu unterbinden?
- Wieviele Kilometer (ggf. pro Fahrzeug) müssten zusätzlich (bzw. könnten weniger) kumuliert werden, um ein cph-Ziel zu erreichen?
- Wieviele Beanstandungen müssten vermieden werden (bzw. dürften noch zusätzlich auftreten), um ein cph-Ziel zu erreichen?

Insbesondere letztere Fragestellung kann dabei helfen, in Verbindung mit einer Analyse der tatsächlich beobachteten Fehler nach Folge (auch finanziell), Art, Ursache und Häufigkeit basierend auf den fortwährend gepflegten Beanstandungsdaten, geeignete Maßnahmen zur Qualitätsoptimierung zu quantifizieren und schließlich zu priorisieren.

Bei $\Sigma LL_i^b > LL_{12MIS}^b$ wird der Chi-Quadrat-Schätzer verringert (sonst überhöht). $\Sigma LL_i^b < LL_{12MIS}^b$ tritt nur zu sehr frühen Prognosezeitpunkten bei ggf. zusätzlich sehr kleiner Anzahl der im Feld befindlichen Fahrzeuge ($=\Sigma i$) oder vergleichsweise großen LL_{12MIS} auf.

Der Formparameter b der unterstellten Weibull-Verteilung wirkt sich ebenfalls auf den cph-Wert aus. Folgendes kann festgestellt werden:

- Bei $LL_i < LL_{12MIS}$ ($LL_i > LL_{12MIS}$) nimmt der cph-Wert bei größer werdendem b -Wert zu (ab).
- Je größer (kleiner) Σi , desto kleiner (größer) ist dieser Effekt.

3. Validierung

Die Anwendung der Prognose erfolgt retrospektiv auf eine Fahrzeugbaureihe, abgek. FBR und einer Benzinmotor-Baureihe, abgek. MBR. Die Validierung besteht hier aus dem Vergleich der prognostizierten mit den ermittelten realen cph-Werten, vgl. Bilder 4 und 5, für die Produktionsdaten während der Anlaufphase zwischen 02.12.2013 und 08.12.2014, also (von Start of Production, abgek. SOP, bis Erreichen der Kammlinie). Es wurden 3 Stichtage als Prognosezeitpunkte festgelegt, 3 Monate, 6 Monate und 9 Monate nach Markteinführung (15.03.2014) der betrachteten Fahrzeug-Baureihe.

Beim Vergleich der cph-Prognose mit dem realen Wert ist generell zu beachten, dass sich für die reale cph je nach betrachtetem Daten-Abzugsdatum (entspricht Betrachtungszeitpunkt) in

Folge des Erstzulassungs- und Gutschriftsverzugs unterschiedliche Werte ergeben. Die realen Quoten sind in Bild 4 und Bild 5 als durchgezogene Linien dargestellt. Je später die Betrachtung bezogen zum Produktionsdatum, für das prognostiziert wird, desto höher der Anteil der Beanstandungen, die dem entsprechenden Produktionsdatum zugeschrieben werden und die das Kriterium erfüllen, nach maximal 12 Monaten im Feld aufgetreten zu sein (vgl. 12MiS-Prognose). Das ist für alle sechs realen cph-Kurven (vgl. Bild 4 und Bild 5) gegeben.

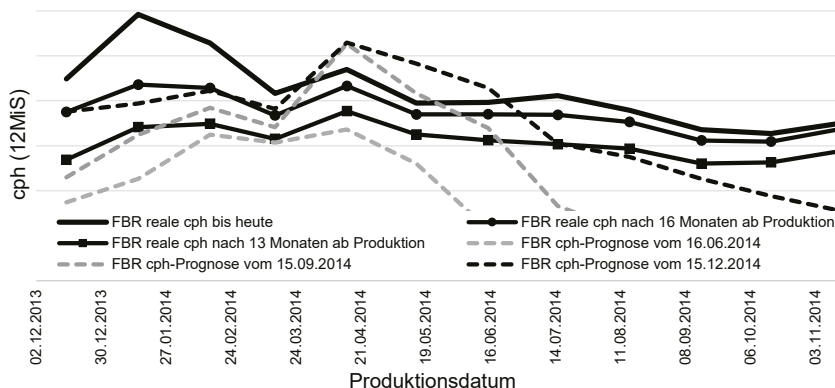


Bild 4: cph-Prognosen im Vergleich mit der retrospektiv realen cph-Quote für die betrachtete Fahrzeug-Baureihe

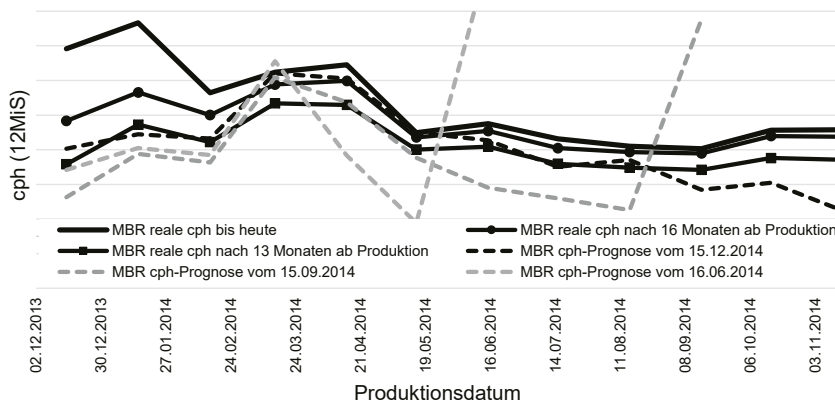


Bild 5: cph-Prognosen im Vergleich mit der retrospektiv realen cph-Quote für die betrachtete Motor-Baureihe

Die Bilder 4 und 5 zeigen die Verläufe der 12MiS-Prognosen zu den Prognosezeitpunkten 16.06.2014, 15.09.2014 und 15.12.2014 sowie den Verlauf der realen cph-Werte, bezogen auf Produktionsdaten der Anlaufphase (KW49 2013 bis KW51 2014).

Die im ersten Drittel des betrachteten Zeitraums (vgl. Bild 4) produzierten Fahrzeuge weisen einen höheren tatsächlichen Erstzulassungsverzug auf, da diese noch vor Markteinführung (15.03.) produziert wurden. Diesen Fahrzeugen (Survivor) wird bei Pauschalisierung des Erstzulassungsverzugs fälschlicherweise ein zu frühes Erstzulassungsdatum als Zufallsgröße zugewiesen, was zu einer größeren generierten Laufleistung durch längere Betriebszeit und letztendlich zu einer zu positiv gerechneten Beanstandungsquote führt. Dem wurde hier Rechnung getragen, indem für den Zeitraum Dezember 2013 bis März 2014 der hier zutreffende Mittelwert angesetzt wurde. Im mittleren Bereich passt die Prognose recht gut, wobei sich einzelne Ausreißer durch wochenindividuelle Einflüsse, wie z.B. auffällig hohe Beanstandungszahlen in einem bestimmten Zeitraum erklären lassen. Im letzten Drittel der jeweiligen Prognosen ist erneut der Einfluss des Erstzulassungsverzugs zu erkennen. Durch dessen Einfluss sinkt der Anteil der zugelassenen Fahrzeuge im Feld und folglich auch die Beanstandungen bzw. kumulierten Kilometer. Der Effekt der Extrapolation der Beanstandungen (vgl. Gleichung (3)) nimmt zu, je größer die Differenz zwischen mittlerer Laufleistung pro Fahrzeug (Tachostand) und Ziellaufleistung ist.

Ein Vergleich der jeweiligen Prognosen zum realen cph-Wert ist in Tabelle 1 für die FBR und Tabelle 2 für die MBR zusammengefasst. Betrachtet wird für den Wertevergleich zwischen den Prognosen verschiedener Zeitpunkte der jeweils mittlere Bereich des cph-Graphen, aus Gründen wie in Kapitel 2.4 hervorgehoben und in [1] benannt. Für die einzelnen Prognosen werden für diesen Vergleich deswegen 2 Monate (= 2 Prognosepunkte) vor Prognosedatum (16.06.2014, 15.09.2014 und 15.12.2014) sowie einheitlich der erste Monat (= 1 Prognosepunkt) nach Produktionsbeginn wegen der anfangs für eine Prognose unzureichenden Datengrundlage ausgeblendet.

Die Verläufe der drei Prognosen sind grundsätzlich gleichgerichtet (vgl. Bild 4 und Bild 5), sie weisen untereinander fast immer dieselben Knicke auf, folgen dabei prinzipiell den Verläufen der realen cph-Kurven der verschiedenen Zeitpunkte und nähern sich mit aktuellerem Prognosezeitpunkt dem Verlauf der realen Beanstandungsquote an. Dies zeigt auch der Wertevergleich in Tabelle 1 und Tabelle 2, vgl. Mittelwert mit Standardabweichung. Die Mittelwerte der Abweichungen errechnen sich als Durchschnitt der absoluten Differenz zwischen Prognosewerten und realen Werten, jeweils bezogen auf den realen Wert.

Für FBR und MBR verbessert sich die Übereinstimmung erwartungsgemäß, je später der Prognosezeitpunkt (Ausnahme: FBR, 13 Monate). Gemittelt über alle Prognosezeitpunkte verglichen gegen die realen cph-Werte bis heute, beträgt der Mittelwert der Abweichung bei FBR 25,9% und bei MBR 31,2%, die Standardabweichung bei FBR 15,4% und bei MBR 21,4%.

Tabelle 1: Wertevergleich der 12MiS-Prognosen, FBR

FBR		reale cph nach 13 Monaten		reale cph nach 16 Monaten		reale cph bis heute	
Prognosezeitpunkt ↓	Betrachteter Zeitraum	Mittelwert der Abweichung	Standardabweichung	Mittelwert der Abweichung	Standardabweichung	Mittelwert der Abweichung	Standardabweichung
16.06. 2014	Jan. 2014 - April 2014	13,6%	12,0%	27,8%	12,1%	38,8%	14,1%
15.09. 2014	Jan. 2014 - Juli 2014	20,7%	15,5%	20,1%	15,6%	26,0%	18,2%
15.12. 2014	Jan. 2014 - Okt. 2014	23,2%	14,6%	19,0%	11,3%	23,3%	10,8%
Gemittelt		19,6%	14,9%	20,1%	13,4%	25,9%	15,4%

Tabelle 2: Wertevergleich der 12MiS-Prognosen, MBR

MBR		reale cph nach 13 Monaten		reale cph nach 16 Monaten		reale cph bis heute	
Prognosezeitpunkt ↓	Betrachteter Zeitraum	Mittelwert der Abweichung	Standardabweichung	Mittelwert der Abweichung	Standardabweichung	Mittelwert der Abweichung	Standardabweichung
16.06. 2014	Jan. 2014 - April 2014	26,2%	9,5%	33,8%	11,9%	40,4%	20,2%
15.09. 2014	Jan. 2014 - Juli 2014	24,6%	15,8%	33,1%	18,8%	39,5%	21,4%
15.12. 2014	Jan. 2014 – Okt. 2014	15,8%	9,8%	19,2%	15,2%	24,8%	18,9%
Gemittelt		19,8%	13,0%	25,4%	17,5%	31,2%	21,4%

4. Zusammenfassung und Ausblick

Die prinzipielle Anwendbarkeit des Prognoseverfahrens bzgl. Beanstandungen in frühen Phasen des Feldeinsatzes wurde dargelegt. Eine Validierung des Verfahrens wurde durch eine retrospektive Auswertung anhand zweier Projekte (Fahrzeug- und Motorbaureihe) für jeweils 3 Prognosezeitpunkte während der Anlaufphase der Produktion durchgeführt. Dabei wurde festgestellt, dass die Prognose im Mittel als zu positiv verglichen mit dem realen Beanstandungswert ermittelt wurde. Die Prognosegüte steigt mit zunehmender Datengrundlage, also mit späterem Prognosezeitpunkt. Für den Vergleich der Prognose mit der realen cph-Kurve nach 13 Monaten zeigt sich in diesem Anwendungsfall die beste Übereinstimmung.

Die Formel zur Prognose der cph ist sensibel bezüglich ihrer Eingangsgrößen, weshalb sie möglichst genauer Eingangsparameter bedarf. Die Dynamik des Anlaufs erschwert dies, durch

besonders starke Abweichungen zu angenommenen Eingangsparametern bezüglich Erstzulassung und Laufleistung.

Die weitere Anpassung des hier dargestellten Beanstandungsprognose-Verfahrens sollte sich mit der weitergeführten Präzisierung befassen. Insbesondere könnten die Einflüsse der Verteilungsermittlung und Simulation der Zufallszahlen, der anlaufspezifischen Eingangsparameter oder des belastungsabhängigen Ausfallverhaltens, z.B. über den Formparameter der Weibull-Verteilung, untersucht werden.

Liegen Erfahrungswerte aus mehreren retrospektiv angewandten Prognosen vor, lässt sich der Unterschied zwischen Prognose und realem cph-Wert antizipieren und bereits in der Prognose berücksichtigen, etwa durch Korrekturfaktoren.

Literaturangaben

- [1] Jordan, C., Jacobi, A. u. Köttermann, T.: Zuverlässigkeitsprognose (12 Months in Service) im Anlauf eines Neuproduktprojekts. Schneller die Qualitätssituation bewerten. In: 27. VDI-Fachtagung Technische Zuverlässigkeit 2015. Entwicklung und Betrieb zuverlässiger Produkte ; Leonberg bei Stuttgart, 20. und 21. Mai 2015. Verein Deutscher Ingenieure: VDI-Berichte, Bd. 2260. Düsseldorf: VDI-Verl. 2015, S. 257–268
- [2] Jacobi, A., Jordan, C. u. Kaganova, E.: Zuverlässigkeitsmanagement für den Anlauf von Neuproduktprojekten in der Automobilindustrie. In: 28. VDI-Fachtagung Technische Zuverlässigkeit 2017. Entwicklung und Betrieb zuverlässiger Produkte ; Leonberg bei Stuttgart, 17. und 18. Mai 2017. Verein Deutscher Ingenieure: VDI-Berichte, Bd. 2307. Düsseldorf: VDI-Verl. 2017
- [3] Abernethy, R. B.: The new Weibull handbook. North Palm Beach, Florida: R.B. Abernethy 2006
- [4] Verband der Automobilindustrie e.V.: Zuverlässigkeitssicherung bei Automobilherstellern und Lieferanten. Zuverlässigkeits-Methoden und -Hilfsmittel, Bd. 3.2. Berlin: Henrich Druck + Medien GmbH 2016

Einsatz textiler Sensoren für die Lebensdauervorhersage von Maschinen



Smart Maintenance am Beispiel einer Fünfachfräsmaschine mit Hilfe textiler Fasern als Sensoren

Dr.-Ing. **Christian Döbel**, TITK e. V., Rudolstadt

Kurzfassung

Mit Hilfe textiler, piezoelektrischer Sensorfasern wurde der Verschleißzustand einer Fünfachfräsmaschine erkannt. Die Sensoren erfassen dabei den aktuellen Körperschall im Material.

Abstract (optional)

The state of wear of an five-axis milling machine was detected with the help of textile, piezo-electric sensor fibers. The sensors detect the current structure-borne sound inside the material.

1. Motivation

Nicht erkannter Maschinenverschleiß oder unvorhergesehene Ereignisse führen häufig zu nicht mehr beeinflussbaren Stillstandszeiten von Anlagen in einer Fertigung. Wenn daraufhin die die gesamte Fertigung zu einem ungünstigen und nicht planbaren Zeitpunkt zum Erliegen kommt, entstehen nicht unerhebliche vermeidbare finanzielle Schäden.

Deshalb wurden im Rahmen eines Projekts des TITK e. V. gemeinsam mit einem Bearbeitungszentrum textile Sensorfasern dazu benutzt, den aktuellen Verschleißzustand der Hauptspindel in einer Fünfachfräsmaschine zu messen. Die textilen Piezosensoren haben den Vorteil, sehr flexibel hinsichtlich Form und Länge zu sein.

In der mittelständischen Firma sind 31 Fünfachfräsmaschinen (Baujahr 1997 ... 2015) im Einsatz, von denen jede etwa alle vier Jahre wegen einer defekten Spindel ausfällt. Die durchschnittliche Stillstandszeit nach einem solchen Ausfall beträgt 10 Arbeitstage. Ziel der Verschleißerkennung ist es deshalb, Maschinenausfälle vorherzusagen, um den Gesamtprozess besser planen zu können. Damit könnten intelligenter geplante Wartungen zu deutlich

mehr Liefertreue führen, wenn beispielsweise Betriebsmittel vorsorglich ausgetauscht werden können und damit ein Stillstand des gesamten Fertigungsprozesses vermieden wird.

2. Textile Piezosensoren

Textile Piezosensoren bieten eine Kombination aus mechanischen und funktionalen Eigenschaften, die sich mit anderen Sensoren nicht darstellen lassen. Mechanisch sind sie einerseits sehr flexibel und eignen sich durch die freie Formbarkeit auch zum Einbau in bestehende Systeme, andererseits ermöglicht deren Endlosfertigung eine wirtschaftliche Herstellung. Funktional ist mit der Möglichkeit, sehr lange Sensoren herzustellen, eine integrale Messung von Schwingungen über lange Wege hinweg möglich, außerdem können die Sensoren sehr breitbandig Signale messen (unterer Hz- bis MHz-Bereich).

Die am TITK entwickelten textilen Piezosensoren sind schichtweise aufgebaut, wie in Bild 1 gezeigt ist. Die innere Schicht (Innenleiter) ist das eigentliche Strukturelement des Funktionsmaterials und hat die Aufgabe, die Reißfestigkeit des Sensors von 8.9 cN/tex (nach DIN EN ISO 2062) zu gewährleisten [1]. Die zweite Funktion des Innenleiters ist es, den elektrischen Messstrom zu tragen, der in den Sensor hinein fließt.

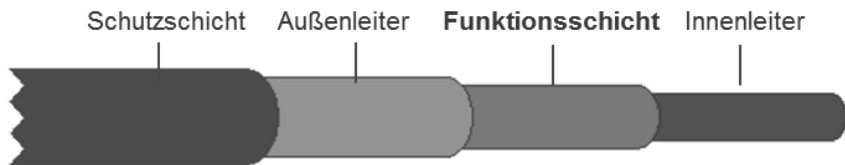


Bild 1: Aufbau eines textilen Sensors.

Am TITK wird der Innenleiter typischerweise aus Polyethylen (PE) als Matrixpolymer hergestellt, das mit leitfähigen Partikeln (typischerweise Polyblak) zu dessen Funktionalisierung gefüllt ist. Damit ist diese Werkstoffkombination nach dem Schmelzspinnen textil verstreckbar, um die Reißfestigkeit zu gewährleisten, zum anderen bilden die Füllpartikel elektrische Strompfade aus, um das Signal zu führen. Der Widerstand liegt im Bereich mehrerer MOhm pro Meter, was für die Anwendung als Sensor ausreichend ist (bei einer Kapazität von

$C_0 = 2.6 \frac{pF}{mm}$). Der Innenleiter hat einen Durchmesser von etwa 620 μm .

Die Funktionsschicht besteht aus Polyvinylidenfluorid (PVDF). Dieses piezoelektrische Polymer wird ca. 40 μm dick aufgebaut. Infolge einer äußeren Anregung des polarisierten PVDF kommt es zur Gitterverschiebung des Werkstoffs und damit zur Verschiebung der La-

derungsschwerpunkte. Daraufhin wird eine elektrische Spannung zwischen den Grenzschichten zum Innenleiter sowie zum Außenleiter erzeugt, solange die äußere mechanische Anregung, etwa das Auftreffen einer Körperschallwelle, vorliegt.

Der Außenleiter ist eine nur ca. 0.1 µm dicke Aluminiumschicht, die aufgedampft wird und den Stromkreis zur elektrischen Signableitung schließt, während als Schutzschicht meist ein Lacksystem eingesetzt wird, um eine elektrische Isolation des Sensors zu gewährleisten.

Der Aufbau des gesamten Sensors mit den Anschlusskontakten ist in Bild 2 mit typischen Abmessungen gezeigt. Die Kontaktierung wird derzeit mit Silberleitlack realisiert, während weitere Möglichkeiten bereits in der Erprobung sind.

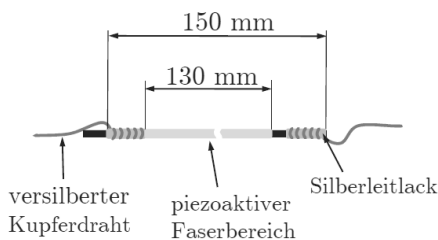


Bild 2: Aufbau gesamter Piezosensor [2].

3. Möglichkeiten der Verschleißerkennung an Fräsmaschinen

Verschleißerkennung an Fräsmaschinen wird heutzutage üblicherweise durch die Messung der Stromaufnahme des Antriebs (elektrisch) realisiert. Alternativ können die mechanischen Größen Drehzahl, Drehmoment und Kraft des Antriebs gemessen werden, um Anomalien gegenüber einem Referenzzustand festzustellen. Auch die Ermittlung von Druck (z. B. Anpressdruck des Werkzeugs ans Werkstück) oder die Prozesstemperatur geben Aufschluss über einen möglichen Verschleiß [3].

Eine weitere Möglichkeit, Materialermüdungen zu erkennen, ist der Körperschall. Dieser wird durch die Spindelrotation verursacht und durch die Reibung zwischen dem Werkzeug (beispielsweise einem Schneidwerkzeug) und dem Werkstück verändert [4]. Das übertragene Spektrum ist dabei abhängig von der Geometrie der Körper, die sich berühren, sowie des Materialzustands [5]. Dieser Zusammenhang bildet die Grundlage des Projekts.

Möglichkeiten zur Messung des Körperschalls sind Beschleunigungsaufnehmer oder Klopfensensoren sowie MEMS mit integrierter Signalauswertung. Diese Systeme jedoch lassen

sich nicht direkt in Werkstoffe integrieren, wo die Körperschallwelle die größte Informationsdichte besitzt und messen zudem nur punktuell.

Deshalb werden im vorliegenden Projekt Textilfasern zur integralen Messung des Körperschalls über lange Entfernungen (bis zu mehreren Metern) eingesetzt. Integral bedeutet dabei, dass die Faser die Summe der mechanischen Belastungen, die die Faser über die gesamte Länge erfährt, in elektrische Spannung umwandelt [6].

Um mit Hilfe des Körperschalls eine Aussage über den Verschleißzustand des Werkzeugs treffen zu können, darf dieser nur vom Werkzeugverschleißzustand verändert werden, indem alle anderen Einflüsse definiert sind. In Bild 3 sind typische Änderungen des Körperschallspektrums in einem bestimmten Zeitraum (4 ... 10 s) gezeigt, der im vorgestellten Projekt einen Ausschnitt des Verschleißzeitraums symbolisiert.

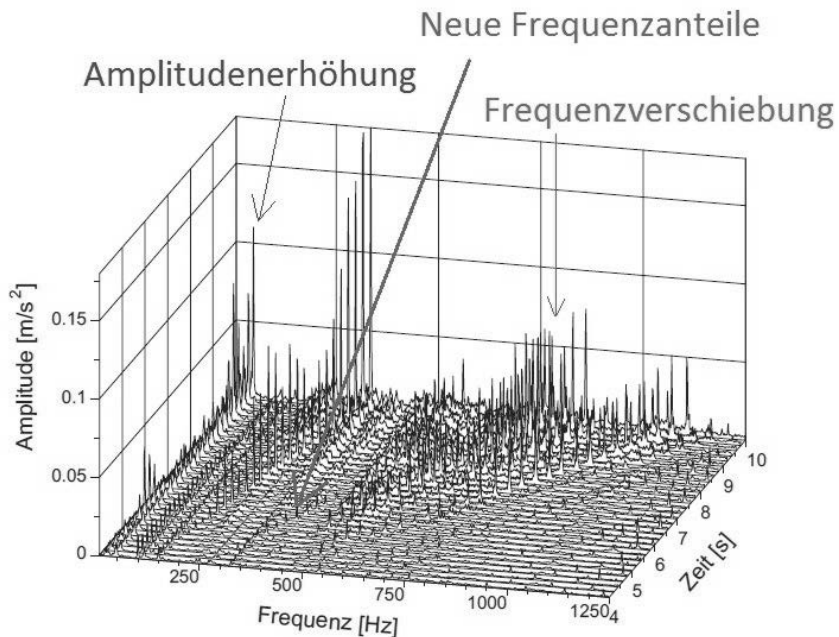


Bild 3: Wasserfallanalyse zeigt die Veränderung des Körperschallspektrums [7].

Zum einen können sich infolge des Werkzeugverschleißes zusätzliche Resonanzfrequenzen herausbilden (beispielsweise Oberwellen infolge von Mikrorissen), zum anderen auch

Resonanzfrequenzen verschieben. Das ist etwa der Fall, wenn sich die Festigkeit des Werkzeugs ändert (z. B. durch Delamination). Auch können Energieanteile in einem bestimmten Frequenzbereich wachsen, was sich durch eine Amplitudenerhöhung in Bild 3 darstellt. Ziel der Untersuchungen war es, diese Artefakte im Spektrum mit Hilfe einer Funktionsfaser zu identifizieren und bestimmten Verschleißerscheinungen des Werkzeugs zuzuordnen [8].

4. Umsetzung im Projekt

Im Rahmen des Projekts wurden für die spätere Integration der Sensorfasern in Werkstoffe verschiedene, adhäsive Möglichkeiten zur elektrischen Kontaktierung untersucht, indem die Piezofasern an silberbeschichtete Kupferleitungen zur Signalauskopplung (Durchmesser: 70 μm) geklebt wurden. Die Verbindungen wurden mechanisch sowie elektrisch charakterisiert, wobei sich der Silberleitlack G3692 (Plano GmbH) als beste Variante erwies (mechanische Spannung der Kontaktstelle: 0.54 cN/tex bei einem elektrischen Übergangswiderstand von 45 k Ω im Bereich 0.1 ... 10 kHz). In Bild 4 sind die Kontaktstellen der Piezofaser dargestellt.

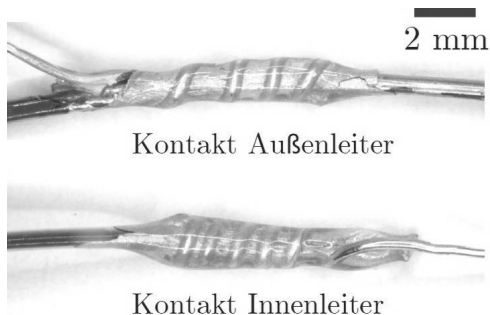


Bild 4: Kontaktstellen des Außen- sowie des Innenleiters [2].

Danach wurde eine polymere Piezofaser als Sensor am Spindelstock einer Fünffachsfräsmaschine angebracht mit dem Ziel, Mikrobewegungen an der Oberfläche des Spindelstocks zu messen. Die formschlüssige Verbindung wurde mit Hilfe eines Aralditharz-Systems erreicht. Aufgrund des Handlings wurde die Faser kraftschlüssig auf die Oberfläche geklebt, womit der Oberflächenschall als Indikator für den Körperschall gemessen werden konnte¹. Die

¹ In einem nächsten Schritt ist die Integration geplant, um tatsächlich den Körperschall aufzunehmen.

Dämpfung des Klebers kann aus der Erfahrung vorangegangener Untersuchungen vernachlässigt werden, die aufgeklebte Faser ist in **Bild 5** gezeigt.



Bild 5: Eingespanntes Werkzeug mit Spindelstock und aufgeklebter Piezofaser.

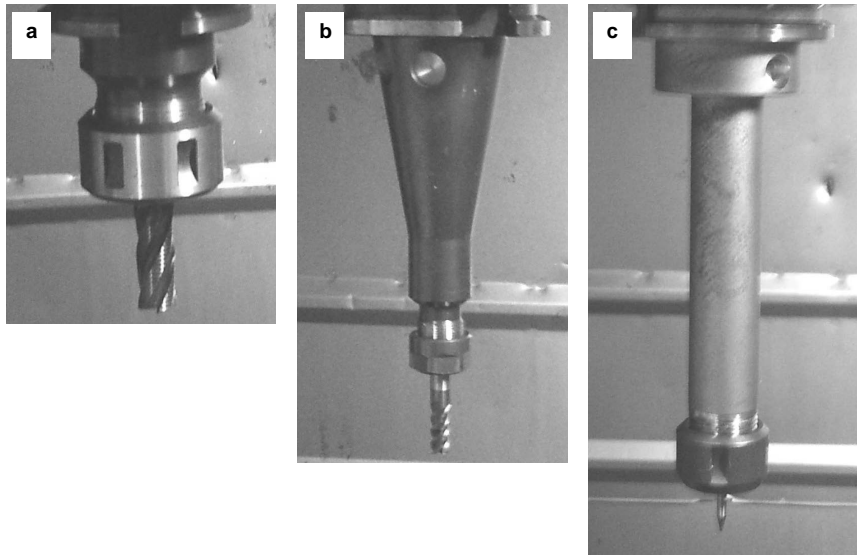


Bild 6: Untersuchungen mit drei verschiedenen Werkzeugen
(a: kurzes, b: mittleres und c: langes Werkzeug).

Die Prüflinge selbst sind eine verschlissene (kurz vor dem Auswechseln) sowie eine neue Spindel. Mit beiden wurden Versuche nach dem gleichen Versuchsplan durchgeführt. Der Versuchsplan umfasste die Bearbeitung eines definierten Stahlquaders („Standardkörper“). Dabei wurden drei Werkzeuge (kurz, mittel, lang, s. Bild 6), drei Rotationsgeschwindigkeiten (500, 6 000, 12 000 U/min), zwei verschiedene Fräsprogramme (Standard, Leerlauf) sowie zwei Rotationsrichtungen (Linkslauf, Rechtslauf) miteinander kombiniert.

Vor dem Spindeltausch wurden mit der verschlissenen Spindel insgesamt 13 Versuche durchgeführt und dabei die ausgekoppelten Körperschallspektren gemessen. Ein Ausschnitt aus einer Messung ist in Bild 7 gezeigt, bei der drei verschiedene Rotationsgeschwindigkeiten miteinander verglichen werden. Die Amplituden sind sehr gering, da sich zwischen dem Antrieb als Energiequelle und dem Gehäuse (Spindelstock) ein sehr guter Dämpfungskörper (als Energiesenke) befindet. In den Messungen sind jeweils die charakteristischen Frequenzen (Drehzahl, Oberwellen) für die Rotationsgeschwindigkeiten von 500, 6 000 und 12 000 U/min sichtbar.

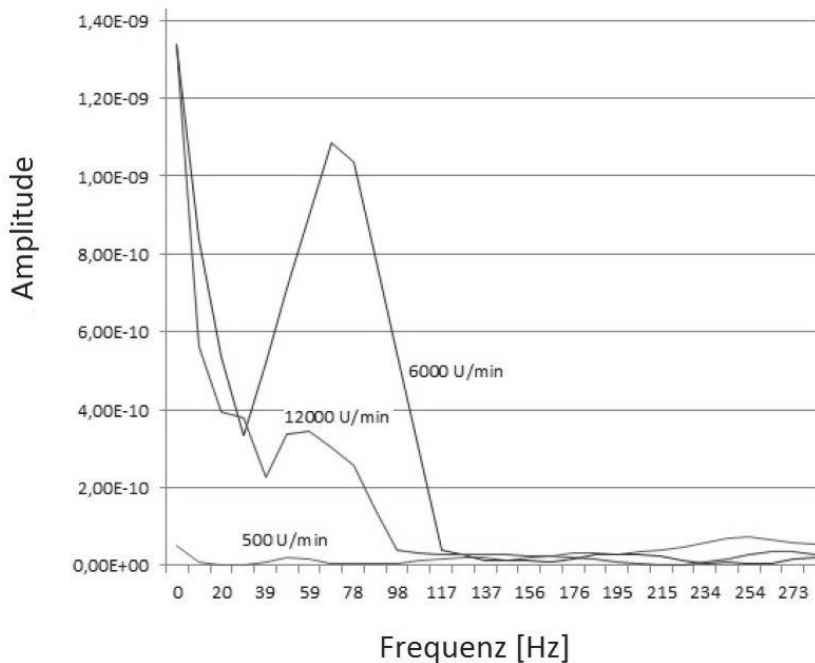


Bild 7: Fouriertransformation bei einem Arbeitspunkt (kurzes Werkzeug).

Anschließend wurden nach dem Spindeltausch die gleichen Messungen anhand des Versuchsplans mit einer neuen Spindel (Prüfling 2) wiederholt und die Frequenzspektren mit denen des ersten Prüflings miteinander verglichen. Daraus ergaben sich insgesamt acht potentielle Merkmale, die zunächst eine Gut-Schlecht-Unterscheidung zwischen Verschleiß und Neuzustand erlauben. Am Beispiel der Signalleistung als potentielles Merkmal sind eine verschlissene und eine neue Spindel in Bild 8 und Bild 9 gegenübergestellt. Die Randbedingungen sind jeweils 500 U/min im Rechtslauf und ein mittleres Werkzeug.

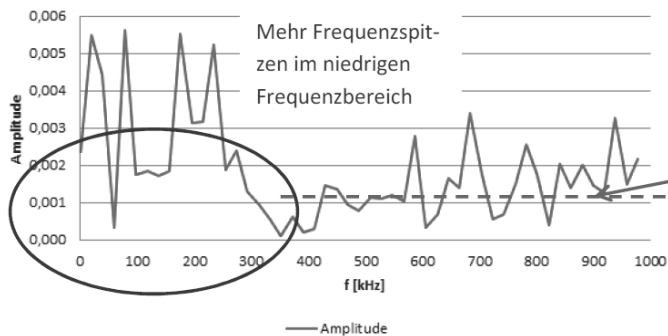


Bild 8: Spektrum einer verschlissenen Spindel (Signalleistung s. braune, gestrichelte Linie).

Einerseits ist hier feststellbar, dass der Trend der mittleren Signalleistung beim verschlissenen Werkzeug eher konstant bleibt (Anstieg beim neuen Werkzeug). Andererseits sind mehr Resonanzfrequenzen im unseren Frequenzbereich im verschlissenen Zustand sichtbar.

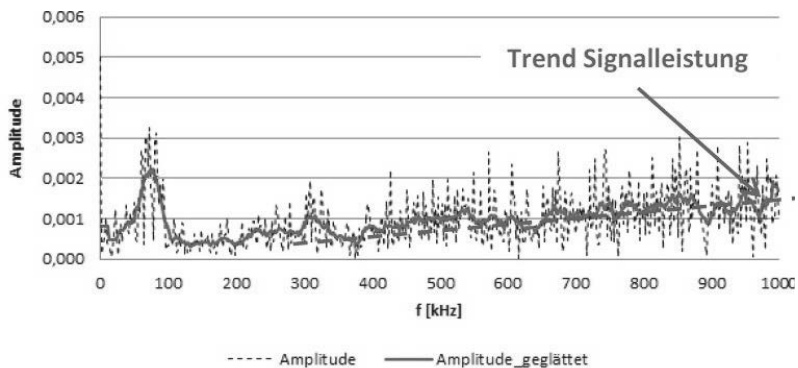


Bild 9: Spektrum einer neuen Spindel.

Um aus den potentiellen Merkmalen im Signal zur Unterscheidung einer verschlissenen von einer neuen Spindel die signifikante Merkmale herauszufiltern, wurden anschließend zwei Schritte unternommen:

- ➔ Die Merkmale wurden auf physikalische Plausibilität überprüft und
- ➔ Die Messungen wurden gezielt auf signifikante Unterschiede zwischen dem Gut- sowie dem Verschleißzustand hinsichtlich des jeweils potentiellen Merkmals untersucht.

Eine solche Untersuchung ist für typische Messungen in Bild 10 dargestellt. Die drei Messungen mit einer neuen Spindel (Kurven NEU_1 ... NEU_3) stark beieinander, am Beispiel der alten Spindel ist mit ALT_2 ein typischer und mit ALT_1 ein atypischer Verlauf der Kurve gezeigt.

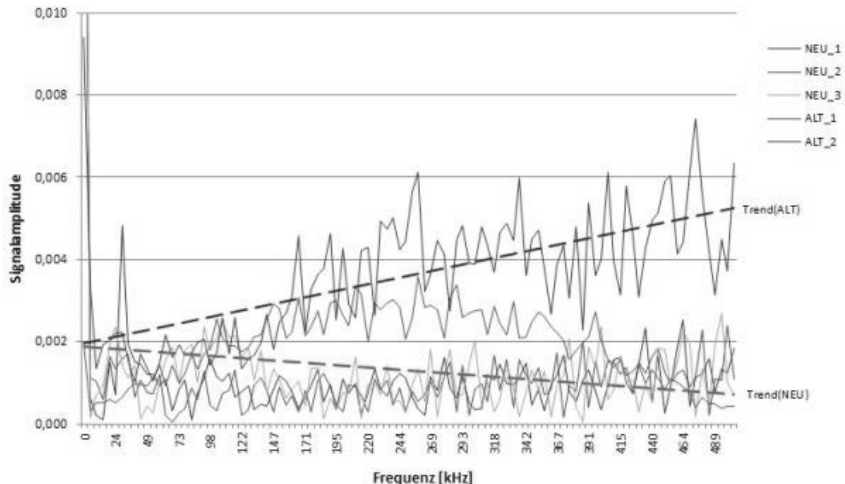


Bild 10: Trends der Signalamplituden im niederfrequenten Bereich einer neuen (rot gestrichelt) und einer verschlissenen (blau gestrichelt) Spindel im Vergleich.

Stehen hinreichend viele und voneinander unabhängige Messungen zur statistischen Auswertung der Trendgeraden zur Verfügung, kann die Trennschärfe zwischen der neuen und denen der verschlissenen Spindel statistisch quantifiziert werden, indem eine Funktion existiert, die einen verschlissenen von einem neuen Prüfling unterscheidet. Im vorliegenden Projekt wurden für die alte und die verschlissene Spindel jeweils 25 Messungen ausgewertet, um eine statistische Aussage (Mittelwert und Streuung der Parameter) abzuleiten.

Nur wenn die Trennschärfe groß genug ist, lassen sich gute und verschlissene Spindeln eindeutig anhand dieses Signalmerkmals klassifizieren, was bei diesem Merkmal der Fall ist. In Bild 11 sind die Konfidenzbereiche des Signalmerkmals gegenübergestellt. Hier ist bereits eine eindeutige (lineare) Unterscheidung zwischen einer guten und einer schlechten Spindel anhand dieses Signalmerkmals möglich ist.

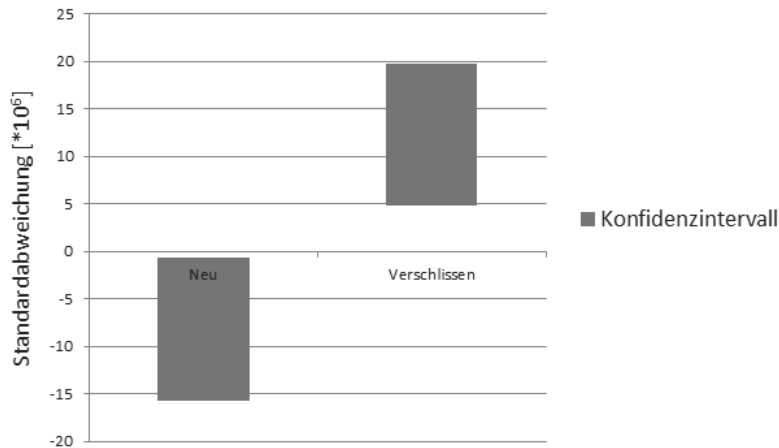


Bild 11: Vergleich der Konfidenzbereiche der neuen und der verschlissenen Spindel im Vergleich für das Signalmerkmal „Trend der Signalamplitude“. Anhand der Stichprobe ist eine eindeutige Klassifizierung zwischen „Neu“ und „Verschlissen“ möglich.

Außerdem ist der Trend der Leistungsdichte im niederfrequenten Bereich auch physikalisch erklärbar. Aus diesem Grund wird aus dem potentiellen ein signifikantes Merkmal im Sinne der Unterscheidung zwischen einer verschlissenen und einer neuen Spindel. Ziel ist es, hinreichend viele signifikante Merkmale zu identifizieren, um durch ihre Kombination eine sichere Vorhersage des Verschleißzustands zu ermöglichen. Insgesamt wurden vier signifikante Signalmerkmale gefunden:

- ➔ Messung der anteiligen Signalleistung im Bereich 500 Hz ... 500 kHz im Vergleich zur gesamten Signalleistung bis 500 kHz
- ➔ Trend der Signalamplitude bis 500 kHz
- ➔ Abnahme der Signalamplitude im niederfrequenten Bereich
- ➔ Messung der Anstiegszeiten an den Signalfanken im Zeitsignal

Da die Messungen an der Spindel nur vier Wochen auseinanderlagen und der Spindelstock samt Sensor nicht getauscht wurde, ist der Verschleiß des Sensors selbst vernachlässigbar. In mehreren vorangegangenen Untersuchungen wurden Sensoren über einen deutlich längeren Zeitraum hinweg im Klimaschrank gelagert und erzeugten bei gleicher Anregung vor

und nach der Beaufschlagung identische Signalamplitude, woraus auf eine lange Verschleißzeit geschlossen wird.

5. Ausblick: Überwachung per App/ Vorteile der Smart Factory

Im nächsten Schritt wurde eine App zur Visualisierung der Daten programmiert. Ziel ist es, den Verschleißzustand einerseits sicher zu erkennen und andererseits dem Maschinenbediener direkt auf dem Smartphone anzuzeigen. In Bild 12 findet sich die Oberfläche mit der Auswertung des aktuellen Signalspektrums im Vergleich zu einem Referenzspektrum zu einem definierten Zeitpunkt, an dem von einer nicht verschlissenen Spindel ausgegangen werden kann.

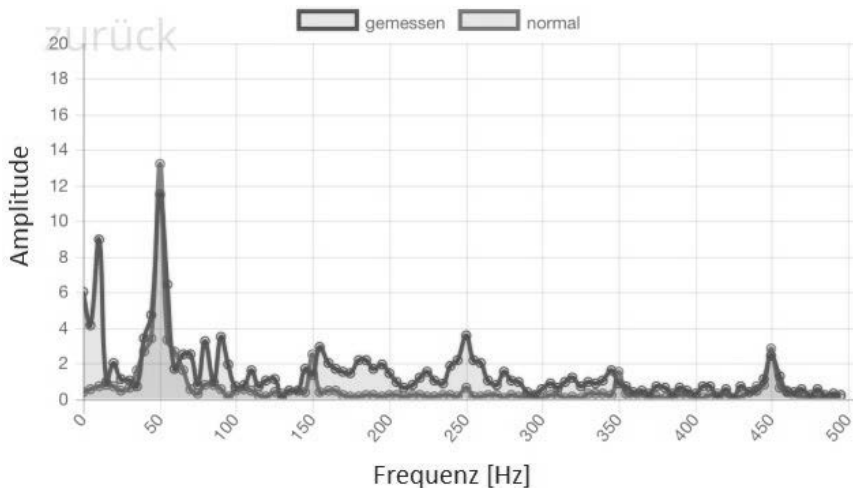


Bild 12: Diskrete Fouriertransformation zweier Signale der Piezofaser im Vergleich visualisiert in einer Smartphone-App [9]. Die graue Kurve ist das Referenzsignal bei einem definierten Gutzustand (nach einer Kalibrierung der App), die grüne Kurve zeigt die aktuelle Messung.

Das ausgegebene Signal in Bild 12 ist als Dummysignal in einen Mikrocontroller implementiert worden, da die Anbindung an die Piezofaser noch nicht mit dem Mikrocontroller verbunden wurde. Allerdings ist bereits eine Datenübertragung vom Controller, an den die Piezofaser angeschlossen ist, an die App möglich. Auch mehrere Verschleißstadien wurden dabei simuliert, die die App korrekt widerspiegelt.

6. Fazit

Im Rahmen des Projekts wurden textile Sensorfasern dazu eingesetzt, den Verschleiß eines Werkzeugs einer Fünffachsfräsmaschine zu erkennen und zu quantifizieren. Dabei wurde der Körperschall, der die Faser erreicht, permanent breitbandig gemessen und (piezoelektrisch) in elektrische Energie gewandelt, bevor das Körperschallspektrum mit Hilfe der (Fast) Fouriertransformation berechnet wurde. Es konnten im Signalspektrum dabei Merkmale identifiziert werden, die den Gutzustand vom Verschleißzustand unterscheiden konnten.

Aufgrund der geringen Anzahl der Messwerte ist eine hinreichende Aussage bezüglich Signifikanz der Signale nicht möglich, sodass nur qualitative Aussagen getätigt werden konnten. Auch wurde die Korrelation zwischen den Signalen und bestimmten Fehlerbildern, etwa dem Spindelbruch, nicht hergestellt. Dazu sind weitere Messungen nötig.

Der Vorteil der Nutzung von Funktionsfasern liegt in der freien Formgebung der Sensorik. Insbesondere bei der Nachrüstung der Sensoren in bestimmte Systeme ist es entscheidend, dass sich die zusätzlichen Sensoren an die Gegebenheiten bestehender Systeme anpassen. Das war im vorliegenden Projekt ohne weiteres möglich. Dem wachsenden Bedarf an der Signalauskopplung sowie -verarbeitung wurde durch die Programmierung einer App zur Visualisierung Rechnung getragen.

Neben der technologischen Weiterentwicklung der textilen Funktionsfasern stehen alle Schritte der Signalerhebung und -interpretation im Rahmen von Nachfolgeprojekten im Vordergrund, um eine vorausschauende Verschleißerkennung zu ermöglichen.

Literaturnachweise:

- [1] TITK, Technische Produktinformation Faserförmige Piezosensoren/ Energiewandler, www.titk.de, Aufruf vom 17.2.2017
- [2] Ritter, F., Diplomarbeit „Faserbasierte Sensierung verschiedener Prozess- und Funktionsparameter von Faserverbundbauteilen für Industrie 4.0“, 2016.
- [3] Wirth, R., Maschinendiagnose an Industriegetrieben, 1998.
- [4] Scheer, C., Reimche, W., Bach, F.-W., Schallemissions- und Waveletanalysen zur frühzeitigen Schadenserkenkung an hochbelasteten rotierenden Bauteilen, In: DGZfP-Jahrestagung, Hannover, 2007.
- [5] Heinrich, M., Rabe, U., Grabowski, B., Valeske, B., Methodische Entwicklung der akustischen Resonanzanalyse zur zerstörungsfreien Erkennung von ur- oder umgeformten Serienteilen mit unzulässigen Geometrieabweichungen – simulationsbasierte Modellüberlegungen, DACH-Jahrestagung, 2015.
- [6] RTE Akustik und Prüftechnik GmbH, Messung und Bewertung der Eigenfrequenzen von Komponenten und Baugruppen in Fertigung und Labor.
- [7] Deckers, J., Dissertation: Entwicklung einer Low-Cost-Körperschallsensorik zur Überwachung des Verschleißverhaltens von wälz- und gleitgelagerten Kreiselumpen kleiner Leistung, 2001.
- [8] Döbel, C., Schadensanalyse in Rotorblättern mit Hilfe textiler, sensorischer Fasern, VDI-Konferenz Schwingungen und Dynamik von Windkraftanlagen, Vortrag und Manuskript im Tagungsband, Bremen, 2016.
- [9] Engelmann, V, Bachelorarbeit: HTML5 App zur Messdatenerfassung einer Piezo-Faser, TU Ilmenau, 2016.

Bootstrap-Monte-Carlo-Simulation von Zuverlässigkeit und Aussagewahrscheinlichkeit bei periodischer Instandhaltung



Bootstrap Monte Carlo Simulation of Reliability with Confidence Level regarding Periodical Maintenance

Frank Müller, M.Sc.,
Institut für Maschinenelemente, Universität Stuttgart;
Prof. Dr.-Ing. **Peter Zeiler**,
Hochschule Esslingen, Fakultät Mechatronik und Elektrotechnik;
Prof. Dr.-Ing. **Bernd Bertsche**,
Institut für Maschinenelemente, Universität Stuttgart

Kurzfassung

In diesem Beitrag wird zunächst die Bootstrap-Monte-Carlo-Simulation (BMCS) als neue Methode, mit welcher erstmals ein Vertrauensbereich der Zuverlässigkeit mit periodischer Instandhaltung für allgemeine technische Systeme angegeben werden kann, vorgestellt. Die Methode kann basierend auf einer Stichprobe aus Lebensdauerdaten oder einer Lebensdauerverteilung mit Angabe des zugrundeliegenden Stichprobenumfangs durchgeführt werden. Sie besitzt keine Einschränkungen hinsichtlich der Verteilungsart und erlaubt sowohl den Zuverlässigkeitsnachweis als auch die -prognose.

Zur Methodenverifikation wird zunächst das Bootstrap-Verfahren zur Berechnung eines Vertrauensbereichs einer Verteilungsfunktion untersucht und mit betaverteilten Vertrauensbereichen verglichen. Schließlich wird die Genauigkeit der über die BMCS bestimmten Vertrauensbereiche der Zuverlässigkeit einer periodisch instandgehaltenen Komponente analysiert und anhand eines Vergleichs mit einem mittels der Momentenmethode bestimmten Vertrauensbereich quantitativ bewertet. Abschließend wird eine Parameterstudie der Eingangsgrößen der BMCS durchgeführt sowie ferner ein Anwendungsszenario betrachtet.

Abstract

In this paper, the bootstrap Monte Carlo simulation (BMCS) as a new method for the determination of reliability with confidence level for general technical systems with periodical measures is presented. The method is either based on the pure samples of failure times or the mean value distributions in combination with the corresponding sample size of reliability. It does not require special distribution types and allows the reliability prediction as well as the

reliability demonstration. Besides the verification of the new method and a parameter study, a case study is presented within this paper.

1. Einleitung

Wichtigster Aspekt zur Beschreibung der Qualität eines Produktes ist dessen Zuverlässigkeit. Um diese zu beschreiben, werden Lebensdauerdaten betrachtet. Die Auswertung erfolgt dabei in der Regel über die Analyse einer mittelwertigen Verteilung, z.B. der Medianverteilung, kombiniert mit einem Vertrauensbereich. Dieser beschreibt die Wahrscheinlichkeit, mit welcher von der Stichprobe auf die Grundgesamtheit, d.h. auf die wahre Zuverlässigkeit der Produkte, geschlossen werden kann.

Mit bisher bekannten Methoden kann zu einer einzelnen Komponente die Zuverlässigkeit mit Vertrauensbereich bestimmt werden. Neben analytischen stehen approximative und simulationsbasierte Verfahren zur Verfügung [8]. Werden Systeme aus mehreren Teilsystemen bzw. Komponenten betrachtet, so kann mit exakten oder approximativen Methoden, wie der Momentenmethode, die Systemzuverlässigkeit mit Vertrauensbereich berechnet werden [11]. Ferner kann die Stichprobengröße bestimmt werden, mit welcher eine gewisse Zuverlässigkeit mit vorgegebener Aussagewahrscheinlichkeit nachgewiesen werden kann [10].

Für allgemeine Systeme, die einem Instandhaltungsprozess unterliegen, gibt es bisher keine Methode, mit der auch für deren Zuverlässigkeit die Aussagewahrscheinlichkeit angegeben werden kann. Basierend auf der Verfügbarkeitsprognose mit Vertrauensbereich [9] und dem Nachweis der Verfügbarkeit mit Vertrauensbereich [6] kann mit der in diesem Beitrag vorgestellten Bootstrap-Monte-Carlo-Simulation (BMCS) als neue Methode die Zuverlässigkeit mit Aussagewahrscheinlichkeit eines Einzelsystems, welches periodischen Instandhaltungen unterworfen ist, erstmals berechnet werden.

Zunächst wird in diesem Beitrag die BMCS als neue Methode, mit welcher erstmalig die Prognose und der Nachweis der Zuverlässigkeit mit Aussagewahrscheinlichkeit für Systeme mit planmäßigen Instandhaltungen möglich sind, vorgestellt. Zur Methodenverifikation wird zum einen das Bootstrap-Verfahren zur Berechnung eines Vertrauensbereichs untersucht und mit betaverteilten Vertrauensbereichen verglichen. Zum anderen wird die Genauigkeit der BMCS anhand einer Komponente mit periodischer Instandhaltungsmaßnahmen verifiziert. Durch Vergleich mit mittels der Momentenmethode berechneten Vertrauensbereichen wird die BMCS quantitativ bewertet. Ferner wird anhand einer Parameterstudie der Einfluss der Eingangsparameter verdeutlicht.

2. Zuverlässigkeit bei periodischer Instandhaltung

Periodische Instandhaltungen werden häufig genutzt, um negative Einflüsse von Alterung oder Verschleiß zu minimieren. Zur Beschreibung des Systemverhaltens bei planmäßiger Erneuerung kann das periodische Instandhaltungsmodell genutzt werden [1].

Bild 1 verdeutlicht den Instandhaltungsplan in der zeitlichen Abfolge. Nach jeder, in konstanten Zeitabständen T_{PM} ($PM = \text{Preventive Maintenance}$) durchgeführten Maßnahme gilt die betrachtete Einheit als neuwertig. Die Dauer der Instandhaltung selbst wird vernachlässigt. Da das Ausfallverhalten vor und nach einer Instandhaltung als stochastisch unabhängig angenommen wird, ergibt sich die Zuverlässigkeitsfunktion einer gewarteten Einheit zu [1]:

$$R_{PM}(t) = R(T_{PM})^k \cdot R(t - k \cdot T_{PM}) \quad (1)$$

für $k \cdot T_{PM} \leq t \leq (k+1) \cdot T_{PM}$ und $k = 0(1)\infty$

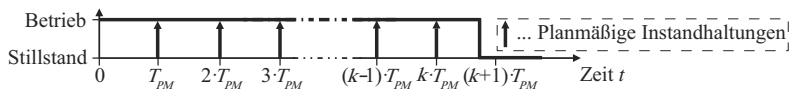


Bild 1: Instandhaltungsplan unter periodischen Maßnahmen in der zeitlichen Abfolge.

Fällt eine Komponente vor der nächsten periodischen Erneuerung aus, wird sie nicht weiter instandgehalten. Neben der Berechnung nach Gleichung 1 kann die Zuverlässigkeit bei periodischer Instandhaltung auch über eine Monte-Carlo-Simulation nach Bild 2 approximativ bestimmt werden.

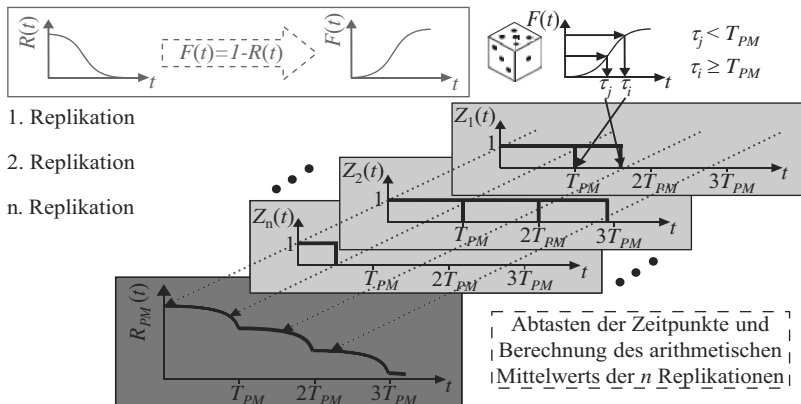


Bild 2: Monte-Carlo-Simulation der Zuverlässigkeit bei periodischer Instandhaltung [5].

Für jede Monte-Carlo-Replikation wird innerhalb des betrachteten Zeitraums ein einzelner Verlauf des Zustandsvektors $Z(t)$ gebildet, indem aus der Lebensdauerverteilung $F(t)$ zufälli-

ge Lebensdauern τ_i gezogen werden. Ist die Lebensdauer größer oder gleich dem Instandhaltungsintervall ($\tau_i \geq T_{PM}$), so wird dem Zustandsindikator bis zum Zeitpunkt der nächsten Instandhaltung der Wert 1 (funktionsfähig) zugeordnet. Andernfalls nimmt er für die zufällige Lebensdauer den Wert 1 und anschließend für den kompletten restlichen Betrachtungszeitraum den Wert 0 (ausgefallen) an. D.h. nach einem Ausfall ist die Komponente auch in den folgenden Instandhaltungsintervallen nicht funktionsfähig. Nachdem für jede Monte-Carlo-Replikation ein Zustandsverlauf gefunden wurde, kann die Zuverlässigkeit bei periodischer Instandhaltung $R_{PM}(t)$ für jeden Zeitpunkt t_j des Betrachtungszeitraums als arithmetischer Mittelwert des Wertes aller Zustandsindikatoren abgeleitet werden.

3. Aussagewahrscheinlichkeit und Vertrauensbereich

Die Aussagewahrscheinlichkeit P_A zur Zuverlässigkeit R_U kann über das Integral der Dichtefunktion mit der Zuverlässigkeit R_U als unteren Grenzwert bestimmt werden zu [10]:

$$P_A = P(R_U \leq R \leq 1) = \int_{R_U}^1 \varphi(R') dR' \quad (2)$$

Ihre Berechnung basiert auf der Dichtefunktion der Zuverlässigkeit $\varphi(R)$. Ein Vertrauensbereich wird als Intervall um einen Schätzwert definiert, innerhalb welchem die Zufallsgröße mit einer bestimmten Wahrscheinlichkeit, der Aussagewahrscheinlichkeit liegt. In der Regel wird ein Vertrauensbereich symmetrisch zum Median gelegt, so dass ein 90 %-Vertrauensbereich eine 5 %- und 95 %-Vertrauensgrenze besitzt. Um Vertrauensbereiche zu bestimmen, können unterschiedliche Methoden (siehe Tabelle 1) genutzt werden (in Anlehnung an [8]).

Tabelle 1: Methoden zur Bestimmung von Vertrauensbereichen.

Analytische Methoden	Approximative Methoden	Simulationsbasierte Methoden
Beta-Binomial-Intervall	Fisher-Matrix-Intervall	Bootstrap-Methode:
Likelihood-Ratio-Intervall	Wald Methode	Nicht-parametrisch
Bayes-Vertrauensintervall	Wilson Score Intervall	Parametrisch

Analytische und approximative Methoden erfordern meist eine analytische Beschreibung der Problemstellung. Häufig liegt dies jedoch nicht vor, z.B. wenn die Monte-Carlo-Simulation genutzt wird, so dass insbesondere die Bootstrap-Methode für einen solchen Anwendungsfall gut geeignet ist. Häufig werden in der Zuverlässigkeitstechnik betaverteilte Vertrauensbereiche bestimmt, d.h. die Dichtefunktion des Vertrauensbereiches der Zuverlässigkeit R_i gehorcht der Betaverteilung [11]:

$$\varphi(R_i) = \frac{1}{\beta(A_i, B_i)} R_i^{A_i-1} (1 - R_i)^{B_i-1} \quad (3)$$

4. Bootstrap-Methode

Bootstrapping [2] als Verfahren zur statistischen Auswertung, welches erstmals von Bradley Efron beschrieben wurde, eignet sich insbesondere als Analyseverfahren, wenn die Parameter einer gegebenen Stichprobe nicht über andere, analytische Methoden bestimmt werden können. Die Bootstrap-Methode kann in zwei Arten unterteilt werden (siehe Bild 3): nicht-parametrisches und parametrisches Bootstrapping.

Ausgangspunkt des nicht-parametrischen Bootstrapping ist eine n Ausfallzeiten umfassende Stichprobe $T = \{t_1, t_2, \dots, t_n\}$. Durch Ziehen mit Zurücklegen und ohne Beachtung der Reihenfolge von n zufälligen Zeiten aus T kann eine Bootstrap-Stichprobe \tilde{T} bestimmt werden. Anschließend werden die Parameter θ der Bootstrap-Stichprobe mit bekannten Parameterschätzverfahren, z.B. der Maximum-Likelihood-Methode [1], ermittelt, so dass eine Bootstrap-Verteilungsfunktion $\tilde{F}(t)$ (sogenannte Realisierung) gefunden werden kann.

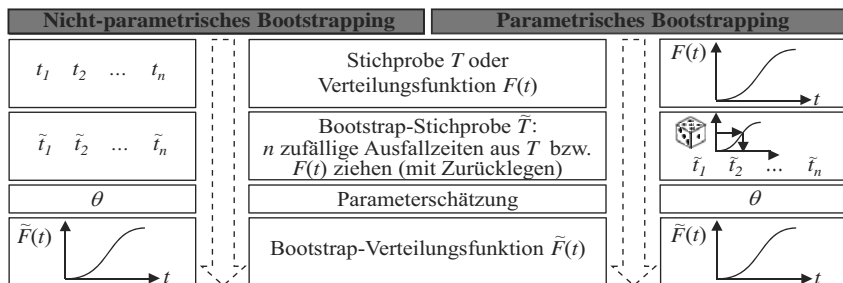


Bild 3: Nicht-parametrische (links) und parametrische (rechts) Bootstrap-Methode.

Analog wird das parametrische Bootstrapping basierend auf einer Verteilungsfunktion $F(t)$, durchgeführt. Über Pseudozufallszahlen und die Inversionsmethode werden n Zufallszeit aus der Verteilungsfunktion $F(t)$ generiert. Diese Zeiten bilden die Bootstrap-Stichprobe \tilde{T} , aus welcher die Realisierung $\tilde{F}(t)$ geschätzt werden kann. Zur Ermittlung eines Vertrauensbereichs können unterschiedliche Bootstrap-Typen eingesetzt werden [3], [4]:

- | | |
|--|--------------------------------------|
| • Empirischer Bootstrap-Vertrauensbereich, | • Bootstrap- t -Vertrauensbereich, |
| • Standard-Bootstrap-Vertrauensbereich, | • BC α -Vertrauensbereich, |
| • Perzentil-Bootstrap-Vertrauensbereich, | • etc. |

Im Folgenden wird nur der empirische Bootstrap-Vertrauensbereich weiter betrachtet.

Unabhängig vom Typ des Bootstrap-Vertrauensbereichs, muss die Bootstrap-Methode entsprechend ihrer Anzahl an Replikationen B -mal wiederholt werden, um den Bootstrap-Vertrauensbereich zu einer gegebenen Verteilungsfunktion zu bestimmen. Damit werden B Realisierungen $\tilde{F}_i(t)$, $i = 1(1)B$, generiert. Anschließend wird die Kurvenschaar der B Realisierungen $\tilde{F}_i(t)$ für jeden Zeitpunkt t_j statistisch ausgewertet, so dass für jeden Zeitpunkt eine

empirische Bootstrap-Vertrauensdichte $\varphi(\tilde{F}(t_j))$ gefunden wird. Über die Perzentile unterhalb diesen Dichtefunktionen kann der Vertrauensbereich zur ursprünglichen Verteilungsfunktion bestimmt werden. Bild 4 verdeutlicht die Vorgehensweise graphisch.

Um die Genauigkeit des Bootstrap-Vertrauensbereichs zu untersuchen, wird zur weibullverteilten Lebensdauerverteilung mit $T = 1.103,26$ h, $b = 4,0$ und $MTTF = 1.000$ h sowohl ein Bootstrap- als auch ein betaverteilter Vertrauensbereich mit 90 % Aussagewahrscheinlichkeit bestimmt. Der betaverteilte Vertrauensbereich dient als Referenzwert. Anschließend können die Mittelwerte der Beträge der Abweichungen Δ sowie deren Standardabweichungen σ zwischen den jeweiligen Vertrauensgrenzen bestimmt werden. Der Stichprobenumfang n sowie die Anzahl der Bootstrap-Replikationen B wird variiert. Bild 5 zeigt beispielhaft die Berechnungsergebnisse für die 95 %-Vertrauensgrenze. Für die 5 %-Vertrauensgrenze kann ein identisches Verhalten beobachtet werden.

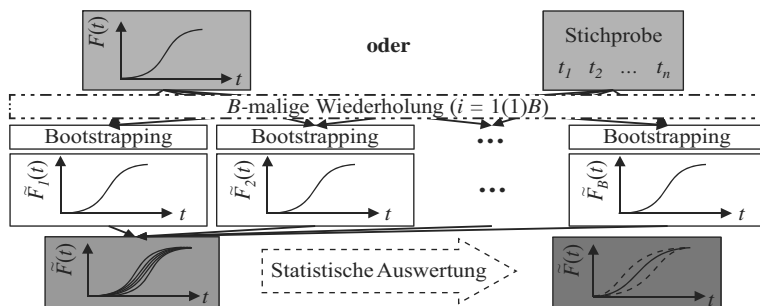


Bild 4: Bestimmung eines Vertrauensbereichs mit der Bootstrap-Methode.

Für $B \geq 200$ Bootstrap-Replikationen und einem Stichprobenumfang $n \geq 100$ werden die Abweichungen vernachlässigbar klein. In gewissem Umfang kann durch Erhöhung der Bootstrap-Replikationen der Einfluss eines geringeren Stichprobenumfangs ausgeglichen werden, weshalb immer eine große Anzahl an Bootstrap-Replikationen anzustreben ist.

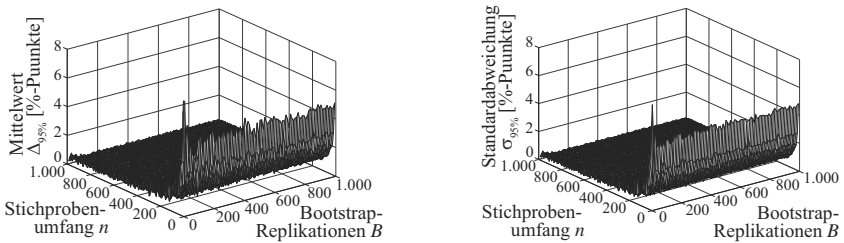


Bild 5: Mittelwert Δ (links) und Standardabweichung σ (rechts) der Beträge der Abweichungen zwischen der 95 %-Vertrauensgrenzen des Bootstrap-Vertrauensbereichs und des betaverteilten Vertrauensbereichs [5].

5. Bootstrap-Monte-Carlo-Simulation

Wird die Zuverlässigkeit bei periodischer Instandhaltung $R_{PM}(t)$ mit Hilfe des periodischen Instandhaltungsmodells nach Gleichung 1 berechnet, so kann keine Aussagewahrscheinlichkeit angegeben werden, da die Berechnung auf der mittelwertigen Zuverlässigkeitsfunktion $R(t)$ basiert und deren statistische Güte nicht miteinbezogen wird. Um einen Vertrauensbereich zu berechnen, wird die Bootstrap-Methode mit der Monte-Carlo-Simulation kombiniert. Mit der Bootstrap-Monte-Carlo-Simulation (BMCS) kann damit erstmals eine Aussagewahrscheinlichkeit bestimmt werden. Bild 6 zeigt den prinzipiellen Ablauf der BMCS.

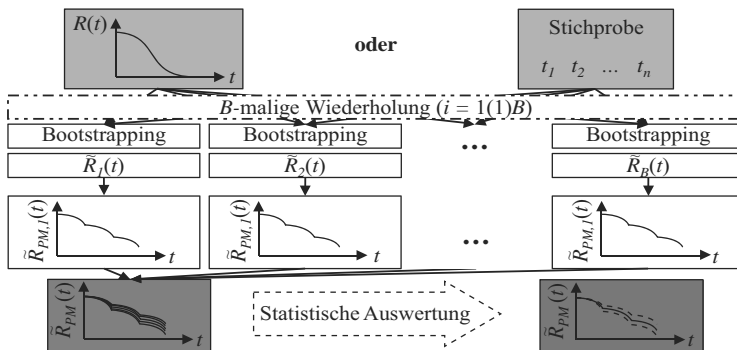


Bild 6: Bootstrap-Monte-Carlo-Simulation der Zuverlässigkeit und Aussagewahrscheinlichkeit bei periodischer Instandhaltung.

Zunächst wird eine Realisierung $\tilde{R}_i(t)$ der Zuverlässigkeitsfunktion ohne periodische Maßnahmen mit Hilfe von Bootstrapping generiert. Sofern anstelle einer Verteilung eine Stichprobe mit Ausfallzeiten gegeben ist, kann hierfür die nicht-parametrische Bootstrap-Methode

genutzt werden. Anschließend wird mit Hilfe der Monte-Carlo-Simulation nach Bild 2 eine Realisierung $\tilde{R}_{PM,i}(t)$ der Zuverlässigkeit bei periodischer Instandhaltung bestimmt.

Die Realisierung $\tilde{R}_{PM,i}(t)$ kann neben der Monte-Carlo-Simulation auch mit Hilfe von Gleichung 1 bestimmt werden (siehe [5]). Da die Monte-Carlo-Simulation die leistungsfähigere und allgemeinere Methode darstellt, wird im Folgenden nur die Berechnung der Zuverlässigkeit bei periodischer Erneuerung über die Monte-Carlo-Simulation detaillierter betrachtet. Diese liefert bei einer ausreichend hohen Anzahl an Monte-Carlo-Replikationen vergleichbare Ergebnisse zur Berechnung nach Gleichung 1.

Durch entsprechend der Bootstrap-Replikationenanzahl B -malige ($i = 1(1)B$) Wiederholung können B Realisierungen $\tilde{R}_{PM,i}(t)$ generiert werden. Die erhaltene Kurvenschaar wird anschließend zu jedem Zeitpunkt t_j des Betrachtungszeitraums statistisch ausgewertet, so dass zu jedem Zeitpunkt eine empirische Bootstrap-Vertrauensdichte $\varphi(\tilde{R}_{PM}(t_j))$ aus den einzelnen Funktionswerten der B Realisierungen abgeleitet werden kann. Über die Perzentile unterhalb dieser Dichtefunktionen kann der Vertrauensbereich zur Zuverlässigkeit mit periodischer Instandhaltung berechnet werden.

Bild 7 zeigt beispielhaft die Zuverlässigkeit bei periodischer Instandhaltung mit 90 %-Vertrauensbereich. Die Zuverlässigkeit einer Einheit ohne planmäßige Instandhaltung wird über die Weibullverteilung mit $T = 1.103,26$ h, $b = 4,0$ und $MTTF = 1.000$ h beschrieben, wobei für deren Ermittlung $n = 20$ Stichprobenelemente analysiert wurden. Als Instandhaltungsintervall wurde $T_{PM} = 1.000$ h angenommen. Die BMCS wurde mit $B = 1.000$ Bootstrap- und $n_{MC} = 10.000$ Monte-Carlo-Replikationen durchgeführt.

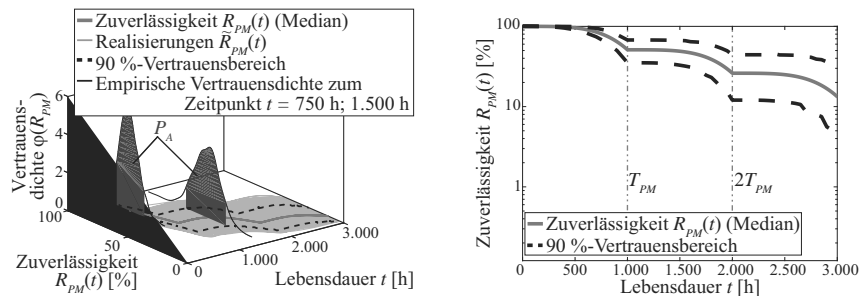


Bild 7: Zuverlässigkeit bei periodischer Instandhaltung mit 90 %-Vertrauensbereich.

Die BMCS bietet ein hohes Potenzial zur Anwendung bei allgemeinen technischen Systemen. Neben der Prognose ist auch der Nachweis von Zuverlässigkeiten mit Aussagewahrscheinlichkeit möglich. Ferner kann die BMCS auch zur Integration der Aussagewahrschein-

lichkeit in die Berechnung von Verfügbarkeiten [5], [6], [9] genutzt werden. Die BMCS liefert als approximative Methode Näherungslösungen der Aussagewahrscheinlichkeiten.

6. Verifikation der Bootstrap-Monte-Carlo-Simulation

Zur Verifikation der über die BMCS ermittelten Aussagewahrscheinlichkeit kann der Vertrauensbereich zur Zuverlässigkeit bei periodischer Instandhaltung auch mit Hilfe der Momentenmethode und eines betaverteilten Vertrauensbereichs der Zuverlässigkeit ohne periodische Maßnahmen bestimmt werden. Die m Faktoren der Gleichung 1 werden genutzt und analog zu einem aus m Komponenten bestehenden Seriensystem [11] der Erwartungswert der Zuverlässigkeit bei periodischer Instandhaltung berechnet. Die Parameter A_i und B_i sind durch den betaverteilten Vertrauensbereich der Zuverlässigkeit ohne periodische Instandhaltung bekannt. Für den Erwartungswert $E(R_{PM})$ und die Varianz $Var(R_{PM})$ der Zuverlässigkeit mit periodischer Instandhaltung ergibt sich nach [7], [11]:

$$E(R_S) = \prod_{i=1}^m \frac{A_i}{A_i + B_i} \quad (4)$$

$$Var(R_S) = \prod_{i=1}^m \frac{A_i \cdot (A_i + 1)}{(A_i + B_i) \cdot (A_i + B_i + 1)} - (E(R_{PM}))^2 \quad (5)$$

Daraus lassen sich die Bestimmungsgleichungen für die Parameter A_{PM} und B_{PM} der betaverteilten Vertrauensdichte $\varphi(R_{PM})$ ableiten. Die über die Momentenmethode berechneten Vertrauensbereiche sind ebenfalls approximativ, da die tatsächliche Vertrauensdichte mit einer Betaverteilung angenähert wird. Der ermittelte Erwartungswert ist jedoch der exakte Wert. Bild 8 vergleicht den 90 %-Vertrauensbereich der Momentenmethode und der BMCS. Die Eingangsdaten bleiben identisch zu den Berechnungen in Bild 7.

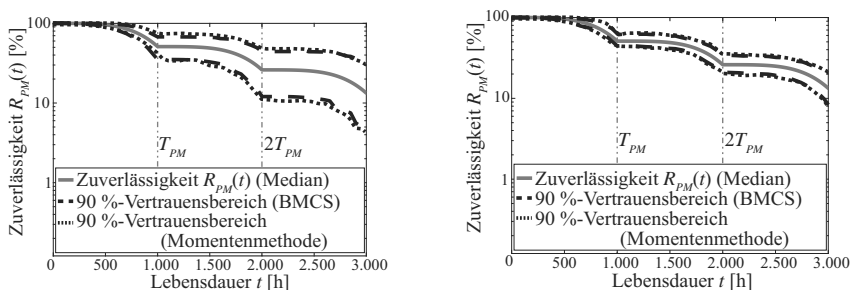


Bild 8: Vergleich des 90 %-Vertrauensbereich der BMCS bzw. der Momentenmethode für einen Stichprobenumfang $n = 20$ (links) und $n = 500$ (rechts).

Bereits bei geringem Stichprobenumfang ($n = 20$, Bild 8 links) weisen beide Vertrauensbereiche einen vergleichbaren Verlauf auf. Bei großem Stichprobenumfang ($n = 500$, Bild 8 rechts) sind diese nahezu deckungsgleich. Die Abweichungen bei geringem Stichprobenumfang werden stark durch die hohe statistische Unsicherheit der geringen Datenbasis beeinflusst. Der Vertrauensbereich wird desto enger, je größer der Stichprobenumfang ist.

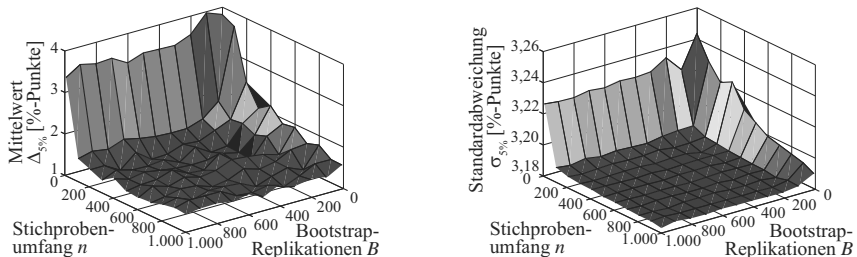


Bild 9: Mittelwert Δ (links) und Standardabweichung σ (rechts) der Beträge der Abweichungen zwischen der über die BMCS ($n_{MC} = 1.000$) bzw. die Momentenmethode berechneten 5 %-Vertrauensgrenzen.

Der Einfluss des Stichprobenumfangs wird zudem deutlich, sofern die Mittelwerte der Beträge der Abweichungen Δ sowie deren Standardabweichung σ zwischen den jeweiligen Vertrauensgrenzen der BMCS bzw. der Momentenmethode betrachtet werden. Bild 9 zeigt diese Kennwerte unter variierendem Stichprobenumfang bzw. variierender Bootstrap-Replikationsanzahl für die 5 %-Vertrauensgrenze. Die BMCS wurde mit $n_{MC} = 1.000$ Monte-Carlo-Replikationen durchgeführt. Für die 95 %-Vertrauensgrenze kann ein identisches Verhalten beobachtet werden.

Bis zu einer Bootstrap-Replikationsanzahl $B = 200$ und einem Stichprobenumfang von $n = 100$ nehmen die Abweichungen stark ab. Für einen Stichprobenumfang $n \geq 200$ in Kombination mit $B \geq 200$ Bootstrap-Replikationen sind die Abweichungen vernachlässigbar.

7. Anwendungsszenario

Im Folgenden wird eine Produktionsanlage betrachtet, für welche eine geeignete, periodische Instandhaltungsstrategie gesucht wird. Es soll bis zum Zeitpunkt $t = 750$ h eine Mindestzuverlässigkeit von 50 % mit 95 % Aussagewahrscheinlichkeit erreicht werden. Für die Produktionsanlage wurden aus Zuverlässigkeitstests $n = 20$ Ausfalldaten (siehe Tabelle 2) ermittelt. Aufgrund produktionstechnischer Randbedingungen können die Instandhaltungen nur in periodischen Abständen von 250 h oder 500 h durchgeführt werden.

Tabelle 2: Ausfalldaten der Produktionsanlage [h].

177,5	252,5	305,0	353,5	402,0	455,0	515,5	591,0	695,5	881,0
221,0	280,0	329,5	377,5	428,0	484,0	551,0	638,0	770,5	1.097,0

Die Analyse der Ausfalldaten nach Tabelle 2 mit 90 %-Vertrauensbereich mit Hilfe der Bootstrap-Methode ($B=1.000$) ergibt den in Bild 10 dargestellten Verlauf der Zuverlässigkeit. Dieser ähnelt einer lognormalverteilten Zuverlässigkeit mit $\mu=6,0896$, $\sigma=0,5$ und $MTTF=500$ h. Ohne Instandhaltungsmaßnahmen kann die B_{50} -Lebensdauer mit 95 % Auswahrscheinlichkeit nur bis zum Zeitpunkt $t=375$ h nachgewiesen werden.

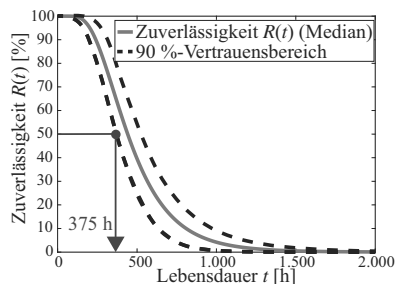


Bild 10: Zuverlässigkeit ohne Instandhaltungsmaßnahmen und 90 %-Vertrauensbereich.

Um die Zuverlässigkeit zu erhöhen, wird zunächst das Instandhaltungsintervall $T_{PM} = 500$ h gewählt und die Zuverlässigkeit bei periodischer Instandhaltung mit 90 %-Vertrauensbereich mittels BMCS ($B=1.000$, $n_{MC}=10.000$) bestimmt. Wie Bild 11 links zeigt, können die Anforderungen bei dieser Instandhaltungsstrategie ebenfalls nur bis $t=375$ h garantiert werden. Sofern das Instandhaltungsintervall weiter verkleinert wird ($T_{PM}=250$ h, Bild 11 rechts), können die Forderungen erfüllt werden. Die B_{50} -Lebensdauer kann mit 95 % Aussagewahrscheinlichkeit bis zum Zeitpunkt $t=900$ h nachgewiesen werden. Folglich ist die Instandhaltungsstrategie mit periodischen Maßnahmen in den Zeitabständen $T_{PM}=250$ h zu wählen.

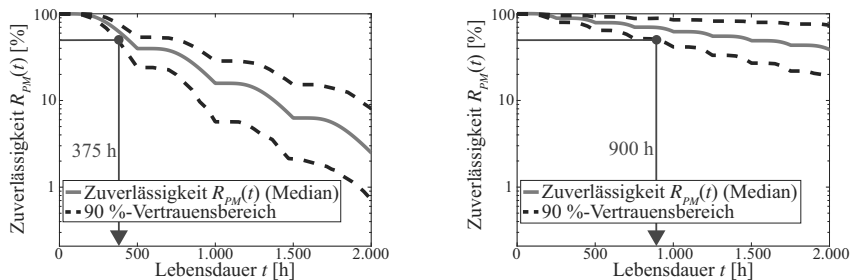


Bild 11: Zuverlässigkeit mit Instandhaltungsmaßnahmen

($T_{PM} = 500$ h links, $T_{PM} = 250$ h rechts) und 90 %-Vertrauensbereich.

8. Zusammenfassung

Als neue Methode zur Prognose und zum Nachweis der Zuverlässigkeit und Aussagewahrscheinlichkeit periodisch instandgehaltener Systeme wird die Bootstrap-Monte-Carlo-Simulation (BMCS) vorgestellt. Sie erlaubt die Analyse basierend auf einer aus Ausfallzeiten bestehenden Stichprobe oder einer mittelwertigen Verteilungsfunktion unter Angabe des Stichprobenumfangs. Die BMCS setzt keine Einschränkungen hinsichtlich der Verteilungsart, d.h. es können z.B. Weibull-, lognormal- oder Exponentialverteilungen berücksichtigt werden.

Zur Verifikation der BMCS wurde zunächst die Bootstrap-Methode zur Ermittlung eines Vertrauensbereichs untersucht. Für eine ausreichend große Anzahl der Bootstrap-Replikationen ($B \geq 200$) bzw. einen großen Stichprobenumfang ($n \geq 100$) sind die Abweichungen des Bootstrap-Vertrauensbereichs zu einem betaverteilten Vertrauensbereich vernachlässigbar. Die Güte der BMCS wurde ferner anhand der Zuverlässigkeit einer periodisch instandgehaltenen Komponente untersucht, indem der über die BMCS bestimmte Vertrauensbereich mit einem mittels der Momentenmethode bestimmten Vertrauensbereich verglichen wurde. Für eine ausreichend große Anzahl an Bootstrap- und Monte-Carlo-Replikationen ($B \geq 200$, $n_{MC} \geq 1.000-10.000$) liefert die BMCS als approximative Methode sehr gute Ergebnisse.

Nach einer abschließend durchgeführten Parameterstudie der Eingangsgrößen der BMCS wurde in einem Anwendungsszenario eine Produktionsanlage mit periodischer Erneuerung betrachtet. Durch die Möglichkeit zur Anwendung bei allgemeinen technischen Systemen ist die BMCS sehr innovativ. Sie kann auf die Berechnung von Verfügbarkeiten mit Angabe einer Aussagewahrscheinlichkeit [6], [9] übertragen werden und bietet ein hohes Potenzial zur Integration allgemeiner Instandhaltungsmaßnahmen bzw. Analyse individueller Kenngrößen.

Literaturverzeichnis

- [1] Bertsche, B.; Lechner, G.: Zuverlässigkeit im Fahrzeug- und Maschinenbau, Ermittlung von Bauteil- und System-Zuverlässigkeiten. Berlin, Heidelberg, New York: Springer, 2004.
- [2] Efron, B.; Tibshirani, R.J.: An Introduction to the Bootstrap. Boca Raton, London, New York, Washington: Chapman & Hall/CRC, 1993.
- [3] Lim, J.-H.; Shin, S.W.; Kim, D.K.; Park, D.H.: Bootstrap Confidence Intervals for Steady-State Availability. In: Asia-Pacific Journal of Operational Research, Vol. 21, No. 3, 2004, S. 407-419.
- [4] Meeker, W.Q.; Escobar, L.A.: Statistical Methods for Reliability Data. New York, Chichester, Weinheim, Brisbane, Singapore, Toronto: John Wiley & Sons, Inc., 1998.
- [5] Müller, F.: Integration der Aussagewahrscheinlichkeit in die Berechnung und Simulation der Verfügbarkeit. Masterarbeit, Universität Stuttgart, 2015.
- [6] Müller, F.; Zeiler, P.; Bertsche, B.: Availability demonstration with confidence level based on reliability and maintainability. In: Proc. of Annual Reliability and Maintainability Symposium (RAMS 2017), Orlando, USA, 2017.
- [7] Thompson, W.E.; Haynes, R.D.: On the Reliability, Availability and Bayes Confidence Intervals for Multicomponent Systems. In: Naval Research Quarterly, Vol. 27, 1980, S. 345-358.
- [8] Yang, G.: Reliability Engineering. Hoboken, New Jersey: John Wiley & Sons, Inc., 2007.
- [9] Zeiler P.; Müller, F.; Bertsche, B.: New methods for the availability prediction with confidence level. In: Proc. of European Safety and Reliability Conference (ESREL 2016), Glasgow, Schottland, 2016.
- [10] Zeiler, P.; Bertsche, B.: Component reliability allocation and demonstration test planning based on system reliability confidence limit. In: Proc. of European Safety and Reliability Conference (ESREL 2015), Zürich, Schweiz, 2015.
- [11] Zeiler, P.; Bertsche, B.: Zur Berechnung der Systemzuverlässigkeit mit Aussagewahrscheinlichkeit – Methodenvergleich und Anwendung zur Zuverlässigkeitsallokation und -testplanung. In: 27. VDI Fachtagung Technische Zuverlässigkeit 2015, VDI-Berichte 2260, Leonberg, Deutschland, 2015.

Robustheitsoptimierung innerhalb des Entwicklungsprozesses durch Integration von Fertigungstoleranzen in die Simulation

C. Hayer, Dr.-Ing. **S. Fiebig**, **J. Sellschopp**,
Volkswagen AG, Braunschweig;
Prof. Dr.-Ing. **T. Vietor**, Technische Universität Braunschweig

Kurzfassung

Um die steigenden Anforderungen an automobile Strukturbauteile zu erfüllen, werden im Produktentstehungsprozess verstärkt rechnergestützte Optimierungsmethoden eingesetzt. Dabei ist es wichtig, die Robustheit der so generierten Bauteile frühestmöglich zu betrachten. Streuungen der Bauteileigenschaften im Toleranzbereich können zu einer Abweichung der Prüfergebnisse gegenüber den auf dem Nennbauteil basierenden Vorhersagen aus der Simulation führen. Bauteile, bei denen dadurch die Anforderungen am Prüfstand verletzt werden, müssen daraufhin durch konstruktive Maßnahmen robuster gestaltet werden. Dies verursacht einen zusätzlichen Zeitaufwand, verbunden mit einer Erhöhung von Kosten und Gewicht. Dieser Beitrag zeigt, wie Fertigungstoleranzen bereits in der Simulation optimiert werden können. Zunächst werden sensible Bereiche mithilfe aufwandsminimierter Versuchsplanung in der Simulation identifiziert. Mithilfe eines analytischen Ersatzmodells wird das Bauteil durch Anpassungen der Toleranzbereiche robuster gestaltet. Die optimierte Tolerierung ermöglicht es, den Leichtbauvorteil aus der Strukturoptimierung in die Prototypenphase zu übertragen. Mit einer verbesserten Vorhersage der Prüfergebnisse können zusätzlich unnötige Entwicklungsschleifen vermieden werden, wodurch erhebliche Zeit- und Kosteneinsparungen erzielt werden.

1. Einleitung

In einem immer anspruchsvolleren Umfeld aus steigendem Wettbewerb und strengeren gesetzlichen Richtlinien in der Automobilindustrie spielt Leichtbau eine immer wichtigere Rolle [1]. Die häufigste Aufgabenstellung bei Bauteiloptimierungen ist die Minimierung des Gewichts unter Berücksichtigung von technischen Anforderungen, wie zum Beispiel Festigkeit und Steifigkeit. Eine der wesentlichen Maßnahmen um das Leichtbaupotential maximal auszuschöpfen ist eine möglichst enge Auslegung des Nennbauteils an die Grenzen der Anforderungen, ohne diese zu verletzen [2]. Streuungen der Geometrie und der Materialeigen-

schaften innerhalb des Toleranzbereiches werden bei der Auslegung in der Regel nicht berücksichtigt. Diese führen jedoch zu einer Abweichung der Prüfergebnisse gegenüber den Vorhersagen aus der Simulation. So lässt sich erst mithilfe der Prüfergebnisse eine Aussage über die Robustheit des Bauteils treffen. Verletzt ein Prototyp die Anforderungen, muss der Bauteilentwurf durch konstruktive Maßnahmen robuster gestaltet werden. Dies verursacht einen zusätzlichen Zeitaufwand, erhöhte Kosten und ein höheres Gewicht. Ein robustes Bauteil wird im weiteren Verlauf dadurch definiert, dass die Anforderungen in jeder tolerierten Konfiguration erfüllt werden. Taguchi, dem das Konzept von robustem Design zugesprochen wird, entwickelte eine experimentelle Methode zur Optimierung der Robustheit, die später auch für die Anwendung in der Simulation angepasst wurde [3,4,5,6]. Das Ziel ist dabei, die Robustheit durch Anpassungen des Nennbauteils zu optimieren. Neben der robusteren Gestaltung durch konstruktive Maßnahmen ist die Anpassung der Toleranzbereiche eine weitere Möglichkeit, die Streuung der Prüfergebnisse zu beeinflussen. Die Berücksichtigung von Toleranzen im Produktentwicklungsprozess beschränkt sich bisher auf das Toleranzmanagement, dessen Ziel es ist, die Fertigbarkeit eines Bauteils sicherzustellen. Dabei wird keine Information darüber gegeben, wie sich die Toleranzen auf die technischen Anforderungen auswirken. Um diesen Einfluss so früh wie möglich quantifizieren zu können, muss die Streuung der Prüfergebnisse bereits in der Simulation ermittelt werden [7]. Um dies sicherzustellen, muss theoretisch jede tolerierte Konfiguration dargestellt und simuliert werden. Nur so können die Konfigurationen, in denen die Anforderungen nicht erfüllt werden, herausgefiltert werden. Deswegen ist es nötig eine Methode zu entwickeln, die für eine große Anzahl an Designparametern angewandt werden kann. Diese soll es nicht nur ermöglichen, die Streuung der Prüfergebnisse vorherzusagen sondern diese auch durch eine Optimierung der Toleranzbereiche in den gewünschten Bereich zu verschieben.

2. Integration einer Robustheitsoptimierung in den Entwicklungsprozess

Die Methode, mit der diese Fragestellung beantwortet wird, kann in zwei separate Phasen unterteilt werden. In der ersten Phase, der Robustheitsanalyse, wird die Streuung der Prüfergebnisse ermittelt. In der zweiten Phase, der Robustheitsoptimierung, werden die Toleranzbereiche optimiert. Bild 1 zeigt die Integration beider Phasen in den aktuellen Entwicklungsprozess.

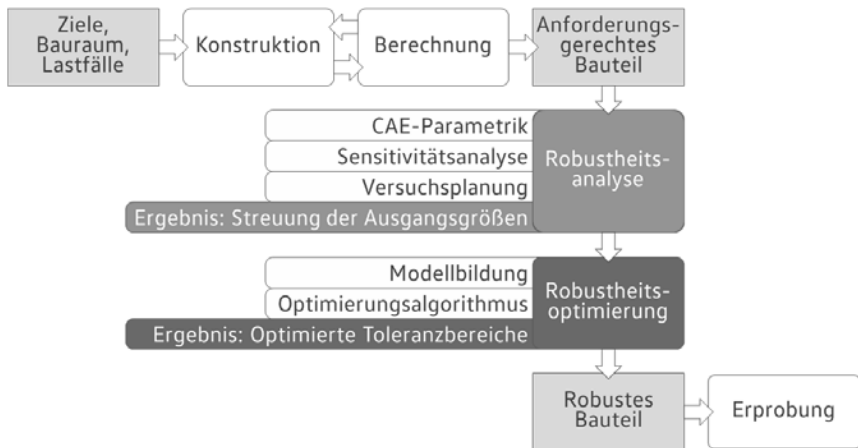


Bild 1: Entwicklungsprozess mit integrierter Robustheitsoptimierung

Diese Methode nimmt als Ausgangspunkt einen Bauteilentwurf, der die technischen Anforderungen erfüllt. Als Ergebnis erhält man eine Tolerierung desselben Entwurfs, mit der die Anforderungen in jeder tolerierten Konfiguration eingehalten werden. In der Analysephase wird zunächst ein parametrisiertes CAE-Modell erstellt, mithilfe dessen alle Abweichungen zum Nennbauteil innerhalb des Toleranzbereiches dargestellt werden können. Die resultierende Anzahl an Designparametern ist in der Regel zu hoch, um mithilfe von Simulationen in einem angemessenen Zeitrahmen eine Aussage über die Streuung der Ausgangsgrößen (Steifigkeit, Festigkeit,...) treffen zu können. Um die Anzahl an Parametern zu reduzieren, werden mithilfe der Taguchi Methodik die Parameter identifiziert, die den größten Einfluss auf die relevanten Ausgangsgrößen haben. Anhand eines effizienten Versuchsplans wird mit der reduzierten Anzahl an Parametern die Streuung der Ausgangsgrößen quantifiziert. Um den Rechenaufwand für die Optimierung der Toleranzbereiche zu minimieren, wird in der zweiten Phase ein analytisches Ersatzmodell erstellt. Dieses basiert auf den Ergebnissen aus der Versuchsplanung, die in der ersten Phase zur Quantifizierung der Streuung genutzt wurde. Das Ersatzmodell dient dazu, ohne weiteren Simulationsaufwand die Toleranzen zu optimieren. Ziel der Optimierung ist die Minimierung der Summe der Größen der Toleranzbereiche. Dabei wird auf die Einhaltung der Anforderungen in jeder tolerierten Konfiguration geachtet.

3. Robustheitsanalyse

3.1 Parametrisches CAE-Modell

Um jede Konfiguration im Toleranzbereich darstellen zu können, müssen alle Simulationsdaten modifiziert werden. Informationen wie Blechdicken, Materialeigenschaften oder Intensität und Richtung von Lasteinleitungen können in der Regel direkt gesteuert werden. Um eine größtmögliche Präzision der Geometriemodifikation zu gewährleisten, wird die kleinste zur Verfügung stehende Auflösung zur Parametrisierung genutzt. Die Veränderung der Geometrie wird daher über die Verschiebung von FEM-Knoten gesteuert. Diese wird den Anforderungen angepasst, die durch die Toleranzen gegeben werden [8]. Üblicherweise wird zwischen Veränderungen der Geometrie in der Ebene (in-plane) und außerhalb der Ebene (out-of-plane), die durch das Blech definiert wird, unterschieden. Dabei werden die Knoten immer normal zum Bauteil verschoben. In Bild 2 wird an einem einfachen Beispiel veranschaulicht, wie diese Veränderungen an einem Blechbauteil aussehen können.

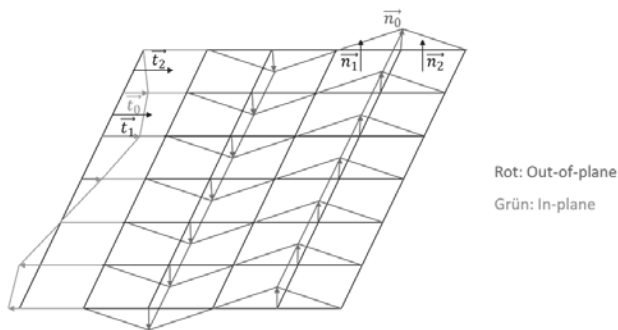


Bild 2: Geometriemodifikation am Beispiel einer Platte

Die in-plane Geometriemodifikation wird bei Beschnitten von Blechbauteilen eingesetzt. Die Richtung der Verschiebung (in Bild 2: \vec{t}_0) ergibt sich aus den Normalen zu den an einen Knoten grenzenden, freiliegenden Kanten (in Bild 2: \vec{t}_1 und \vec{t}_2). Bei der out-of-plane Verschiebung (in Bild 2: \vec{n}_0) ergibt sich die Richtung aus den Normalen zu den an einen Knoten grenzenden Elementen (in Bild 2: \vec{n}_1 und \vec{n}_2). Um die so entstehende Anzahl an Parametern zu reduzieren, werden die Elemente in der Praxis gruppiert, um physikalisch sinnvolle Modifikationen zu generieren. Die Gruppierung wird durch den Benutzer individuell festgelegt, sollte aber den Geometriemodifikationen entsprechen, die durch den Fertigungsprozess entstehen können.

3.2 Sensitivitätsanalyse

Um den Zeitaufwand zu verringern, wird in einem zweiten Schritt die Anzahl an Parametern reduziert. Dazu werden die Parameter, die den größten Einfluss auf die relevanten Ausgangsgrößen haben, mithilfe einer Sensitivitätsanalyse identifiziert. Diese wird in der Regel über Versuchspläne gesteuert, die sich hauptsächlich durch die Anzahl an Versuchen und die Genauigkeit der Quantifizierung unterscheiden. Um mit einer geringen Anzahl an Versuchen komplexe Zusammenhänge zwischen den Parametern darzustellen kann auf die orthogonalen Felder oder Taguchi-Matrizen zurückgegriffen werden [9]. Die Effektivität dieser Versuchspläne wird in mehreren Arbeiten beschrieben. In Kombination mit einer statistischen Auswertung der Simulationsergebnisse bilden diese Felder die Basis für die Taguchi-Methodik. Um die Anzahl an Berechnungen zu minimieren werden für jeden Parameter zwei Stufen gewählt. Stufen beschreiben festgelegte Zustände der Parameter. In der hier beschriebenen Methode sind die zwei gewählten Stufen die obere und die untere Grenze des Toleranzbereichs ($P=T^u$ oder $P=T^o$). Hat man zum Beispiel eine nominale Blechdicke von 3mm, mit einem Toleranzbereich von $\pm 0,15\text{mm}$, wären die zwei Stufen für diese Blechdicke 2,85mm und 3,15mm. Orthogonale Felder sind so aufgebaut, dass für jeden Parameter jede Stufe in gleichem Anteil über die Versuche verteilt ist. In Tabelle 1 wird der Zusammenhang zwischen der Anzahl an Parametern und der Anzahl an Versuchen zusammengefasst.

Tabelle 1: Größe der orthogonalen Felder

Anzahl Parameter n_p	Anzahl Versuche n_v	Anzahl Versuche mit $P=T^u / P=T^o$
2 - 3	4	2
4 - 7	8	4
$2^k - 2^{k+1} - 1$	2^{k+1}	2^k

Die Auswertung der Versuche erfolgt mithilfe einer Varianzanalyse [10]. Der Sobol-Wert, der zur Quantifizierung des Einflusses eines Parameters berechnet wird, stellt das Verhältnis zwischen der Varianz, die durch diesen Parameter verursacht wird, und der Varianz, die durch alle Parameter verursacht wird [11,12]. In der Regel hat eine Eingangsgröße einen signifikanten Einfluss auf die Ausgangsgrößen, wenn der errechnete S-Wert größer ist als 3 [6]. So lässt sich die Anzahl der Parameter je nach Anwendungsbeispiel um bis zu 99% reduzieren.

3.3 Ermittlung der Streuung der Ausgangsgrößen

Mit der reduzierten Anzahl an Parametern ist es möglich mithilfe einer weiteren Versuchsplanung eine Aussage über die Streuung der Ausgangsgrößen zu treffen. Um einen optimalen Versuchsplan zu finden, gibt es mehrere Kriterien, die es ermöglichen Versuchspläne zu vergleichen. Dazu gehören zum Beispiel die Orthogonalität, D-Optimalität, oder das Raumfüllungskriterium [13]. Verschiedene Studien zeigen, dass für eine große Anzahl an Parametern das „Orthogonal Array Design“ (OAD) am effektivsten ist, um das Spektrum der Eingangsgrößen abzudecken [14]. Dieser Versuchsplan nimmt als Basis das in der Literatur weit verbreitete „Latin Hypercube Sampling“ (LHS) [15] und ergänzt dieses um eine zusätzliche gleichmäßige Unterteilung des Wertebereichs. Dabei wird beachtet, dass jede dieser Unterteilung zu gleichen Teilen gefüllt wird. Bild 3 veranschaulicht diese Ergänzung an einem Beispiel mit zwei Parametern, deren Wertebereiche für das Orthogonal Sampling in jeweils 3 Teilbereiche unterteilt wurden.

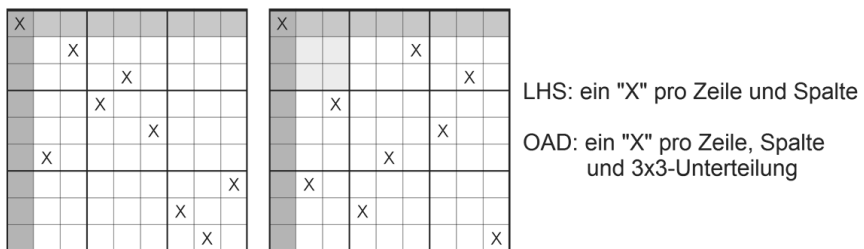


Bild 3: Vergleich zwischen LHS (links) und OAD (rechts)

Im Gegensatz zum LHS wird beim OAD sichergestellt, dass die Versuchspunkte gleichmäßig über den Wertebereich verteilt sind und alle Extremitäten des Wertebereichs abgedeckt sind. Anhand des so erstellten Versuchsplans kann die Streuung der Ausgangsgrößen ermittelt werden. Die Ergebnisse werden in Abschnitt 5 an einem Beispiel veranschaulicht.

4. Robustheitsoptimierung

4.1 Modellbildung

Mit der Modellbildung beginnt die zweite Phase des Prozesses, in der die Toleranzbereiche optimiert werden. Mit einem Ersatzmodell wird ein analytischer Zusammenhang zwischen Eingangs- und Ausgangsgrößen gebildet. Dadurch entfallen in der Optimierung die zeitaufwendigen Simulationen. Der erste Ansatz zur Bildung eines analytischen Modells wird von Box erarbeitet [16]. Mithilfe der „Response Surface Methodology“ wird hier ein linearer Zu-

sammenhang zwischen Input und Output erstellt. Dieser ermöglicht es jedoch nicht, Wechselwirkungen zwischen den verschiedenen Parametern darzustellen. Zu den moderneren Ansätzen der Modellbildung zählt der Ansatz des „Polynomial Chaos“, der auf Wieners Theorie des homogenen Chaos basiert [17]. Seit der Einführung wurde diese Methode an vielen Beispielen auf strukturmekanische Fragestellungen angewandt [18][19]. Bei der „Polynomial Chaos Expansion“ (PCE) werden Polynome als Basisfunktion für das Modell gewählt. Eine detailliertere Beschreibung dieser Methode kann in [20] gefunden werden. Die Form der genutzten Polynome hängt von der Verteilung der Eingangsparameter im Wertebereich ab. Im Fall einer gleichmäßigen Verteilung, wie sie bei der Tolerierung angenommen wird, werden die Legendre-Polynome gewählt. Üblicherweise reicht es bei strukturmekanischen Problemstellungen einfache Wechselwirkungen zu berücksichtigen. Deswegen werden zunächst nur Polynome zweiter Ordnung verwendet. Bei der PCE wird in der Regel ein nichtlineares Ersatzmodell erstellt. In den Fällen, in denen die Wechselwirkungen zwischen den Parametern keinen Einfluss haben, ist das Ersatzmodell linear.

4.2 Optimierung der Toleranzbereiche

Bei der Optimierung der Toleranzbereiche muss darauf geachtet, dass die Reduzierung der Toleranzbereiche in sensiblen Bereichen durch eine Erweiterung der Toleranzbereiche in anderen Bereichen ausgeglichen wird. So wird sichergestellt, dass durch die Anpassung der Toleranzen keine unnötig hohen Fertigungskosten entstehen. In der Optimierung werden für jeden Designparameter die untere und die obere Grenze des Toleranzbereichs (T^u , T^o) als Variable gewählt. Als Zielfunktion (F) wird die Gesamtgröße der Toleranzbereiche maximiert:

$$F(T^u, T^o) = \sum_{i=1}^{n_p} T_i^o - T_i^u$$

n_p ist hier die Anzahl an Designparametern. Die Nebenbedingungen (G) der Optimierung sind eine Ober- und Untergrenze für den Toleranzbereich, eine Mindestgröße für den Toleranzbereich, und die technischen Anforderungen, die als Restriktion (R) zusammengefasst werden. Die minimale Größe für den Toleranzbereich ist hierbei abhängig vom gewählten Fertigungsverfahren.

$$\left. \begin{aligned} T_i^u &\geq T_{\min}, T_i^o \leq T_{\max}, T_i^o - T_i^u \geq dT \\ R(T^u, T^o) &\geq 0 \end{aligned} \right\} G(T^u, T^o) \geq 0$$

So entsteht ein Ungleichungssystem mit mehr Nebenbedingungen als Variablen. Die Form des Ersatzmodells ist ausschlaggebend dafür, ob das Problem linear oder nichtlinear ist. Ein

lineares Problem (LP) wird es mithilfe des Simplex-Verfahren gelöst [21]. Bei dieser Methode wird zunächst eine Startlösung bestimmt. Ausgehend von dieser Lösung, beginnt ein iteratives Verfahren. In jedem Schritt wird versucht, eine neue Lösung mit besserem Zielfunktionswert zu konstruieren, bis dies nicht mehr möglich ist. Um einen Schritt zu machen, bewegt man sich entlang des Polyeders der möglichen Lösungen. Dieser wird aus den sogenannten Basislösungen gebildet. Wie diese Basislösungen erhalten werden, wird in [21] erläutert. Besitzt (LP) ein Optimum, wird es mit dem Simplex-Verfahren ermittelt. Im Fall eines nichtlinearen Problems (NP) findet diese Methode ein lokales, jedoch nicht zwangsläufig ein globales Optimum. Das gefundene Optimum hängt in diesem Fall von der Startlösung ab. Für (NP) wird daher die Lagrange Methode verwendet [22]. Mithilfe der Lagrange Multiplikatoren $\lambda = (\lambda_k \geq 0)_{1 \leq k \leq n_N}$ wird eine neue Zielfunktion (L) generiert:

$$L(T^u, T^o, \lambda) = F(T^u, T^o) + \sum_{k=1}^{n_N} \lambda_k G_k(T^u, T^o)$$

n_N ist die Anzahl an Nebenbedingungen. Die neue Zielfunktion muss zwei Konditionen erfüllen:

$$\begin{cases} \vec{\nabla}_{T^u, T^o} L(T^u, T^o, \lambda) = 0 \\ \min(\lambda_k, G_k(T^u, T^o)) = 0, \forall k \in \{1 \dots n_N\} \end{cases}$$

So ergibt sich ein Gleichungssystem mit $(n_p + n_N)$ Gleichungen und $(n_p + n_N)$ Variablen, dass sich einfach lösen lässt. Die so errechneten $[T_i^u; T_i^o]$ bilden nun die neuen Toleranzbereiche für das Bauteil.

5. Robustheitsoptimierung am Beispiel eines Hilfsrahmens

Die Robustheitsoptimierung wird in diesem Abschnitt an einem Hilfsrahmen veranschaulicht. Die Geometrie, Randbedingungen und Anforderungen ähneln vorhandenen Bauteilen, ohne dabei einem realen Produkt exakt zu entsprechen. Bild 4 zeigt die Geometrie des betrachteten Bauteils.

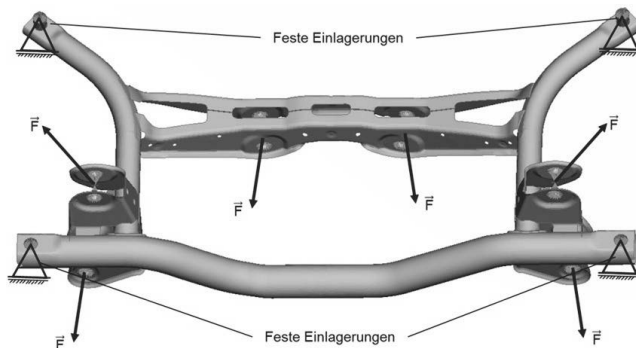


Bild 4: Geometrie des Hilfsrahmens

Der Hilfsrahmen wird an 4 Punkten fest gelagert. In vier Lastfällen werden Brems- und Beschleunigungskräfte an den 6 in Bild 4 beschriebenen Kräfteinleitungspunkten eingebracht. Zusätzlich werden vier verschiedene Steifigkeiten ausgewertet. Das Bauteil wird mit 72006 Shell-Elementen vernetzt. Zur Modifikation der Geometrie werden 59 Parameter verwendet. Dies sind 9 Blechdicken, 32 Beschnittkanten (in-plane Knotenverschiebung), 14 Auszugshöhen von Taschen (out-of-plane Knotenverschiebung) und 4 zusätzliche Parameter, mithilfe derer die Geometrie des vorderen Rohrs (in Bild 4 unten) verändert wird. Im ersten Schritt wird die Anzahl der Parameter in der Sensitivitätsanalyse reduziert. Bei 59 Parametern sind dafür 64 Simulationen notwendig. Bild 5 zeigt die Verteilung der S-Werte über die Designparameter.

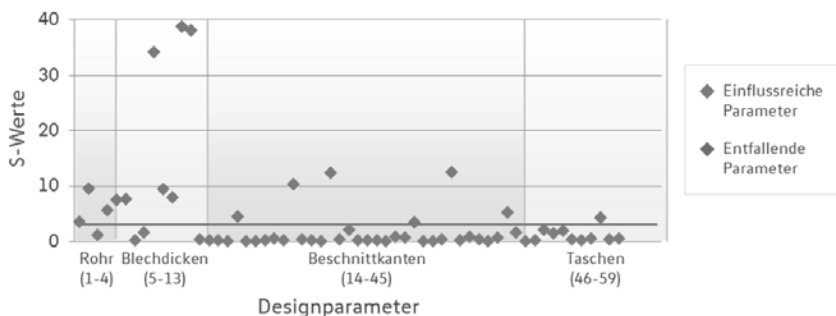


Bild 5: S-Werte für die Designparameter des Hilfsrahmens

Die Anzahl der Parameter wird im ersten Schritt auf 17 Parameter reduziert, die über alle Parameter-Kategorien verteilt sind. Anhand dieser Parameter wird ein Versuchsplan erstellt, um die Streuung der Ausgangsgrößen zu quantifizieren. Für 17 Eingangsgrößen, deren Wertebereich jeweils in 4 Teilbereiche unterteilt wird, werden 256 Simulationen benötigt. Zusätzlich werden die Parameter, die einen geringen Einfluss auf die Ausgangsgrößen haben, zufällig in ihrem Toleranzbereich gestreut. So kann ermittelt werden, welcher Anteil der Konfigurationen die Anforderungen verletzt. Um die verschiedenen Anforderungen zu vereinheitlichen, wird die normierte technische Anforderung definiert. Dieser stellt das Ergebnis im Verhältnis zu den Grenzen der Anforderungen dar. Ist dieser Wert kleiner als 1, wird keine Anforderung verletzt. Ist der Wert größer wird mindestens eine Anforderung verletzt. In Bild 6 ist dieser Wert prozentual dargestellt.

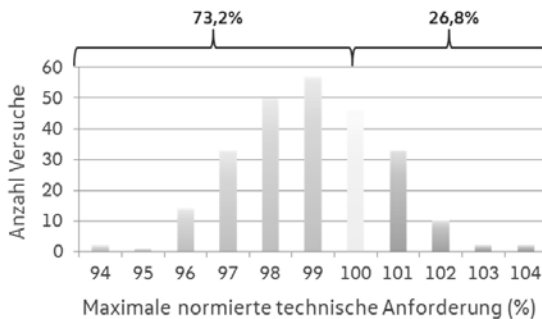


Bild 6: Streuung der Ausgangsgrößen (256 Versuche)

In 25% der Konfigurationen wird mindestens eine Anforderung verletzt. Durch die Anpassung der Toleranzbereiche sollen alle Anforderungen in jeder Konfiguration erfüllt werden. Um das benötigte PCE-Modell zu bilden, werden zusätzliche 342 Berechnungen durchgeführt. Auch dabei werden die Parameter, die einen geringen Einfluss auf die Ausgangsgrößen haben, zufällig gestreut. Dabei entsteht ein nichtlineares Ersatzmodell. Der gewählte Optimierungsalgorithmus ist daher die Lagrange-Methode. Die Toleranzen, die als Ausgangspunkt dienen, sowie die aus der Optimierung resultierenden Toleranzen sind in Tabelle 2 zusammengefasst. Als Unter- und Obergrenze für die Optimierung wird der doppelte initiale Toleranzbereich gewählt.

Tabelle 2: Alter und neuer Toleranzbereich für die 17 Einflussreichen Parameter

Rohr	Anpassung des Toleranzbereichs (%)	Beschnittkante	Anpassung des Toleranzbereichs (mm)
1	[-1 ; 1] ► [-0,43 ; 0,4] ↘ 59%	1	[-0,5 ; 0,5] ► [0,26 ; 0,84] ↘ 42%
2	[-1 ; 1] ► [-0,39 ; 0,43] ↘ 59%	2	[-0,5 ; 0,5] ► [-0,47 ; 1] ↗ 47%
3	[-1 ; 1] ► [-1,26 ; 2] ↗ 63%	3	[-0,5 ; 0,5] ► [-0,45 ; 0,47] ↘ 8%
Blechedicke	Anpassung des Toleranzbereichs (mm)	4	[-0,5 ; 0,5] ► [-0,62 ; 0,48] ↗ 9%
1	[2,88 ; 3,12] ► [2,94 ; 3,24] ↗ 25%	5	[-0,5 ; 0,5] ► [-0,89 ; 0,4] ↗ 29%
2	[2,88 ; 3,12] ► [2,92 ; 3,24] ↗ 35%	Auszugshöhe	Anpassung des Toleranzbereichs (mm)
3	[2,88 ; 3,12] ► [3,06 ; 3,24] ↘ 23%	1	[-0,5 ; 0,5] ► [-0,46 ; 0,25] ↘ 29%
4	[2,88 ; 3,12] ► [2,98 ; 3,24] ↗ 8%	2	[-0,5 ; 0,5] ► [-1 ; 0,47] ↗ 47%
5	[2,88 ; 3,12] ► [2,88 ; 3,24] ↗ 48%		
6	[2,88 ; 3,12] ► [2,88 ; 3,24] ↗ 50%		
7	[2,88 ; 3,12] ► [3,02 ; 3,24] ↘ 9%		

↗ Vergrößerung des Toleranzbereichs (%) ↘ Reduzierung des Toleranzbereichs (%)

Die Gesamtgröße der Toleranzbereiche wird in der Optimierung um 7% vergrößert. Die Einhaltung der Anforderungen in jeder Konfiguration wird mithilfe des Ersatzmodells sichergestellt. Um dieses Ergebnis zu prüfen wird, wie schon in der ersten Phase des Prozesses, die Streuung der Ausgangsgrößen ermittelt, diesmal mit den neuen Toleranzbereichen. Der Vergleich zu der Streuung im Ausgangsbauteil (in grau) wird in Bild 7 dargestellt.

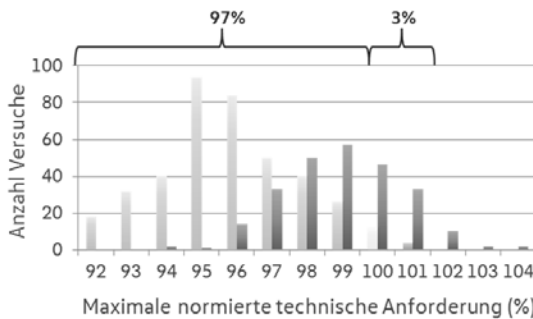


Bild 7: Streuung der Ausgangsgrößen mit neuen Toleranzbereichen (400 Versuche)

Bei 3% der Versuche die Anforderung trotz der Optimierung verletzt wird. Dabei beträgt die maximale Überschreitung des Spannungsgrenzwerts 0,6%, und die maximale Unterschreitung des Steifigkeitsgrenzwerts 0,8%. Die Differenz zwischen den Aussagen aus der Optimierung und der Berechnung kann durch die Ungenauigkeit des Ersatzmodells erklärt wer-

den. Um dieser Abweichung entgegenzuwirken, kann zum Beispiel ein zusätzlicher Sicherheitsfaktor in die Optimierung integriert werden. Um die Genauigkeit des Ersatzmodells zu verbessern kann auch die Ordnung der Polynome bei der PCE erhöht werden. Werden die Berechnungsergebnisse dann auch am Prüfstand bestätigt, kann mit der neuen Tolerierung das Gewichtspotential aus der Entwurfsphase in die weiteren Entwicklungsphasen übertragen werden.

6. Zusammenfassung und Ausblick

Mit der Robustheitsanalyse ist es möglich geworden, die Robustheit eines Bauteils zu ermitteln. Das Leichtbaupotential, dass durch rechnergestützte Methoden in der Entwurfsphase generiert wird, kann zudem durch die Optimierung der Toleranzbereiche in die weiteren Entwicklungsschritte übertragen werden. Mithilfe aufwandsreduzierter Versuchsplanung werden die Designparameter identifiziert, die einen signifikanten Einfluss auf die relevanten Eingangsgrößen haben. Der Einfluss einzelner Parameter, sowie komplexere Wechselwirkungen werden mithilfe eines analytischen Ersatzmodells quantifiziert. Dieses wird dazu genutzt, die Toleranzbereiche der Eingangsgrößen so zu optimieren, dass alle Anforderungen in jeder erlaubten Konfiguration erfüllt werden. Zeit- und gewichtsintensive Anpassungen des Bauteils entfallen und ein leichtes und robusteres Bauteil kann in einem schnellen und effizienten Prozess entwickelt werden. Zudem entfallen Prototypenschleifen, die durch die nicht-Erfüllung der Anforderungen am Prüfstand verursacht werden. Eine Toleranzoptimierung am Beispiel des Hilfsrahmens zeigt, welche Ergebnisse mit dieser Methode erzielt werden können. Allerdings wird hier auch deutlich, dass die Genauigkeit des analytischen Ersatzmodells noch Verbesserungspotential bietet. Neben der Möglichkeit, die Ordnung der Polynome in der Modellierung zu erhöhen, kann die Genauigkeit zum Beispiel durch eine lokale Näherung der Ausgangsgrößen im Bereich der Grenzen der Anforderungen erhöht werden. Die zunächst für Blechbauteile entwickelte Methode kann nun auch für Gussbauteile angepasst werden, um die gesamte Bandbreite der automobilen Bauteile abdecken zu können. Ein weiteres Potential bietet die Zielfunktion der Optimierung. Anstatt die Gesamtgröße der Toleranzbereiche zu maximieren, kann zum Beispiel die Größe des kleinsten Toleranzbereichs maximiert werden. Mithilfe eines Kostenmodells kann außerdem eine Verbindung zwischen der Größe der Toleranzbereiche und der Fertigungskosten erstellt werden. Durch die Integration eines solchen Modells als Zielfunktion in die Optimierung kann Übertragbarkeit in den Werkzeugbau erheblich verbessert werden.

Literatur

- [1] Bestendaufnahme Leichtbau in Deutschland, VDI-Kurzstudie Projekt I C 4-10/15, Verein Deutscher Ingenieure, Berlin, 2015
- [2] Schumacher, A.: Optimierung mechanischer Strukturen – Grundlagen und industrielle Anwendungen, 2. Aufl. 2013. Berlin, Heidelberg: Springer-Verlag Vieweg 2013
- [3] Taguchi, G.: Introduction to Quality Engineering: Designing Quality into Products and Processes. Tokyo: Quality Resources (Asian Productivity Organization) 1986
- [4] Phadke M.S.: Quality Engineering using Robust Design. Upper Saddle River: Prentice Hall 1989
- [5] Dimitriou, P.; Peng, Z.; Lemon, D.; Gao, B. et al.: Diesel Engine Combustion Optimization for Bio-Diesel Blends Using Taguchi and ANOVA Statistical Methods, SAE Technical Paper 2013-24-0011, 2013
- [6] Güven O.: Application of the Taguchi method for parameter optimization of the surface grinding process. Materials Testing: Vol. 57, No. 1, 2015, 43-48
- [7] Trosset, M. W.: Taguchi and Robust Design. Technical report 96-31. Department of Computational & Applied Mathematics. Rice University. 1997
- [8] DIN-EN-ISO-1101: Geometrische Produktspezifikation (GPS) – Geometrische Tolerierung – Tolerierung von Form, Richtung, Ort und Lauf (ISO 1101:2004); Deutsche Fassung EN ISO 1101:2005. 2008
- [9] Klein B.: Versuchsplanung-DOE: Einführung in die Taguchi/Shainin-Methodik, 4. Auflage. Oldenburg: Walter de Gruyter GmbH & Co KG 2014
- [10] Siebertz, K.; van Debber, D.; Hochkirchen, T.: Statistische Versuchsplanung – Design of Experiments (DoE). Berlin, Heidelberg: Springer-Verlag 2010
- [11] Janon, A.: Analyse de sensibilité et réduction de dimension - Application à l'océanographie. Université de Grenoble Dissertation. 2006
- [12] Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Saisana, M.; Tarantola, S.: Global Sensitivity Analysis. The Primer. Chichester: John Wiley & Sons, 2008
- [13] Myers, R. H.; Montgomery, D.C.: Response Surface Methodology: Process and Product Optimization Using Designed Experiments. New York: Wiley 1995
- [14] Kalla, S.: Use Of Orthogonal Arrays, Quasi-Monte Carlo Sampling, And Kriging Response Models For Reservoir Simulation With Many Varying Factors. Regional Engineering College Master Thesis. Warangal 2002

- [15] Loh, W.: On Latin Hypercube Sampling. *The Annals of Statistics*, Vol. 24, No. 5, 1996, 2058-2080
- [16] Box, G. E. P.; Wilson, K.B.: On the Experimental Attainment of Optimum Conditions (with discussion). *Journal of the Royal Statistical Society Series B* 13(1), 1951, 1–45
- [17] Xiu, D.; Karniadakis, G.E.: The Wiener-Askey polynomial chaos for stochastic differential equations. *Journal of Scientific Computing* 24(2), 2002, 619–644
- [18] Kewlani, G.; Crawford, J.; Iagnemma, K.: A polynomial chaos approach to the analysis of vehicle dynamics under uncertainty. *Vehicle System Dynamics*, 2012, 1-26
- [19] Rosic, B.; Matthies H.G.: Computational Approaches to Inelastic Media with Uncertain Parameters. *Journal of Serbian Society for Computational Mechanics* 2(1), 2008, 28-43
- [20] Zein, S.; Colson, B.; Glineur, F.: An Efficient Sampling Method for Regression-Based Polynomial Chaos Expansion. *Communications in Computational Physics* 13, 2013, 1173-1188
- [21] Nelder, J.A.; Mead, R.: A simplex method for function minimization. *Computer Journal* 7, 1965, 308-313
- [22] Bertsekas, D.P.: *Constrained Optimization and Lagrange Multiplier Methods*. London: Academic Press 1982

Lebensdaueranalyse auf Basis von multidimensionaler Zuverlässigkeits- und Robust Design Simulation



Integrale Betrachtung der Robusten Zuverlässigkeit

Dipl.-Ing. **S. Kemmler**, M.Sc. **M. Dazer**, Prof. Dr.-Ing. **B. Bertsche**,
Institut für Maschinenelemente, Universität Stuttgart; TTI GmbH;
Dr.-Ing. **T. Lepold**,
Knorr-Bremse, Systeme für Nutzfahrzeuge GmbH, München

Kurzfassung

Konventionelle Methoden des *Robust Design* verfolgen größtenteils Ansätze zur Varianzbetrachtung, die potentiell über der Produkteinsatzzeit auftreten können. Daher geben diese Methoden keine genauen Informationen über das zeitlich funktionale Verhalten der sich verändernden Produkteigenschaften und –anforderungen sowie deren Lebensdauer. Um genaue Lebensdauerprognosen bezüglich dieser Funktionserfüllung zu beschreiben, ist eine kontinuierliche Merkmalsänderung in Form von sogenannten Degradationsmodellen von Vorteil. Diese werden im Allgemeinen durch reale Versuche mit einem hohen Grad an zeitlichem und kostenintensivem Aufwand durchgeführt. Für eine effizientere Ermittlung der Modelle sollten bereits in frühen Phasen des Produktentwicklungsprozesses, virtuelle Degradationsmodelle entwickelt werden. Durch die genaue Kennung von Funktionsausfällen über der Zeit können nicht nur Produkte zielgerichtet ausgelegt, sondern auch unnötige Ressourcen eingespart werden.

Abstract (optional)

Conventional methods of *Robust Design* are largely used for variance consideration, which may potentially occur during the product use. Therefore, these methods do not provide accurate information about the temporal behavior of the changing product properties and requirements. However, in order to describe accurate life predictions with regard to the fulfilling functions, a continuous feature change is advantageous with the aid of degradation models. These models are generally determine with real tests, which are associated with a high degree of time and cost effort. For a more efficient determination of these models, virtual degradation models should be developed in the early stages of the product development process. The products can be targeted designed and even unnecessary resources can be saved by the precise identification of functional failures over time.

1. Einleitung und Motivation

Die Lebensdauer eines Produkts gilt als wichtiges Qualitätsmerkmal. Sie markiert nicht nur die Beschaffenheit und den Zustand eines Produktes, sondern schließt auch den zuverlässigen und sicheren Betrieb innerhalb eines bestimmten Zeitrahmens ein. Wird die Ausfallrate an technischen Bauelementen in Abhängigkeit von deren Lebensdauer betrachtet, ergibt sich ein typischer Verlauf, die sogenannte Badewannenkurve. Das Ausfallverhalten wird hierbei in drei Phasen klassifiziert: Früh-, Zufalls- und Ermüdungsausfälle [1]. Bei den Ermüdungsausfällen altert das Bauteil und repräsentiert eine steigende Ausfallrate. Der Ausfallzeitpunkt kann anhand einer messbaren Degradation in der realen Erprobung beziehungsweise mit virtuellen Degradationsmodellen bereits früher im Entwicklungsprozess prognostiziert werden. Diese Modelle ermöglichen dem Entwickler einerseits eine genauere Berechnung der Funktions-Lebensdauer und andererseits eine Möglichkeit Schwachstellen im System zu ermitteln. Bild 1 zeigt schematisch einen typischen Degradationsverlauf bis zu einem definierten Ausfallkriterium (*End-of-life (EOL)* - Kriterium).

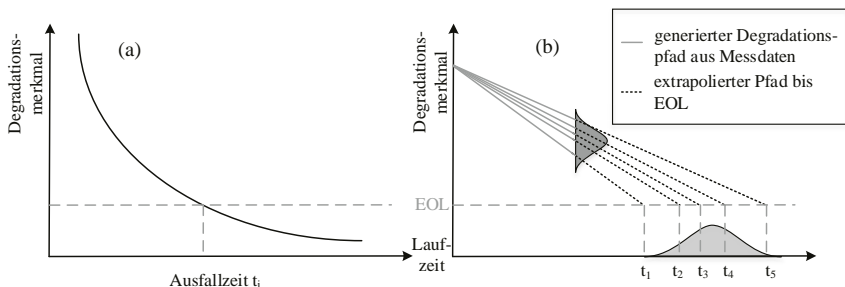


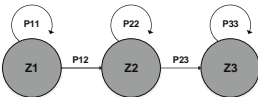
Bild 1: Änderung der Degradation über der Betriebszeit (a) und Zusammenhang zwischen Degradationspfaden und Ausfalldichte (b)

Erste Degradationsmodelle basieren auf der Annahme, dass die Degradation ein zufälliger, stochastischer Prozess (*Wiener Prozess*) in einer definierten Zeit ist [2]. Für die Bestimmung dieser Modelle beziehungsweise die Prognose des Ausfallverhaltens der zu beschreibenden Aufgabenstellung, kann mittels statistischer Modelle erfolgen. Die Modellierung erfolgt auf Basis der Annahme einer Funktion über der Zeit mit multidimensionalen zufälligen Variablen (*General Degradation Path Models*). Beispiele hierfür sind unter anderem die zweistufige Bestimmung der Funktionsparameter (*Nonlinear mixed-effects model*) nach [3], das fraktionell faktorielle Design nach [4] oder das Model mit zufälligen Regressionskoeffizienten nach [5]. Weitere Modelle sind in Tabelle 1 gelistet.

Gegenüber der herkömmlichen Erprobung haben diese Modelle den wirtschaftlich entscheidenden Vorteil einer beschleunigten Erprobung. Dies hat beispielsweise bei Bauteilen mit einer hohen Lebensdauer, unabhängig von deren Zuverlässigkeit, den Vorteil, dass durch

herkömmliche Lebensdauertests kaum oder keine Ausfälle generiert werden können und anhand der Degradationserprobung neben den Ausfallzeiten auch der Verlauf über der Zeit bis zum Ausfall beschrieben werden kann. Dies ermöglicht eine variable Messzeitpunkt-Definition sowie ein vorzeitiges Beenden der Versuchsreihen nach Verhaltenskenntnis sowie deren Extrapolation. Zudem muss durch die Definition eines *EOL*-Kriteriums nicht zwingend bis zum Ausfall getestet werden.

Tabelle 1: Degradationsmodellen nach Stand der Technik und Forschung

Parametrisch	linear	Arrhenius	$y_{ij}(t, T) = \alpha_{ij} - t\beta_i \cdot \exp(-\frac{\gamma}{T})$	[6]
		Random-Coefficients	$y_{ij}(t) = \alpha_{ij}[1 - \exp(-(\beta_i(t))^{\gamma}) + \epsilon_{ij}$	[7]
		Simple-Constant-Rate	$y_{ij}(t) = \alpha_{ij} - \beta_i \cdot t$	[6]
		Simple Path	$\log(y_{ij}(t)) = \log(\beta_i) + \log(t) + \epsilon_{ij}$	[8]
		Random-Regression-Coefficients	$y_{ij}(t) = \alpha_{ij} + \beta_i \log(t_{ij}) + \epsilon_{ij}$	[5]
		Sample-Degradation-Path	$y_{ij}(t) = D_{ij} + \epsilon_{ij}$	[6]
	nicht-linear	Random-Effects	$y_{ij}(t) = D(t, \beta_{ij}) + \epsilon_{ij}$	[9]
		Mixed-Effects	$y_{ij}(t) = D_{ij} + \epsilon_{ij} = D(t; \alpha_{ij}, \beta_i) + \epsilon_{ij}$	[3]
		Integrated S-Curve	$y_{ij}(t) = \alpha_{ij}F(t, \zeta_{ij}) + \epsilon_{ij}(t)$	[10]
		Gompertz	$y_{ij}(t) = \exp(-ce)^{\beta_i t}$	[11]
		Lloyd-Lipow	$R_t = R_{\infty} - \frac{\beta_i}{t}$	[12]
		Potenz	$y_{ij}(t) = \beta_i t^m$	[13]
		Exponentiell	$y_{ij}(t) = \exp(\beta_i t)$	
		Logarithmisch	$y_{ij}(t) = \ln(\beta_i t)$	
Stochastisch	Gamma-Prozess		$y_{ij}(t) = r_0 - D_{ij}$ $f_{X(t)}(x) = Ga(x \lambda(t), \beta_i)$	[14]
	Wiener-Prozess		$y_{ij}(t) = \Lambda(t) + \sigma W(\Lambda(t))$	[2]
	Markov-Prozess			[15]

Zusammenfassend gilt, dass die Degradation einen Ausfall hervorruft, sobald sie den kritischen Degradationswert D_f (*EOL*) erreicht. Infolgedessen hängt die Ausfallwahrscheinlichkeit $F(t)$ im Wesentlichen davon ab, mit welcher Wahrscheinlichkeit P

die Degradationswerte y_i über oder unter dem *EOL*-Kriterium liegen. Hierfür ist die Verteilungsform des Degradationsmerkmals y wichtig, vergleich dazu Bild 1 (b) mit folgender Gleichung:

$$P(t > t_f) = F(t) = F(t, y) = P(D(t, y)) > D_f.$$

Für eine sichere und genauere Aussage einer Prognose auf Basis von Stichproben hinsichtlich Streuung und Regressionsgüte gibt es bislang mehrere Ansätze zur Modellierung der Aussagesicherheit. Zur Verfügung stehen klassische Regressionsmethoden sowie analytische und numerische Methoden zur Verfügung. Hierfür wird anhand theoretischer Ausfallzeiten, welche sich aus dem parametrisierten Degradationsmodell ergeben, eine konventionelle Lebensdauerauswertung durchgeführt. Eine Übersicht gibt hierzu Bild 2.

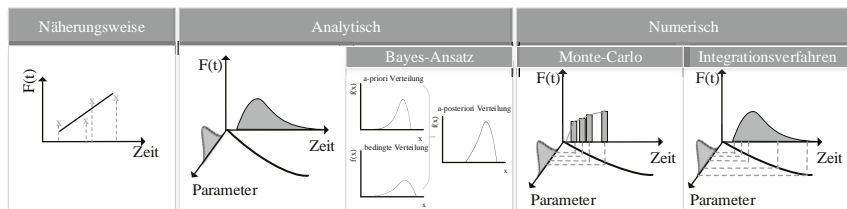


Bild 2: Zuverlässigkeitstechnische Auswertungsmöglichkeiten von Degradationsmodelle

Als weitere Möglichkeit wird in diesem Beitrag eine simulative Alternative gezeigt, die nicht nur die Modellierung der Degradation in der frühen Entwicklungsphase ermöglicht, sondern zudem zur Bestimmung der Aussagesicherheit dient. Hierfür werden stochastische Lebensdauerberechnungen bezüglich Funktionsausfällen mit Hilfe von simulativ ermittelten kontinuierlichen Merkmalsänderungen anhand einer Überlastkupplung einer Nachstelleinheit für Nutzfahrzeugsbremssystemen vorgestellt. Dabei werden die schädigungsrelevanten Effekte, wie Verschleiß, hinsichtlich der Funktionsdegradation mit *Finite Element (FE)*-Berechnungen ermittelt. Damit ist es möglich Zuverlässigkeitsprognosen für funktionales Versagen zu prognostizieren.

2. Allgemeine Vorgehensweise

Basierend auf der bekannten Vorgehensweise in Kapitel 2.1 wird als Vergleich die in diesem Beitrag zu Grunde liegende Vorgehensweise vorgestellt.

2.1 Klassische Vorgehensweise

In Bild 3 wird eine allgemeine Vorgehensweise gezeigt. Zu Beginn müssen in der ersten Phase Produkteigenschaften für typische Belastungen, wie beispielsweise Nutzungszyklen, Temperatur, Druck oder Strom, zur Degradationsbeschreibung ermittelt werden. Im Anschluss werden Zeitintervalle sowie *EOL*-Kriterien bestimmt und definiert. In der zweiten Phase wird das Modell auf der generierten Datenbasis sowie im Abgleich zu bestehenden Modellen erstellt. Jeder Degradationsprozess benötigt ein eigenes Modell nach dem jeweiligen

degradierten Merkmal, wie Risswachstum oder Widerstandsänderung. Im Anschluss erfolgen die Abschätzung der Parameter sowie die Evaluierung der Ausfallwahrscheinlichkeit.

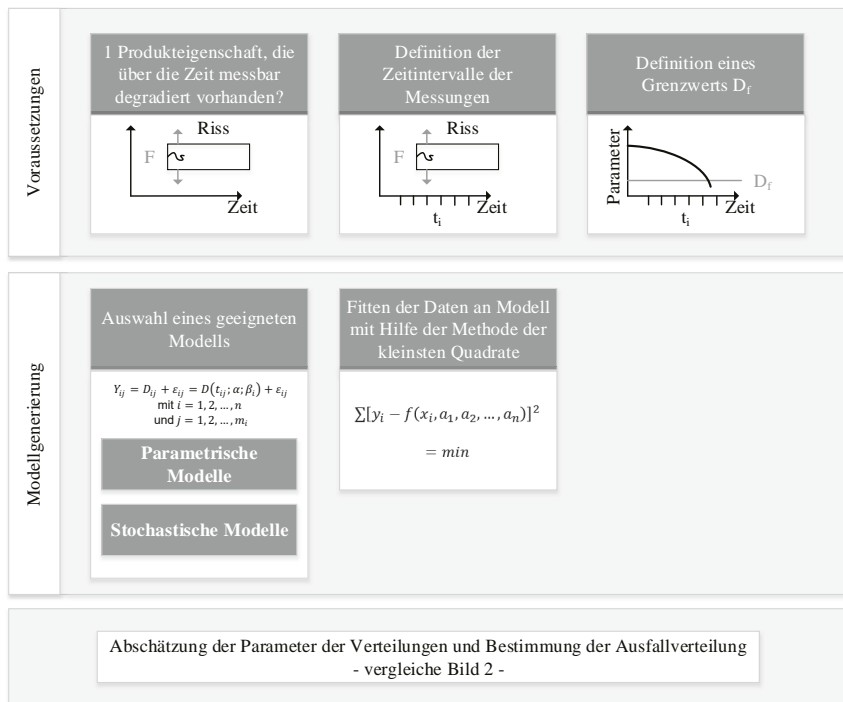


Bild 3: Schematische Vorgehensweise zur Degradationsmodellierung

2.2 Lebensdaueranalyse auf Basis von multidimensionaler Zuverlässigkeits- und Robust Design Simulation

Die Lebensdaueranalyse soll auf Basis des *Robust Design* erfolgen. Der Grundgedanke dabei ist, das System oder die Komponente gegen äußere und innere Störeinflüsse unempfindlich zu gestalten. Hierbei wird nach der klassischen Vorgehensweise nicht explizit die Betriebszeit als eine steuernde Größe betrachtet, sondern als Störgröße, da das Produkt in der geforderten Betriebszeit seine Funktion aufrecht halten soll. Allerdings treten während der Einsatzzeit alterungsbedingte Ermüdungen am Produkt auf. Um das Produkt möglichst robust gegen diese Ermüdungen zu konstruieren, bedarf es eines Designs, das innerhalb der geforderten Einsatzzeit keine beziehungsweise bis zu einem zulässigen Kriterium die Funktion gewährleistet. Diese Herausforderung wird meistens unter der Prämisse einer Überdimensionierung erreicht. Allerdings kann selbst diese unwirtschaftliche Konstruktion außerhalb der Funktionsgrenzen liegen. Um dies zu vermeiden wird entsprechend der Theorie

der *Robusten Zuverlässigkeit* nach [16] eine zeitliche Betrachtung des Funktionsverhaltens unter Berücksichtigung sämtlicher Wechselwirkungen verfolgt. Zur Lebensdaueranalyse auf Basis von multidimensionaler Zuverlässigkeits- und *Robust Design* Simulation wird mittels einer stochastischen Degradationsanalyse die Funktionserfüllung innerhalb der Einsatzzeit betrachtet. Großteils sind solche virtuellen Degradationssimulationen mit einem Mehraufwand verbunden, die eine Effizienz infolge hoher Ressourcenaufwand von beispielsweise Rechenzeit und –kapazität nicht immer gewährleisten. Folglich wird nach der durchgeführten *Robust Design* Definition die 5 % - und 95 % - Quantil-Designs bestimmt sowie ausschließlich mit diesen und dem nominalen Design die Degradation über der Einsatzzeit simuliert und die Ausfallzeit zum *EOL* betrachtet. Ziel ist es, mit den drei durchgeführten Simulationen nicht nur Rechenzeit einzusparen, sondern auch eine Aussagesicherheit respektive für die Population zu bieten. Entsprechend wird hierfür folgende Hypothese aufgestellt:

Die definierten Quantil-Designs aus den Robusten Designs entsprechen zum Zeitpunkt des EOL-Kriteriums der Aussagesicherheit zum jeweiligen Lebensdauer-Quantil.

Zur Überprüfung dieser These wird am nachfolgend beschriebenen Demonstrator das *Robust Design* sowie die virtuelle Degradationssimulation der drei genannten Designs durchgeführt. Zur Bestätigung wird die Simulation, nach erfolgter Validierung mittels realer Degradationsversuche, um die durchgeführte Parameterstudie im *Robust Design* auf die Degradationsmodellierung übertragen. Ziel ist es, die Quantil-Designs in ihrer Definition zu bestätigen. Nach erfolgreicher Bestätigung der Hypothese können somit nicht nur effektive virtuelle Lebensdaueranalysen umgesetzt, sondern auch daraus zielgerichtet reale Nachweisversuche ressourcenreduzierend abgeleitet und geplant werden.

3. Verschleißmodellierung des Demonstrators

In der vorliegenden Funktionsbaugruppe als Anwendungsbeispiel wird eine Überlastkupplung verwendet. Anschließend an die Vorstellung des Demonstrators wird die Vorgehensweise beschrieben und abschließend die Ergebnisse aus virtueller und realer Erprobung diskutiert.

3.1 Das Anwendungsbeispiel

Die Überlastkupplung besteht aus Nabe (1), Flanschring (2) sowie aus rotationssymmetrisch angeordnete Kugeln (3), die durch ihre Kontakteigenschaften das Drehmoment übertragen, vergleiche Bild 4. Bei einer korrekten Funktionsweise übertragen die Kugeln bis zu einem definierten Grenzwert das Drehmoment von der Nabe (1) auf den Flanschring (2). Im Überlastfall rotieren Nabe (1) und Flanschring (2) relativ zueinander, was die Kugeln (3) aus ihren Rastpunkten (RP) drückt. In diesem Zustand kann kein Drehmoment übertragen werden,

vergleiche Bild 4 (b). Die Drehbewegung wird in Folge der Überlastfunktion in eine axial translatorische Bewegung entkoppelt. Das Robustheitsmaß (Zielgröße) ist die Funktion innerhalb eines übertragbaren, normierten Drehmoments von $1,0 \pm 0,2$ [-].

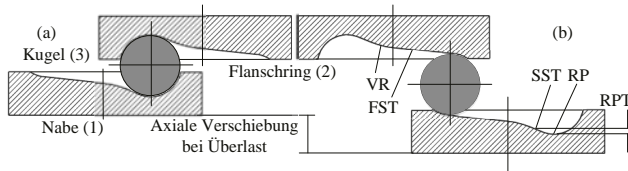


Bild 4: Überlastkupplung im unbetätigten (a) und betätigten (b) Zustand

3.2 Degradationsermittlung

Die Umsetzung der Verschleißsimulation findet mittels *FE* statt. Der Grundgedanke besteht

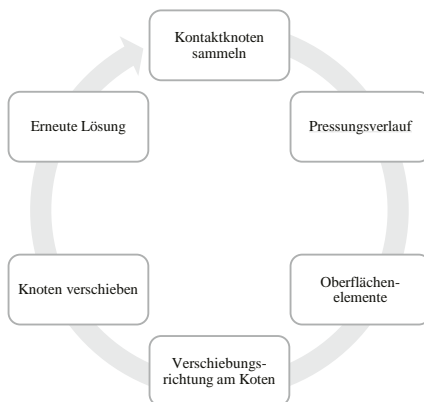


Bild 5: Schematische Vorgehensweise zur Degradationssimulation

darin, alle Knoten des *FE*-Netzes zu sammeln, die während der gesamten Bewegung Kontakt mit den Kugeln erfahren. Diese Knoten sollen in die Wirkrichtung der Pressung in das Bauteilinnere verschoben werden. Die Vorgehensweise ist in Bild 5 aufgezeigt. Der Verschleiß in Form von Materialverlust wird durch Netzknoten-Verschiebung realisiert. Mit den aus der *FE*-Simulation erhaltenen Pressungswerten kann eine einfache Simulation des Materialabtrags aufgrund von adhäsiven Verschleiß modelliert werden. Hierfür werden Derivate des Archard'schen Verschleißmodells nach [17,18] angewandt. Weitere Informationen zur Umsetzung und Anwendung der Verschleißsimulation siehe [19].

3.3 Simulationsergebnisse

Im Folgenden werden die Ergebnisse nach dem *Robust Design*, der Verschleißsimulation der Bahnkontur und der Validierung der Überlastkupplung kategorisiert vorgestellt.

3.3.1 Robust Design der Überlastkupplung

Beginnend mit dem *Robust Design* der Überlastkupplung ist ein Design definiert, welches mit seiner Toleranzdefinition die maximal geforderten Grenzen von $1,0 \pm 0,2 [-]$ (normiert) unter Einwirkung innerer sowie äußerer Störgrößen einhält. Dies zeigt die dargestellte Parameterstudie, vergleiche Bild 6 (a).

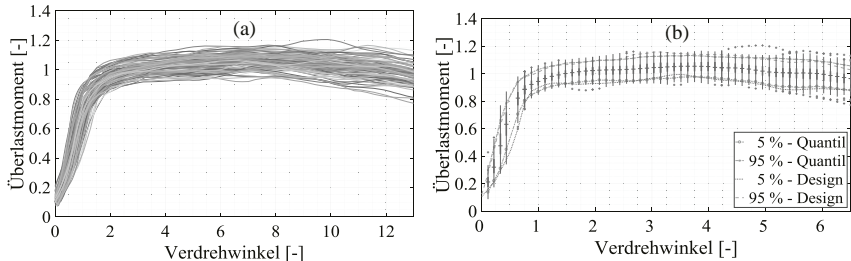


Bild 6: Überlastmoment der Kupplung: *Robust Design* (a), Quantile sowie deren Designs (b)
Für eine robuste Aussagesicherheit werden für die anschließende virtuelle Degradationsmodellierung zuerst die Verläufe der Überlastmomente des *Robust Design* diskretisiert und je Diskretisierungspunkt x ein $p = 5 \%$ und $p = 95 \%$ - Quantil mittels *Linearer Interpolation* für $n = 150$ Designs aus Bild 6 (a) bestimmt:

$$x_{[(n-1)p+1]} + ((n-1)p+1 - [(n-1)p+1])(x_{[(n-1)p+1]} + 1 - x_{[(n-1)p+1]})$$

Für die Zuordnung des jeweiligen Quantil zum entsprechenden Design, wird die Methode der *Kleinsten Fehlerquadrate* der deskriptiven Statistik angewendet:

$$Q(\beta_0; \beta_1) = \sum_{i=1}^n (p_i - \hat{p}_i)^2 = \sum_{i=1}^n (p_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2,$$

substituiert mit:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(p_i - \bar{p})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad \text{und} \quad \hat{\beta}_0 = \bar{p} - \hat{\beta}_1 \bar{x}.$$

Als Ergebnis ergibt sich aus der minimalen Fehlerquadratsumme aller diskretisierten Abstandsberechnungen der in Bild 6 (b) dargestellte Zusammenhang zwischen Quantil und resultierenden Design.

3.3.2 Degradationsergebnisse der Überlastkupplung

Wird die Degradation der Bahnkontur beider Scheiben in Bezug auf das Überlastmoment in Abhängigkeit des Verdrehwinkels aufgetragen, vergleiche Bild 7, beispielhaft für sechs Simulationen, ergibt sich mit zunehmenden Verschleiß und verschiedenen Parameterkombinationen an der Bahnkontur ein stetig zunehmendes (a) und ein stetig abnehmendes (b) Überlastmoment sowie ein sich änderndes Verhalten der Merkmale.

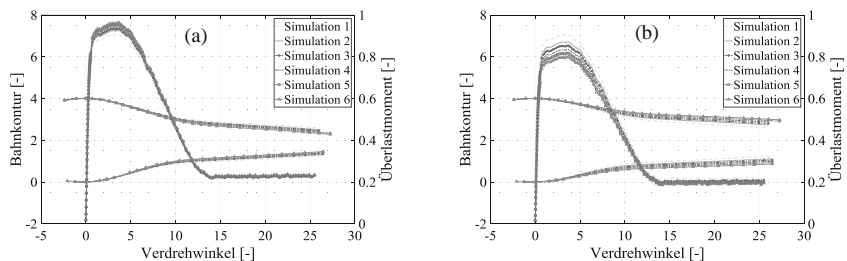


Bild 7: Degradation der Bahnkontur in Bezug auf das resultierende abnehmende (a) und zunehmende (b) Überlastmoment

Der Grund hierfür wird durch die Betrachtung der Degradationsmerkmale der Bahnkontur (*Rastpunkt-Radius* und *-Tiefe*, *Übergangsradius* sowie *steile* und *flache Steigung*) nachvollziehbar, vergleiche Bild 8. Es zeigt sich in allen Merkmalen eine durch die Pressung und Reibarbeit bedingte Abnahme des 95 % Designs, die ein erschwelter Übergang zur rotatorischen Entkopplung indiziert und folglich zu einem zunehmenden Überlastmoment führt. Die Abnahme des 5 % und des Nominal-Designs ergibt sich aufgrund deren Parameterkombination, die eine wechselseitige Beeinflussung auf die Merkmale haben.

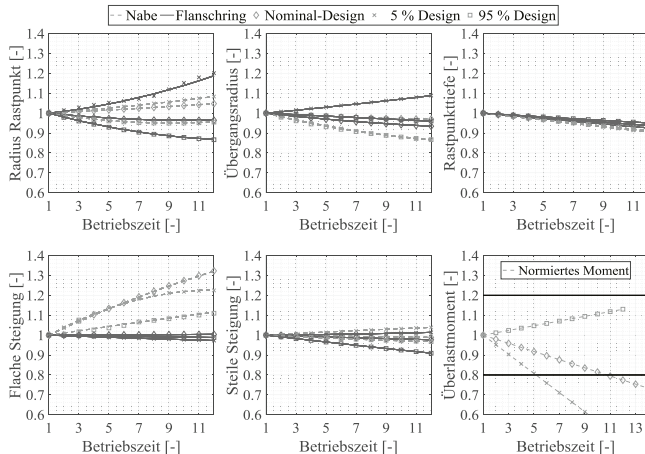


Bild 8: Normierte Degradation des Nominal- und der Quantil-Designs der Bahnkontur-Parameter der Nabe und des Flanschrings und des maximalen Überlastmoments

3.3.3 Validierung der Überlastkupplung

Mittels den vorhandenen Testdaten von sechs Nachstellern ist es möglich, eine erste Modellauswahl zu treffen. Nach erster Einschätzung beschreibt der lineare Modellansatz

(*Simple-Constant-Rate*) nach [7], vergleiche Tabelle 1, den Degradationspfad. Nach dem festgelegten *EOL*-Kriterium können via Monte-Carlo (*MC*)-Simulation und Extrapolation zur Generierung weiterer Degradationspfade der Daten sogenannte *Pseudo-Ausfallzeiten* generiert werden, vergleiche Bild 9. Eine Validierung des Degradationsmodells von realer und virtueller Erprobung kann aus dem Vergleich des normierten Überlastmoments aus Bild 8 und Bild 9 (a) geschlossen werden. Beide Modelle stimmen sowohl im Modell als auch in deren zeitlichen Verhalten überein, vergleiche hierzu vorab das Überlastmoment in Verhältnis zum Nominal-Design in Bild 10.

4. Virtuelle Degradationsmodellierung und Ergebnisinterpretation

Die Schlussfolgerung aus der Simulation zeigt sich in der Validierung der sechs Nachsteller sowie in der *MC*-Simulation, vergleiche Bild 9 und Bild 10. Die realen Versuche weisen sowohl zunehmende als auch abnehmende Überlastmomente auf. Der Streubereich verdeutlicht diesen Trend. Werden die *EOL*-Kriterien betrachtet ergibt sich eine B_5 - Lebensdauer beziehungsweise ein 5 % - Quantil für die untere Ausfallgrenze bei einer normierten Betriebszeit von 4,06 [-] und eine obere Ausfallgrenze von 6,81 [-].

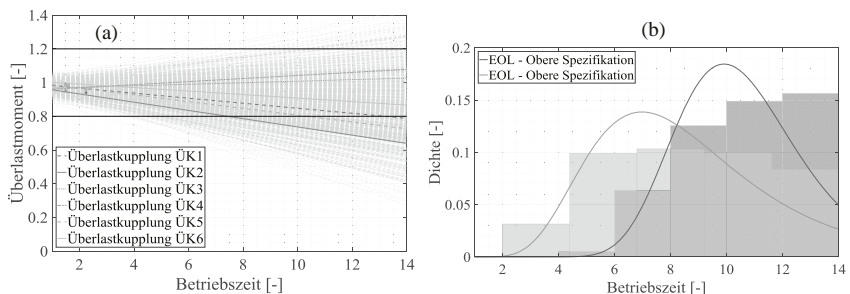


Bild 9: Extrapolation des Überlastmodells zum *EOL* mittels *MC*-Simulation

Zur Überprüfung der Hypothese aus Kapitel 2.2 wird auf Basis der Quantil-Designs sowie dem Nominal-Design der virtuellen Degradations-Simulation eine Aussagesicherheit zum Zeitpunkt des *EOL* prognostiziert. Bild 10 zeigt die Simulation des jeweiligen Momentenverlaufs. Es sind sowohl die Quantil- und Nominal-Designs der *MC*-Simulation auf Basis der realen Versuche aus Bild 9 (a), als auch die virtuelle Degradations-Simulation zusammengefasst. Es wird deutlich, dass alle Designs innerhalb der Degradation der Quantil-Designs liegen beziehungsweise mit konservativeren Ansatz diese umhüllen. Es ergibt sich ein 5 % - Quantil für die untere Ausfallgrenze bei einer normierten Betriebszeit von 3,95 [-] und eine obere Ausfallgrenze von 6,05 [-]. Dies sind im Vergleich zur *MC*-Simulation auf Basis der realen Versuche eine um 3,5 %ige beziehungsweise um 12,6 %ige konservativere Betrachtung.

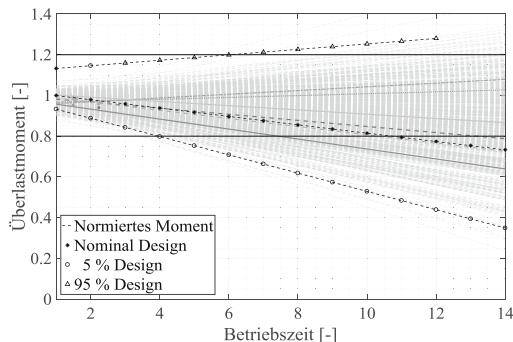


Bild 10: Virtuelle Degradation mit Design-Grenzen als Aussagesicherheit im Vergleich zur realen Erprobung

Die aufgestellte Hypothese zur Definition von Quantil-Designs aus den *Robust Designs* einer virtuellen Degradations-Simulation umfassen zum Zeitpunkt des *EOL*-Kriteriums die Aussagesicherheit zum jeweiligen Quantil, kann folglich bestätigt werden. Entsprechend ist eine integrale Betrachtung der *Robusten Zuverlässigkeit* des vorgestellten Anwendungsbeispiels gegeben.

6. Zusammenfassung

In diesem Beitrag wird eine Lebensdaueranalyse auf Basis von multidimensionaler Zuverlässigkeits- und *Robust Design* Simulation am Beispiel einer Überlastkupplung für Nutzfahrzeugbremsysteme vorgestellt. Die zu Beginn durchgeführte *Robust Design* Untersuchung beschreibt die Quantil- und Nominal-Designs zum Auslegungszeitpunkt vor Inbetriebnahme. Die anschließende virtuelle Degradations-Simulation zeigt, aufgrund der bestehenden Wechselwirkung der Degradationsmerkmale auf den Verlauf der Degradationspfade, ein sich zu- und abnehmendes Momentverhalten. Dieses wird durch die anschließenden realen Versuche und der darauf aufbauenden *MC*-Simulation unter Beweis gestellt. Diese realen Degradationspfade verlaufen innerhalb der Quantil-Designs. Dies bedeutet einerseits eine Bestätigung der aufgestellten Hypothese sowie eine ausreichende und konservative Abdeckung der virtuellen Degradations-Simulation anhand von Quantil- und Nominal-Designs. Dies erlaubt dem Entwickler frühzeitig eine robuste und zuverlässige Auslegung der Produkte sowie eine gezielte Versuchsplanung.

Literaturverzeichnis

- [1] Bertsche, B.; Lechner, G.: Reliability in Automotive and Mechanical Engineering. Berlin, Heidelberg, New York: Springer-Verlag, 2007.
- [2] Doksum, K. A.: Degradation Rate Models for Failure Time and Survival Data. CWI Quart, Nr. 4, S. 195-203, 1991.

- [3] Meeker, W.; Escobar, L.: Using degradation measures to estimate a time-to-failure distribution. *Technometrics*, 40. Jg., Nr. 2, S. 89-99, 1998.
- [4] Hamada, M.: Using statistically designed experiments to improve reliability and to achieve robust reliability. *IEEE Transactions on Reliability*, 44. Jg., Nr. 2, S. 206-215, 1995.
- [5] Lu, J.-C.; Park, J.; Yang, Q.: Statistical inference of a time-to-failure distribution derived from linear degradation data. *Technometrics*, 39. Jg., Nr. 4, S. 391-400, 1997.
- [6] Nelson, W.: Accelerated testing: statistical models, test plans, and data analysis. New Jersey: John Wiley & Sons, Inc. 2004.
- [7] Carey, M.; Koenig, R.: Reliability assessment based on accelerated degradation: A case study. *IEEE Transactions on Reliability*, 40. Jg., Nr. 5, S. 499-506, 1991.
- [8] Lu, J.-C.; Meeker, W.; Escobar, L.: A comparison of degradation and failure-time analysis methods for estimating a time-to-failure distribution. *Statistica Sinica*, S. 531-546, 1996.
- [9] Chiao, C.-H.; Hamada, M.: Analyzing experiments with degradation data for improving reliability and for achieving robust reliability. *Quality and Reliability Engineering International*, 17. Jg., Nr. 5, S. 333-344, 2001.
- [10] Boulanger, M.; Escobar, L.: Experimental design for a class of accelerated degradation tests. *Technometrics*, 36. Jg., Nr. 3, S. 260-272, 1994.
- [11] Winsor, C.: The Gompertz curve as a growth curve. *Proceedings of the national academy of sciences*, 18. Jg., Nr. 1, S. 1-8, 1932.
- [12] Reliasoft: Reliability growth & repairable system data analysis. Reliasoft Corp., 2015.
- [13] Davidovic, V.: Reliability Physics and Engineering: Time-to-Failure Modeling, JW McPherson. Springer, 2012.
- [14] Pandey, M.; Yuan, X.; Van Noortwijk, J.: Gamma process model for reliability analysis and replacement of aging structural components. *Proceedings ICOSAR, Rome*, 2005.
- [15] O'Connor, P.; Kleyner, A.: Practical Reliability Engineering. New Jersey: John Wiley & Sons, Inc., 2012.
- [16] Kemmler, S.; Eifler T.; Bertsche B.: Development of robust, reliable products and processes - using the methodology of Robust Reliability (in print).
- [17] Popov, V.: Kontaktmechanik und Reibung. Heidelberg: Springer-Verlag, 2010.
- [18] Thompson, J.; Thompson, M.: A Proposal for the Calculation of Wear. *International ANSYS Users Conference & Exhibition*, 2006.
- [19] Dazer, M.; Kemmler, S.; Bertsche, B.: Wear simulation of an overload clutch. *WEAR*, 2017 (in print).

Zustandsbasierte Sicherheits- und Verfügbarkeitsanalyse unterschiedlich automatisierter Fahrfunktionen mittels Monte-Carlo-Simulation



Johannes Heinrich, M.Sc., Dr.-Ing. **Fabian Plinke**,
Institut für Qualitäts- und Zuverlässigkeitsmanagement, Wuppertal;
Dr.-Ing. **Jan Hauschild**,
Institut für Qualitäts- und Zuverlässigkeitsmanagement, Hamburg

Kurzfassung

Automatisierte Fahrfunktionen, die den Fahrer durch die Übernahme der Längs- und Querrichtung des Fahrzeuges aktiv entlasten (Teilautomatisierung), bzw. ihn komplett aus dem Regelkreis nehmen (Hochautomatisierung und Vollautomatisierung), müssen für den operativen Betrieb eine hohe Sicherheit und Verfügbarkeit gewährleisten können. Ein Ansatz für die Absicherung der Systeme sind quantitative Sicherheits- und Verfügbarkeitsanalysen. Hierbei ist jedoch zu beachten, dass das alleinige Analysieren von Systemstrukturen nicht mehr ausreichend ist, da auch Faktoren wie Degradationszustände des automatisierten Systems, die Umwelt und der Fahrer bei den Analysen einbezogen werden müssen. Mit dem in dieser Arbeit vorgestellten Modell, einer Kombination aus Zustandsdiagrammen und zeitbasierter Monte-Carlo-Simulation (MCS), ist es möglich, Sicherheits- und Verfügbarkeitsanalysen für hochautomatisierte Fahrerassistenzsysteme auf einer oberen Systemebene durchzuführen. Grundlage für dieses Modell sind Ergebnisse aus einem Forschungsprojekt mit der Bundesanstalt für Straßenwesen [1], welches mit der Unterstützung von Automobilherstellern und Zulieferern durchgeführt wurde.

1. Einleitung

OEMs und Zulieferer forschen bekanntlich zurzeit intensiv in allen Bereichen zum Thema automatisierte Fahrfunktionen. Eine große Herausforderung dabei wird es sein, Methoden zu finden, die die Funktionalitäten in zukünftigen Systemen derart absichern können, dass durch die Funktionen keine Gefahren für den Fahrer, Beifahrer oder Personen außerhalb des Fahrzeugs entstehen können. Hierzu müsste eine große Anzahl an Einflussfaktoren berücksichtigt werden, was jedoch einen immensen Anstieg der technischen und organisatorischen Ressourcen und eine gravierende Verlängerung der Entwicklungszeit zur Folge hätte. Des Weiteren verfügen viele Einflussfaktoren über keine deterministischen Eigenschaften, womit

eine reine analytische Betrachtung nicht möglich ist. Um ein einheitliches und ausreichendes Sicherheitslevel der Funktionalitäten gewährleisten zu können, müssen definierte Prozessschritte während der Produktentwicklung durchgeführt werden. Hierbei sollten auch alternative Methoden, im Vergleich zu den gängigen heutigen, verwendet werden; z.B. auf theoretischem Level oder anhand von Simulationen. Insbesondere Systeme, bei denen der Fahrer die automatisierten Systeme nicht mehr überwachen muss, müssen diese so sicher und zuverlässig wie möglich ausgelegt werden.

Das in dieser Arbeit präsentierte Modell zeigt ein Konzept, mit dem die Sicherheit und Verfügbarkeit von automatisierten Fahrfunktionen berechnet werden kann. Hierbei wird der Verlauf vom aktivierten System bis zum Erreichen des risikominimalen Zustandes aufgrund von diversen Einflussfaktoren dargestellt.

Die für das Modell verwendeten Zustandsdiagramme werden hierbei um besondere Eigenschaften der Transitionen erweitert. Dabei erfolgt die Analyse mittels Monte-Carlo-Simulation (MCS) im Kontext definierter spezieller Entscheidungsregeln.

Als Einflussgrößen und Systemgrenzen für das Modell wurde das 3-Ebenen Modell von Donges [2] zugrunde gelegt; dieses bildet allgemein den Zusammenhang zwischen Fahrer, Fahrzeug und Umwelt als Regelkreis ab.

2. Automatisierte Fahrerassistenzsysteme

Die Entwicklung von automatisierten Fahrerassistenzfunktionen begann bereits im 20. Jahrhundert. So wurde beispielsweise 1981 der erste Prototyp eines Abstandsregeltempomaten (Adaptive Cruise Control) entwickelt; 1995 folgte die Markteinführung im Modell Diamante von Mitsubishi [3]. Bei allen automatisierten Fahrerassistenzfunktionen, die bis heute in Serienfahrzeugen verbaut wurden, ist der Fahrer für die komplette Fahrzeugführung verantwortlich, somit auch für Handlungen, die die automatisierte Fahrerassistenzfunktion initiiert. Für zukünftige Systeme wird diese Eigenschaft entfallen. D.h. dem Fahrer wird es möglich sein, sich vom Verkehrsgeschehen abzuwenden und anderen Tätigkeiten nachzugehen. Durch den steigenden Automatisierungsgrad, der durch eine maschinelle Sinneswahrnehmung und Informationsverarbeitung gekennzeichnet ist, verlagert sich die Verantwortung vom Fahrer auf das automatisierte Fahrzeug.

Bei der Klassifikation von automatisierten Fahrfunktionen werden gegenwärtig die Standards J 3016 der SAE International und der Bundesanstalt für Straßenwesen (BASt) zugrunde gelegt. Für die in dieser Arbeit durchgeführten Untersuchungen wurde die Klassifikation der BASt vorausgesetzt [4]. In dieser werden fünf unterschiedliche Automatisierungsgrade klassifiziert, aufsteigend von *Driver only* bis zu *Vollautomatisiert*. Differenziert wird bei der Einteil-

lung zwischen der Übernahme der Längs- bzw. Querverführung des Fahrzeuges, der Überwachung des Systems, der Übernahmezeit des Fahrers, sowie der Fähigkeit des Systems sich autonom in den risikominimalen Zustand zu überführen (Tabelle 1).

Tabelle 1: Grade der Automatisierung und ihre Definition [4]

Nomenklatur	Fahraufgaben des Fahrers nach Automatisierungsgrad
Vollautomatisiert	<p>Das System übernimmt Quer- und Längsführung vollständig in einem definierten Anwendungsfall.</p> <ul style="list-style-type: none"> • Der Fahrer muss das System dabei nicht überwachen. • Vor dem Verlassen des Anwendungsfalles fordert das System den Fahrer mit ausreichender Zeitreserve zur Übernahme der Fahraufgabe auf. • Erfolgt dies nicht, wird in den risikominimalen Systemzustand zurückgeführt. • Systemgrenzen werden alle vom System erkannt, das System ist in allen Situationen in der Lage, in den risikominimalen Systemzustand zurückzuführen.
Hochautomatisiert	<p>Das System übernimmt Quer- und Längsführung für einen gewissen Zeitraum in spezifischen Situationen.</p> <ul style="list-style-type: none"> • Der Fahrer muss das System dabei nicht überwachen. • Bei Bedarf wird der Fahrer zur Übernahme der Fahraufgabe mit ausreichender Zeitreserve aufgefordert. • Systemgrenzen werden alle vom System erkannt. Das System ist nicht in der Lage, aus jeder Ausgangssituation den risikominimalen Zustand herbeizuführen.
Teilautomatisiert	<p>Das System übernimmt Quer- und Längsführung (für einen gewissen Zeitraum oder/und in spezifischen Situationen).</p> <ul style="list-style-type: none"> • Der Fahrer muss das System dauerhaft überwachen. • Der Fahrer muss jederzeit zur vollständigen Übernahme der Fahrzeugführung bereit sein.
Assistiert	<p>Der Fahrer führt dauerhaft entweder die Querverführung oder die Längsführung aus. Die jeweils andere Fahraufgabe wird in gewissen Grenzen vom System ausgeführt.</p> <ul style="list-style-type: none"> • Der Fahrer muss das System dauerhaft überwachen. • Der Fahrer muss jederzeit zur vollständigen Übernahme der Fahrzeugführung bereit sein.
Driver only	<p>Der Fahrer führt dauerhaft (während der gesamten Fahrt) die Längsführung (Beschleunigung/Verzögerung) und die Querverführung (lenken) aus.</p>

Aktuell verbaute Fahrerassistenzsysteme sind maximal teilautomatisierte Systeme. Die Einführung von hochautomatisierten Systemen wird für ca. 2020 prognostiziert [5]. Vorausset-

zung dafür ist jedoch eine Gesetzestextanpassung im Straßenverkehrsgesetz, in der der Mensch und der Computer der die automatisierte Fahrzeugführung übernimmt rechtlich gleichgestellt werden. Die ersten Funktionalitäten werden nach aktuellem Stand für die Autobahn verfügbar sein, gefolgt vom Stadtverkehr und zuletzt für den ländlichen Bereich.

3. Zustandsbasierte Simulation und Transitionsregeln

Für die Transitionen in dem Modell werden auf der Basis von stochastischen Prozessen besondere Regeln eingeführt und definiert. Damit der stochastische Prozess mit komplexen Strukturen, Abhängigkeiten und beliebigen Verteilungsfunktionen modelliert werden kann, wird im Folgenden die MCS angewandt. Weitere Ausführungen hierzu siehe [6],[7] und [8]. Die zugrunde gelegten Regeln werden in den folgenden Abschnitten erläutert.

3.1.1. Mehrfache Transition

Bei herkömmlichen Zustandsdiagrammen ist üblicherweise nur eine Transition pro Richtung zwischen zwei Zuständen vorgesehen. In dem erarbeiteten Modell wird diese Konvention dadurch erweitert, dass auch mehrere duale Transitionen ($T_{I,II}(1)$ und $T_{I,II}(2)$) möglich sind (Abbildung 1). In der Praxis werden solche Transitionen üblicherweise zusammengefasst, wodurch jedoch die Hintergrundinformationen der Transition verloren gehen. Des Weiteren werden die Zielzustände oftmals geteilt, wodurch die Anzahl der Zustände exorbitant steigen.

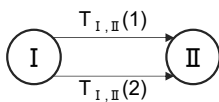


Bild 1: Abbildung einer mehrfach Transition von Zustand I nach Zustand II

3.1.2. Regel-basierte Transition

In stochastischen Prozessen werden für die Transitionen in der Regel speziell ausgewählte Übergangswahrscheinlichkeiten verwendet. Im in dieser Arbeit vorgestellten Modell können jedoch auch unterschiedliche Zufallsgrößen und deren Verteilungen verwendet werden, z.B. Wahrscheinlichkeiten basierend auf stetigen Verteilungen oder relativen Häufigkeiten. Es ist außerdem möglich, dass mehrere Entscheidungen nacheinander getroffen werden müssen, beispielsweise wenn zunächst in Abhängigkeit der Zustandsausgänge ein Ereignis berücksichtigt werden muss und danach eine vom Wechselereignis abhängige Verweildauer zu ermitteln ist.

3.1.3. Abhängige Transition

Die Eigenschaft *abhängige Transition* bedeutet, dass bestimmte Übergänge nur stattfinden können, wenn im Voraus eine definierte Transition durchgeführt wurde. In Verbindung mit der *mehrfach-Transition* können somit verschiedene Pfade über die gleichbleibende Anzahl von Zuständen ausgewählt werden (Abbildung 2).

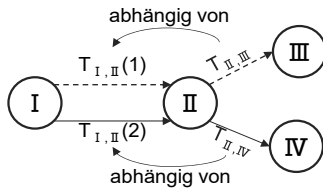


Bild 2: Beispiel einer abhängigen Transition

Die Transition von II nach III kann nur gewählt werden, wenn zuvor der Übergang von I nach II, basierend auf der gestrichelten Linie, also $T_{I,II}(1)$, durchgeführt wurde. Basierte der Übergang hingegen auf der durchgezogenen Linie, so kann vom Zustand II nur nach Zustand IV gewechselt werden. Die Entscheidungsregeln um Zustand II zu verlassen sind somit:

$II \rightarrow III \mid I \rightarrow II$ basierend auf $T_{I,II}(1)$;

$II \rightarrow IV \mid I \rightarrow II$ basierend auf $T_{I,II}(2)$.

3.1.4. Zeit-Situations-basierte System Life-Modellierung

Das Modell unterscheidet zwischen unterschiedlichen Ereignissen für die Aktivierung und die Deaktivierung des Assistenzsystems. Hierbei ist insbesondere der Unterschied zwischen einem Fehler bzw. Reparatur und der Aktivierung/Deaktivierung durch den Fahrer bzw. Verlassen des Anwendungsbereiches wichtig. Einerseits wird die relative Lebensdauer des Systems nach einer Reparatur ($t_{r,1}$, $t_{r,2}$) auf null gesetzt, das System wird hierbei komplett repariert und ist so gut wie neu, auf der anderen Seite stagniert die relative Lebensdauer nach einer Deaktivierung durch den Fahrer und wird nach der Aktivierung des Systems bei dem stagnierenden Zeitpunkt fortgesetzt (Abbildung 3). Das relative Alter des Systems steigt weiter an, während die Zeiten für *Deaktivierung durch Fahrer* oder *Verlassen Anwendungsbereich* genullt werden. Nach einer bestimmten Anzahl von Zyklen der Deaktivierung und Aktivierung tritt der Fall ein, dass die relative Lebenszeit des Systems den vorher generierten Ausfallzeitpunkt erreicht, das System in einen degradierten, aber nicht sicherheitskritischen,

Zustand wechselt und der Fahrer übernehmen muss. Es wird daraufhin eine Selbstreparatur eingeleitet, in der das System beispielsweise einen Reset durchführt. Im Anschluss wird die relative Lebensdauer auf null gesetzt.

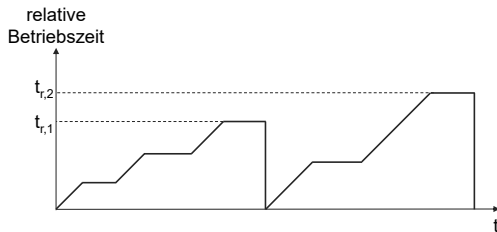


Bild 3: Zeitverlauf der Lebenszeit und der relativen Lebenszeit

4. Simulationsmodell

Das entwickelte Modell basiert auf einem Zustandsdiagramm für hochautomatisierte Fahrerassistenzsysteme. Das System agiert für definierte Anwendungsfälle komplett eigenständig ohne dass der Fahrer es überwachen muss. Der Fahrer muss jedoch jederzeit dazu in der Lage sein, die Fahrzeugführung innerhalb einer bestimmten Übernahmezeit zu übernehmen. Für den Fall, dass er dieser Aufforderung nicht nachkommt oder ein sicherheitskritischer Fehler eintritt, besteht die Möglichkeit, dass sich das System autonom in einen risikominimalen Zustand überführt. Für den Fall, dass ein *Aktionsplan „risikominimaler Zustand“* durchgeführt wird, besteht die Annahme, dass die Zustände die darauf folgen, als absorbierend anzusehen sind.

4.1. Beschreibung der Einflussgrößen

Als Einflussgröße dient das 3-Ebenen-Modell von Donges [2]. In diesem ist der Zusammenhang zwischen dem Fahrzeug, dem Fahrer und der Umwelt als Regelkreis dargestellt.

Der Fahrer kann als Einflussgröße das System direkt aktivieren und deaktivieren. Die Einflussgröße Umwelt charakterisiert sich dadurch, dass das System die Umgebung beobachtet um zu kontrollieren, ob der Anwendungsbereich der automatisierten Funktion vorhanden ist. Das Fahrzeug wird durch einen einfachen übergeordneten stochastischen Prozess dargestellt (Abbildung 4). Hierbei werden sicherheitskritische (SK) und nicht-sicherheitskritische (NSK) Komponentenausfälle mittels MCS simuliert.

Aus Zustand 1 kann das Modell einerseits durch einen nicht-sicherheitskritischen Komponentenausfall in Zustand 2 wechseln, aus dem es jedoch durch eine Reparatur wiederum in Zustand 1 übergehen kann. Andererseits wechselt das System im Falle eines sicherheitskritischen

tischen Komponentenausfalls in den absorbierenden Zustand 3. Dies kann von Zustand 1, aber auch im Anschluss an einen nicht- sicherheitskritischen Komponentenausfall, also Zustand 2, geschehen.

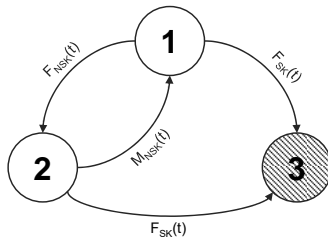


Bild 4: Das Fahrzeug illustriert als einfachen stochastischen Prozess

Mittels der MCS ist es möglich, für die Fehler- und Reparaturübergangswahrscheinlichkeiten unterschiedliche Verteilungsfunktionen anzunehmen, um Ziel und Zeit der Transition zu bestimmen, beispielsweise Weibull- oder Normalverteilung.

4.2. Zustandsbasiertes Modell

Das Zustandsdiagramm für hochautomatisierte Fahrerassistenzsysteme ist in Abbildung 5 abgebildet. Es zeigt unterschiedliche Zustandsverläufe die von einem aktiven System bis hin zu einem deaktivierten System führen. Das deaktivierte System ist hierbei entweder durch das Erreichen des risikominimalen Zustandes, durch den Fahrer oder das System, oder das nicht Erreichen des risikominimalen Zustandes charakterisiert.

4.2.2. Verlassen des Anwendungsbereichs

Für den Fall, dass der Anwendungsbereich der Funktion verlassen wird, muss dieser zuerst abgegrenzt werden. Im Modell werden die drei Bereiche Land, Stadt und Autobahn betrachtet. Diese Bereiche haben jeweils eine relative Häufigkeit, z.B. 15 % Land, 50 % Stadt und 35 % Autobahn. Nachdem der Anwendungsbereich definiert wurde, wird eine Zeit festgelegt, nach der das Fahrzeug den Anwendungsbereich verlässt. Diese Zeit basiert auf einer stetigen Verteilung, die jeweils von dem festgelegten Anwendungsbereich abhängt (Abbildung 7). Die Zeit die sich das System auf der Autobahn befindet sei z.B. länger als die im Stadtverkehr.

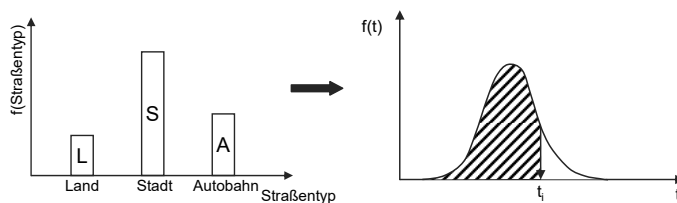


Bild 7: Relative Häufigkeit des Anwendungsbereiches und Dichtefunktion der Zeit, die das Fahrzeug im Anwendungsbereich verweilt

4.2.3. Aktivierung vice versa Deaktivierung durch den Fahrer

Die Zeit für die Aktivierung ($3 \rightarrow 1$) sowie die Deaktivierung ($1 \rightarrow 3$; $2 \rightarrow 3$) des Fahrerassistenzsystems durch den Fahrer ist, abhängig von einer kontinuierlichen Verteilung (Abbildung 8), jedoch unabhängig von weiteren Einflussgrößen.

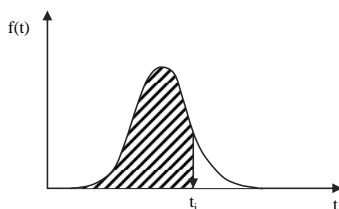


Bild 8: Dichtefunktion der Zeit nach der der Fahrer das System aktiviert/deaktiviert

4.2.4. Fahrer ignoriert

Abhängig von dem Anwendungsbereich (Abbildung 7) werden maximale Verweilzeiten festgelegt. So kann beispielsweise die Übernahmezeit auf der Autobahn größer sein als in der

Stadt. Überschreitet der Fahrer diese Zeiten so ist davon auszugehen, dass der Fahrer die Übernahmeaufforderung ignoriert (2→4).

4.2.5. Aktionsplan risikominimaler Zustand

Die Transition, ausgehend von Zustand 4, ist auf der ersten Ebene abhängig von einer diskreten Verteilung, im Kontext eines erfolgreichen Überganges in den risikominimalen Zustand (4→5) vice versa (4→6). Sollte der Übergang erfolgreich sein, so wird aus einer stetigen Verteilung eine Zeit generiert, die das System benötigt, bis es den risikominimalen Zustand erreicht (Abbildung 9).

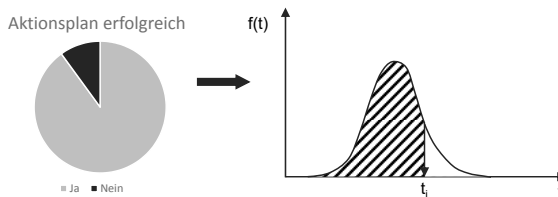


Bild 9: Relative Häufigkeit „Aktionsplan erfolgreich“ und Dichtefunktion für die Zeit nach der das Fahrzeug überführt ist

4.3. Simulationsstruktur

4.3.1. Zustand 1 – „Aktivierung System“

Zustand 1 repräsentiert ein aktiviertes Fahrerassistenzsystem, z.B. einen Autobahnpiloten. Das System funktioniert fehlerfrei und das Fahrzeug befindet sich in dem vorgegebenen Anwendungsbereich. Ausgehend von Zustand 1 ist es möglich, in Zustand 2 zu wechseln. Gründe hierfür sind entweder ein nicht-sicherheitskritischer Komponentenausfall (Transitionszeit t_{NSK}) oder das Verlassen des Anwendungsbereiches (t_{vA}). Von Zustand 1 ist es außerdem möglich, in den risikominimalen Zustand 3 zu wechseln, weil der Fahrer das aktivierte System deaktiviert (t_{deak}) und die Fahrzeugführung wieder übernimmt. Tritt im Zustand 1 ein sicherheitskritischer Komponentenausfall (t_{SK}) auf, wechselt das System in Zustand 4. Um den Zustand zu differenzieren, den das System nachfolgend einnimmt, wird die minimale Zeit der vier möglichen Transitionszeiten ausgewählt (Min[...]) (Tabelle 2).

Tabelle 2: Transitionen Zustand 1

Zustand (i)	Zustand (i+1)	Transitionszeit	Entscheidungsregel
1	2	t_{NSK}	Min [...]
	2	t_{vA}	
	3	t_{deak}	
	4	t_{SK}	

4.3.2. Zustand 2 – „Fahrerübernahmeaufforderung, Start der Verweilzeit“

In Zustand 2 findet die Fahrerübernahmeaufforderung statt und die Verweilzeit startet. Entweder deaktiviert der Fahrer das System (t_{deak}) und das System wechselt in Zustand 3 oder der Fahrer ignoriert die Übernahmeaufforderung ($t_{\text{s,max}}$ (Stadt); $t_{\text{a,max}}$ (Autobahn); $t_{\text{l,max}}$ (Land)) und das System konvertiert in Zustand 4. Außerdem ist die Transition nach Zustand 4 aufgrund eines sicherheitskritischen Komponentenausfalls (t_{SK}) möglich. Um den nächsten Zustand zu lokalisieren wird wieder die minimale Transitionszeit ausgewählt (Tabelle 3).

Tabelle 3: Transition Zustand 2

Zustand (i)	Zustand (i+1)	Transitionszeit	Entscheidungsregel
2	3	t_{deak}	Min [...]
	4	$t_{\text{s,max}}$; $t_{\text{a,max}}$; $t_{\text{l,max}}$	
	4	T_{SK}	

4.3.3. Zustand 3 – „Risikominimaler Zustand, Deaktivierung System“

Zustand 3 repräsentiert einen risikominimalen Zustand der durch die Deaktivierung des Systems durch den Fahrer erreicht wird. Der Fahrer hat in diesem Zustand wieder die Kontrolle über die Fahrzeugführung übernommen und ist für diese verantwortlich. Via Zustand 3 gelangt das System auf zwei Möglichkeiten in Zustand 1. Einerseits können nichtsicherheitskritische Komponentenausfälle durch eine Reparatur beseitigt werden (t_{rep}). Diese Transition ist nur möglich, wenn die Transition 1→2 auf einem nicht-sicherheitskritischen Komponentenausfall basiert. Andererseits kann der Fahrer das System wieder reaktivieren (t_{reak}), hierfür musste das System in Zustand 1 entweder direkt durch den Fahrer deaktiviert oder der Anwendungsbereich des Systems verlassen worden sein (Tabelle 4). Bevor das System wieder aktiviert werden kann, muss jedoch geprüft werden, ob der Anwendungsbereich des Fahrerassistenzsystems vorhanden ist.

Tabelle 4: Transition Zustand 3

Zustand (i)	Zustand (i+1)	Transitionszeit	Entscheidungsregel
3	1	t_{reak}	1→2 Deaktivierung durch Fahrer, Verlassen Anwendungsbereich
	1	t_{rep}	1→2 nichtsicherheitskritischer Komponentenausfall

4.3.4. Zustand 4 – „Aktionsplan risikominimaler Zustand“

Im Zustand 4 startet der Aktionsplan risikominimaler Zustand. Dieser Aktionsplan kann unterschiedliche Ausprägungen besitzen, z.B. das einfache Anhalten auf der aktuellen Fahrspur mit Einschalten der Warnblinkanlage oder das Wechseln auf eine andere Spur, z.B. auf der Autobahn auf den Seitenstreifen. Für die Entscheidung ob der Aktionsplan erfolgreich ist oder nicht, wird aus einer diskreten Ja/Nein-Verteilung ein Wert generiert. Sollte der Aktionsplan erfolgreich sein, wird aus einer stetigen Verteilung eine Zeit (t_{mz}) generiert, die das System für die Überführung benötigt (Tabelle 5).

Tabelle 5: Transition Zustand 4

Zustand (i)	Zustand (i+1)	Transitionszeit	Entscheidungsregel
4	5	t_{mz}	Diskrete Ja/Nein Entscheidung
	6	---	

4.3.5. Zustand 5 – „Risikominimaler Zustand, Deaktivierung System“ (absorbierend)

Zustand 5 charakterisiert einen risikominimalen Zustand, der durch die autonome Überführung des Systems eingetreten ist. Eine direkte Reaktivierung ist hier nicht möglich. Im Falle eines sicherheitskritischen Fehlers müsste das Fahrzeug z.B. erst in einer Werkstatt repariert werden, bevor es wieder betriebsbereit ist.

4.3.6. Zustand 6 – „Risikominimaler Zustand nicht erreicht“ (absorbierend)

Zustand 6 wird erreicht, wenn der Aktionsplan risikominimaler Zustand nicht erfolgreich war. Dies kann auf externen Einflüssen basieren, beispielsweise einem Unfall, z.B. verursacht durch ein anderes Fahrzeug, oder auf einem common mode, bzw. common cause Fehler der weitere Handlungen unterbindet.

5. Beispiel

Durch die MCS kann eine große Anzahl an verfügbaren Daten als Einflussgrößen generiert werden. Zunächst lässt sich der Zustandsverlauf über die Zeit abbilden (Abbildung 10 oben links). Aufbauend darauf ist es möglich, zu untersuchen, wie der Zustandswechsel der Transitionen aus bestimmten Zuständen erfolgt, dargestellt in Abbildung 10 oben rechts. In dieser sind die relativen Häufigkeiten der Transitionen vom Zustand 1 ausgehend in die Zustände 2, 3 und 4, dargestellt in einem Kreisdiagramm. Diese Auswertungsmethoden sind jedoch auch mit Markov-Ketten möglich und benötigen keine komplexen Simulationen. In den weiteren Auswerteschritten werden die Vorteile der Simulation jedoch präsenter. In Abbildung 10 unten links ist die relative Häufigkeit der Gründe dargestellt, weshalb Zustand 1 verlassen wurde. Dies wird dadurch möglich, dass während der Simulation die Transitionsgründe unterschiedlich sein können und bei der Durchführung der Simulation gespeichert werden.

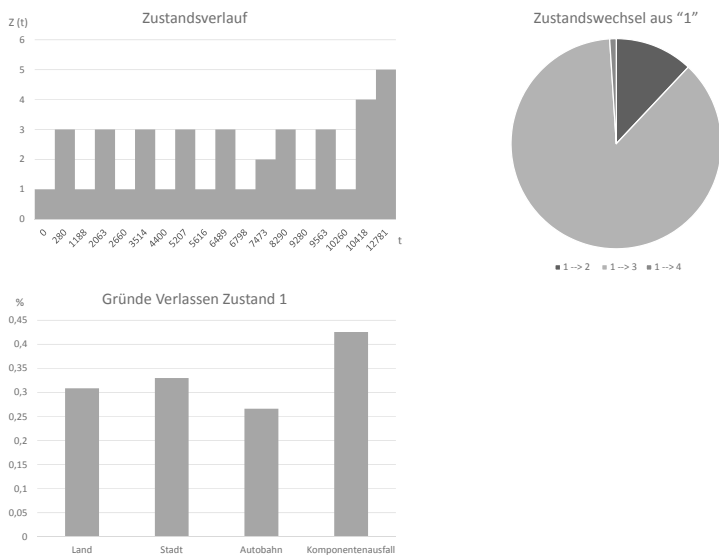


Bild 10: Zustandsverlauf über der Zeit (oben links), Zustände in die das System nach Zustand 1 wechselt (oben rechts), Gründe des Verlassens von Zustand 1 prozentual dargestellt (unten links)

Eine weitere Möglichkeit die Daten aufzubereiten ist, die Aufenthaltsdauern in den jeweiligen Zuständen zu bestimmen. In Abbildung 11 oben links wurden dabei die Zustände „Aktiviert“, „Deaktiviert“, „Reparatur“ und „Ausgefallen“ betrachtet. Unter der Annahme, dass die Zustände 5 und 6 absorbierend sind, ist es möglich, für das System empirische Ausfallwahrscheinlichkeiten zu berechnen. Im analysierten Beispiel wurden zehn Iterationen durchgeführt, wobei das System bei neun Iterationen in einen der absorbierenden Zustände gelangt ist. Somit können neun Stützstellen für eine Ausfallfunktion bestimmt werden (Abbildung 11 oben rechts). Abschließend lässt sich die Verfügbarkeit des Systems berechnen. Werden Zustand 1 und 2 betrachtet als die Zustände, in denen das System verfügbar ist, und die restlichen Zustände als nicht-verfügbar, so ist es möglich, über die Betriebszeit die Verfügbarkeit zu ermitteln. Der Verlauf der Grafik hat in der Regel die Charakteristik, dass sich die Funktion mit steigender Betriebszeit der stationären Verfügbarkeit eines Systems annähert (Abbildung 11 unten links).

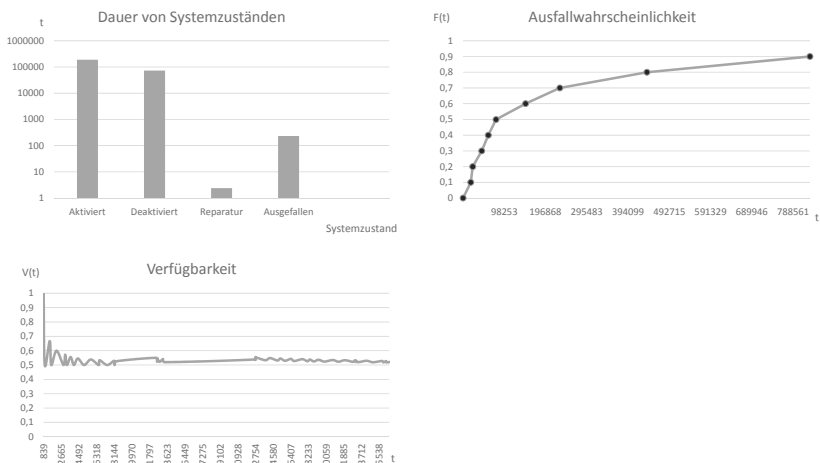


Bild 11: Aufenthaltsdauern in Systemzuständen differenziert nach „Aktiviert“, „Deaktiviert“, „Reparatur“ und „Ausgefallen“ (oben links), Ausfallwahrscheinlichkeit des Systems (oben rechts), Verfügbarkeit des Systems (unten links)

6. Fazit

Das in dieser Arbeit präsentierte Modell zeigt auf einem hohen Systemlevel einen Ansatz für eine Sicherheits- und Verfügbarkeitsanalyse, basierend auf einer zustands- und zeitbasierten MCS für hochautomatisierte Fahrerassistenzfunktionen. Die Basis für die Klassifikation bildet die Einteilung der BAST, welche in Kapitel 2 kurz vorgestellt wurde. Mittels geeigneter, definierter Regeln für die Zustandsübergänge ist es möglich, die Anzahl der Zustände möglichst gering zu halten, ohne jedoch Informationen zu verlieren. Als Einflussgrößen für das entwickelte Modell, insbesondere für die Transitionen, dient das 3-Ebenen-Modell von Donges. Durch die Betrachtung des Systems auf einem hohen Systemlevel ist es möglich, ein konzeptionelles praxisorientiertes Modell zu erstellen. Dabei konnten trotz der zahlreichen möglich Varianten quantitative Daten zur Berechnung von verwertbaren Ergebnissen im Sinne einer sicherheits- und zuverlässigkeitstechnischen Bewertung genutzt werden.

Der Ansatz, Zustandsdiagramme und die MCS zu kombinieren, scheint hierbei ein effizienter Weg ohne Restriktionen (z.B. via Markov) zu sein. Die Grundlage dieses Modells bilden die Arbeiten aus dem Forschungsprojekt für die Bundesanstalt für Straßenwesen [1]. In diesem wurden mit Unterstützung eines Expertenkreises aus der Automobil- und Zuliefererindustrie die unterschiedlichen Automatisierungsgrade hinsichtlich des risikominimalen Zustandes untersucht. Für zukünftige Untersuchungen wäre eine weitere Detaillierung des Systems auf der obersten Ebene möglich, beispielsweise sei hier auf die Einbindung von unterschiedlichen Degradationsstufen verwiesen, die wiederum Einfluss auf die Verweilzeiten haben können. Eine weitere Möglichkeit das Modell weiter zu entwickeln, ist die genaue Betrachtung der einzelnen Zustände und das Herunterbrechen bis auf die technische Ebene, um auf dieser ebenfalls mit Simulationen arbeiten zu können und auf diesem Wege u.a. verwertbare Daten wiederum für die Simulation auf der obersten Ebene zu erhalten.

Mit dem vorgestellten Modell dürfte es zukünftig möglich sein, automatisierte Fahrfunktionen schon im Entwurfsstadium abzusichern, auch ist der Vergleich unterschiedlicher Entwürfe möglich, da durch die MCS die Einflussgrößen relativ einfach verändert werden können. Ein möglicher Einsatzbereich ist auch die Funktionale Sicherheit, auf Basis des Modells lassen sich Metriken zur Validierung und Verifizierung entwickeln.

Das Modell wurde im Kontext automatisiertes Fahrerassistenzsysteme modelliert. Ein prinzipieller Transfer auf weitere Bereiche, zum Beispiel Robotik, ist nach Ansicht der Autoren durchaus möglich.

7. Literatur

- [1] Meyna, A.; Heinrich, J.: Risikobewertung unterschiedlicher Umsetzungsszenarien des Überführens eines automatisch gesteuerten Fahrzeugs in den sog. „risikominimalen Zustand“. Forschungsprojekt im Auftrag Bundesanstalt für Straßenwesen, 2016.
- [2] Donges, E.: Aspekte der aktiven Sicherheit bei der Führung von Personenkraftwagen. In Automobil Industrie Nr. 27, 1982.
- [3] Maurer M.; Gerdes, J.; Lenz, B.; Winner, H.: *Autonomes Fahren: Technische, rechtliche und gesellschaftliche Aspekte*. Berlin: Springer-Verlag, 2015.
- [4] Gasser, T.M. et al.: Rechtsfolgen zunehmender Fahrzeugautomatisierung. Berichte der Bundesanstalt für Straßenwesen, Heft F 83, Wirtschaftsverlag NW, Bremerhaven, 2012.
- [5] Winner, H.; Hakuli, S.; Lotz, F.; Singer, C.: Handbuch Fahrerassistenzsysteme: Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort. Berlin: Springer-Verlag, 2015.
- [6] Hauschild, J.: Beitrag zur Modellierung stochastischer Prozesse in der Sicherheits- und Zuverlässigkeitstechnik mittels Monte-Carlo-Simulation unter Berücksichtigung dynamischer Systemänderungen. Dissertation, Bergische Universität Wuppertal, 2007.
- [7] Meyna, A.; Pauli, B.: Taschenbuch der Zuverlässigkeitstechnik: Quantitative Bewertungsverfahren. 2., überarbeitete und erweiterte Auflage, München: Carl Hanser Verlag, 2010.
- [8] Plinke, F.: Beitrag zur Weiterentwicklung der zuverlässigkeitstechnischen Sensitivitäts- und Ausfallanalyse mittels Monte-Carlo-Simulation. Dissertation, Bergische Universität Wuppertal, 2016.

Vorhersage von Verfügbarkeitszahlen in der Entwicklungsphase neuer Produktgenerationen



M. Sc. **D. Ortmann**,
Dipl.-Ing. (FH) **C. Lisiecki**,
GROB-WERKE GmbH & Co. KG, Mindelheim

Kurzfassung

Der Beitrag stellt ein Verfahren vor, um aus der Analyse von Felddaten Verfügbarkeitszahlen neuer Produktgenerationen zu einem möglichst frühen Zeitpunkt des Produktentstehungsprozesses vorhersagen zu können. Hierzu werden Störmeldungen der aktuellen Produktgeneration im Feld erfasst und den Hauptbaugruppen der Werkzeugmaschine zugewiesen. Somit ist es möglich, den Hauptbaugruppen Verfügbarkeitswerte zu zuordnen. Unter Zuhilfenahme ausgewählter Unterbaugruppen der Stückliste und den Verfügbarkeitswerten wird durch maschinelles Lernen ein Modell trainiert, welches dann Verfügbarkeitswerte für neue Hauptbaugruppen und damit übergeordnet auch für neue Produktgenerationen vorhersagen kann.

Abstract (optional)

This paper presents a method to predict the technical availability of new product generations by analyzing field data. Therefor, alerts are recorded [in the field] and assigned to main components of the tool machine to generate the technical availability of these. The technical availabilities are combined with data from the bill of material to train a machine learning model. This can be used to predict the technical availability of new components and accordingly of new product generations.

1. Einleitung

Produzierende Unternehmen sehen sich mit der ständigen Herausforderung der steigenden Marktdynamiken konfrontiert. Individuell auf den Kunden zugeschnittene Produkte, kürzere Innovationszyklen sowie schwer vorhersehbare Nachfrageschwankungen sind nur einige Beispiele [1]. Hinzu kommt ein steigender Wettbewerbsdruck im Zuge der Globalisierung.

Letzteres führt dazu, dass sich Anbieter in den verkaufsentscheidenden Merkmalen gegenüber der Konkurrenz abgrenzen müssen.

Im Bereich der Werkzeugmaschinen sind neben dem Preis, eine qualitäts-gewährleistende Präzision sowie eine hohe Geschwindigkeit wichtige Verkaufsargumente. Insbesondere gegenüber Kunden aus dem Automobilsektor ist eine weitere wesentliche Kenngröße die technische Verfügbarkeit der Anlage. Diese Kennzahl beeinflusst nicht nur direkt die Ausbringung der Maschinen, sondern auch die Kosten im laufenden Betrieb bspw. für Instandhaltungseinsätze.

1.1 Problemstellung

In der Regel ist die technische Verfügbarkeit Teil des Kaufvertrages. Dies führt zu dem, dass bei nicht Einhalten des zugesagten Wertes, Teilzahlungen zurückgehalten oder Abnahmen nicht erteilt werden. Zum anderen muss die Kenngröße zu einem Zeitpunkt bestimmt werden, da die Werkzeugmaschine noch nicht konstruiert wurde. Hierbei sind grundsätzlich zwei Fälle zu differenzieren:

Im ersten Fall wird eine Werkzeugmaschine der Generation A verkauft. Diese Generation wurde bereits vielfach vertrieben und es existieren Felddaten, die analysiert werden. Da jedoch Werkzeugmaschinen dergleichen Produktgeneration individuell auf den Kunden zugeschnitten werden (bspw. Spannvorrichtungen, Beladekonzept, Greifer) müssen die im Feld erhobenen Verfügbarkeitszahlen auf Baugruppenebene aufgeteilt werden. Dies ermöglicht dem Vertrieb die technische Verfügbarkeit für die, individuell auf den Kunden zugeschnittene, Werkzeugmaschine zu berechnen.

Im zweiten Fall wird eine neue Generation (B) entwickelt. Ausgiebige Tests, um eine technische Verfügbarkeit zu bestimmen, würden den oben beschriebenen kürzer werdenden Innovationszyklen entgegenwirken und zu einer Schwächung der Marktposition führen. Trotzdem müssen durch den Vertrieb Verfügbarkeitszahlen präsentiert werden. Bei den GROB-WERKEN wurde dazu die im nachfolgenden vorgestellte Methode entwickelt, welche diese Kennzahl zu einem frühen Zeitpunkt des Produktentstehungsprozesses vorhersagen kann.

Da für diese Methode ebenfalls Felddaten älterer Produktgenerationen analysiert werden müssen, wird in Kapitel 3.1 zunächst die Generierung von baugruppenbezogenen Verfügbarkeitswerten beschrieben (s. erster Fall).

2. Begriffe und Grundlagen

Im Folgenden werden die begrifflichen Grundlagen für diesen Beitrag gelegt. Hierbei wird zwischen der Definition der wichtigsten Kenngrößen (Kapitel 2.1), dies beinhaltet insbeson-

dere den Begriff der technischen Verfügbarkeit, und einer kurzen Einführung in das Gebiet Big Data und Datenanalyse (Kapitel 2.2) unterschieden.

2.1 Definition der wichtigsten Kenngrößen

Wie oben beschrieben ist die Kennzahl technische Verfügbarkeit ein wesentliches Verkaufsargument. Laut [2] gibt die technische Verfügbarkeit „den prozentualen Anteil der Belegungszeit an, für den die Maschine/Anlage ohne *technischen* Mangel der Produktion zur Verfügung steht“. Die Zeiten zur Behebung der technischen Mängel werden zur technischen Ausfallzeit aggregiert: „Die technische Ausfallzeit ist die Summe aller Ausfallzeiten, die ihre Ursache in Mängeln der Konzeption oder Ausführung einer Maschine/Anlage haben. Sie gehören zur Verantwortung des Herstellers“ [2]. Die Ausfallzeiten, die vom Betreiber zu verantworten sind, werden unter dem Begriff der organisatorischen Ausfallzeit zusammengeführt [2].

Für eine Differenzierung der beiden Begriffe haben die GROB-WERKE, die folgende grafische Darstellung der Zeiten zur Bestimmung der Verfügbarkeit entworfen (s. Bild 1) und die Zeitanteile beim Auftreten einer Störung der organisatorischen und der technischen Ausfallzeit zugeordnet.

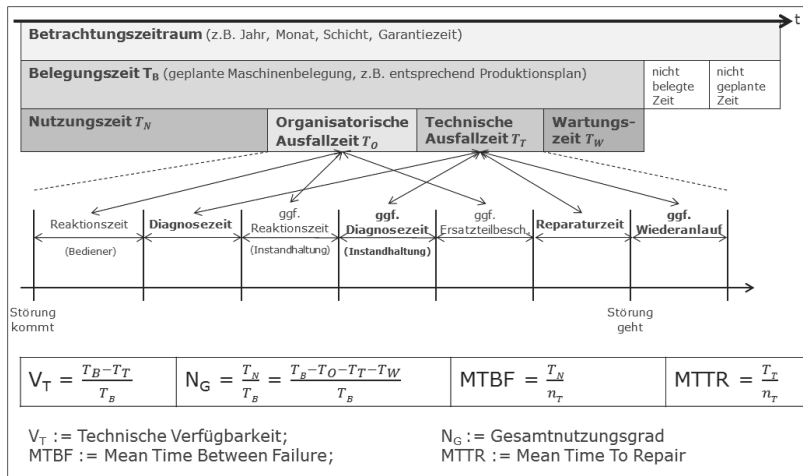


Bild 1: Übersicht der wichtigsten Kenngrößen (i.A.a. [2])

Ein ebenfalls häufig verwendeter Begriff ist die Availability (Verfügbarkeit). Diese bestimmt sich aus dem „Verhältnis der Zeit, in der das System funktionsfähig ist, zur gesamten Zeit“

[3]. Da in diesem Beitrag, jedoch die technische Verfügbarkeit nach [2] verwendet wird, sei lediglich auf die Literatur [3] verwiesen.

2.2 Big Data und Datenanalyse

Da in der Praxis der Begriff **Big Data** häufig nicht klar definiert ist, soll hier eine Eingrenzung mit Hilfe der sogenannten „**5 V's**“ (Volume, Variety, Velocity, Veracity, Value) vorgenommen werden. **Volume** bedeutet, dass die Datenmenge enorm groß ist (>1 Petabyte). **Variety** beschreibt die Datenvielfalt (bspw. Wartungsverträge, E-Mails, Videos). Der Begriff **Velocity** umfasst die Geschwindigkeit mit der Daten generiert, aber auch analysiert werden. Die Datenqualität bzw. Wahrhaftigkeit der Daten wird mit dem Begriff **Veracity** bewertet. **Value** bezeichnet den Wert, der aus den Daten gewonnen werden kann [4 und 5].

Zur Analyse werden häufig Methoden des **Maschinellen Lernens** genutzt. Dies bedeutet „Computer so zu programmieren, dass ein bestimmtes Leistungskriterium anhand von Beispieldaten oder Erfahrungswerten aus der Vergangenheit optimiert wird“ [6]. Mit Hilfe des Modells können sowohl Rückschlüsse auf die Erfahrungswerte als auch Vorhersagen für die Zukunft getroffen werden [6].

Für das Trainieren eines Modells können (je nach Anwendungsfall) verschiedenste Algorithmen genutzt werden. Ein Beispiel ist die Klasse der Entscheidungsbäume. Hierbei handelt es sich um ein hierarchisches Modell, welches aus Entscheidungsknoten und terminalen Blättern besteht [6 und 7].

3. Verfügbarkeitszahlen und Big Data im Werkzeugmaschinenbau

In diesem Kapitel wird aufgezeigt, wie Verfügbarkeitszahlen auf Hauptbaugruppenebene generiert werden können, welche Probleme dabei zu beachten sind und wie diese Daten in Bezug zu Big Data (5V's) eingeordnet werden können.

Vorab sei hier beispielhaft eine vereinfachte Stücklistenstruktur einer Werkzeugmaschine skizziert:

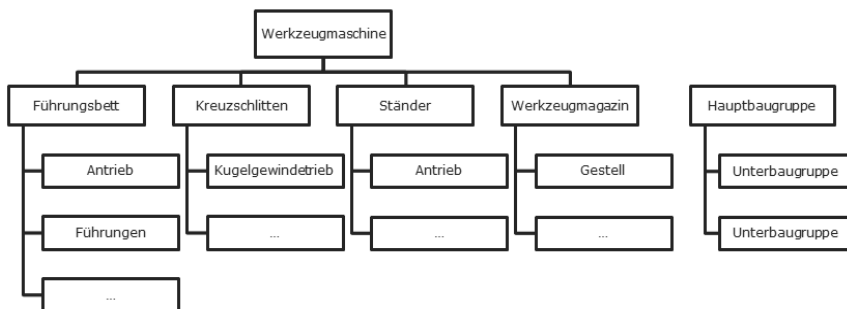


Bild 2: vereinfachte Darstellung der Stücklistenstruktur einer Werkzeugmaschine

3.1 Generierung von Verfügbarkeitszahlen auf Hauptbaugruppenebene

Generell erzeugt jede Werkzeugmaschine im Verlaufe ihres Lebens eine Reihe von Fehlermeldungen. Diese werden über einen gewissen Zeitraum im Alarm-Log gespeichert und sind auf der Maschine verfügbar. Da jedoch nicht jede Störmeldung zu einem technischen Ausfall führt, müssen diese Daten mit weiteren Maschinendaten angereichert werden. In der Regel gibt es hierzu ein Manufacturing Execution System (MES), welches bspw. zusätzlich Achsbewegungen kontrolliert und entsprechend der Bewegungen verschiedene Status ausgibt. So stellt ein MES aggregierte Daten zur Verfügung, in dem die verschiedenen Zustände mit Zeitstempel und eventuell weiteren Informationen vorhanden sind.

Diese Daten müssen nun analysiert werden. Hierzu werden zunächst die einzelnen Status, die das MES ausgibt, analysiert und bewertet, welcher Status der technischen Ausfallzeit zugeordnet werden kann (bspw. Störung und Reparatur). Dabei gilt es zu beachten, dass die Betreiber der Werkzeugmaschinen unterschiedliche MES verwenden, was dazu führt, dass sich die Zustände und deren Berechnung unterscheiden. Hier muss vom Werkzeugmaschinen-Hersteller eine Analyse durchgeführt werden, sodass die verschiedenen Rohdaten der Kunden-MES zunächst in ein einheitliches Format gebracht werden.

Weiterhin können mehrere Fehlermeldungen zum gleichen Zeitpunkt auftreten, was, je nach MES, dazu führt, dass sich Zustände überschneiden oder gar doppelt vorkommen, wodurch bspw. real 10 Sekunden vergehen, aber die Maschine (lt. MES-Daten) 20 Sekunden ausfällt, da der Zustand doppelt vorkommt. Dies würde zu einer Verschlechterung der technischen Verfügbarkeit führen und muss bereinigt werden.

Um nun eine Verfügbarkeit für Hauptbaugruppen zu berechnen, müssen die einzelnen Status, die die technische Ausfallzeit beeinflussen, analysiert werden. Für einen Status, der eine Störmeldung der Maschine enthält, kann die Störmeldung analysiert und einer Hauptbau-

gruppe zugeordnet werden. Für einen Status ohne Störmeldung (bspw. Reparatur) muss ein Zusammenhang zu einer Störmeldung gezogen werden. Eine einfache Möglichkeit ist, den Status ohne Störmeldung dem letzten Status mit Störmeldung zuzuordnen. Somit könnten die Zeiten addiert und entsprechend der technischen Verfügbarkeit der Baugruppe angerechnet werden. Dies erweist sich jedoch nur als sinnvoll, sofern die beiden Status zeitlich betrachtet nicht zu weit auseinander liegen. Andernfalls müssen hier weitere Daten zur Hilfe genommen werden. Die Vorgehensweise zur Generierung von baugruppen-bezogenen Verfügbarkeitszahlen fasst Bild 3 zusammen:

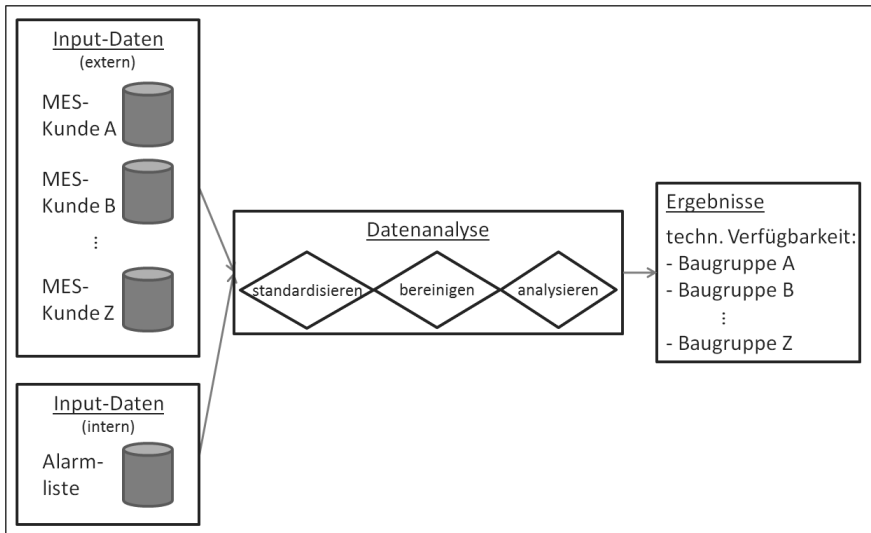


Bild 3: Generierung von baugruppenbezogenen Verfügbarkeitszahlen

3.2 Eingruppierung in den Bereich Big Data

Um die gerade beschriebenen Daten zu klassifizieren, werden diese im Folgenden gemäß der in Kapitel 2.2 vorgestellten Big Data Definition eingeordnet.

Der Bereich **Volume** (Datenmenge) hängt von verschiedenen Einflussfaktoren ab, bspw. wie viele Daten vom MES abgefragt werden und wie diese aggregiert werden. Da jedoch nicht nur eine Maschine im Feld ist, sondern eine Vielzahl von Maschinen bei vielen Kunden, entsteht alleine dadurch eine große Datenmenge. Weiterhin können Daten deutlich schneller als im Sekundentakt generiert werden, womit auch eine entsprechend hohe Geschwindigkeit (**Velocity**) bei der Erzeugung von Daten gewährleistet ist.

Da es sich bei MES-Daten immer um strukturierte Daten handelt, ist der Bereich der **Variety** sicherlich weniger stark ausgeprägt. Die gleiche Bewertung erfolgt für die **Veracity** der Daten, da es sich um bereits aggregierte Daten handelt und (aus Sicht des Werkzeugmaschinenherstellers) nicht immer geklärt werden kann, wie welcher Status innerhalb des MES entsteht. Dennoch haben die Daten einen hohen **Value** (Wert) für den Werkzeugmaschinenhersteller. Die baugruppenbezogenen Verfügbarkeitswerte können im Vertrieb genutzt werden, um den Kunden realistische Kennzahlen für seine Werkzeugmaschine zu präsentieren oder in der Konstruktion/Entwicklung um Schwachstellen der Baugruppen zu verbessern. So kann ein kontinuierlicher Verbesserungsprozess datengetrieben erfolgen und mit Hilfe von Verfügbarkeitszahlen validiert werden.

Tabelle 1 fasst die beschriebene Einordnung noch einmal zusammen, wobei „+“ eine geringe Erfüllung, „++“ eine mittlere Erfüllung und „+++“ eine volle Erfüllung des Teilaspekts von Big Data (aus Sicht eines Werkzeugmaschinenherstellers) bedeuten:

Tabelle 1: Big Data im Werkzeugmaschinenbau

Volume	Variety	Velocity	Value	Veracity
+++	+	+++	++	+

4. Vorhersage von Verfügbarkeitszahlen

In Kapitel 3 wurde gezeigt, wie aus Felddaten Verfügbarkeitswerte für Hauptbaugruppen bestimmt werden können (Kapitel 1.1 erster Fall). In diesem Kapitel wird nun eine Methode beschrieben, mit der Verfügbarkeitskennzahlen für eine neue Produktgeneration bereits in der Entwicklungsphase bestimmt werden können (Kapitel 1.1 zweiter Fall).

4.1 Dateninput für das Modell

Als Grundlage des folgenden Verfahrens dienen die Ergebnis-Daten aus Kapitel 3 (s. Bild 2). Es werden die Verfügbarkeitszahlen für die erfassten Hauptbaugruppen betrachtet. Diese werden nun mit den Komponenten aus den darunterliegenden Stücklistenebenen der Baugruppe ergänzt. Je nach Komplexität ergeben sich dadurch mehrere hundert Teile je Hauptbaugruppe. Da nicht jedes Bauteil einen Einfluss auf die technische Verfügbarkeit hat, wird die Stückliste aggregiert und nur bestimmte Ausschnitte aus der Stückliste werden verwendet. Beispielsweise hat ein Antrieb sicherlich größeren Einfluss auf die technische Verfügbarkeit als ein Blech oder eine Schraube. Zur Einteilung, welche Auszüge aus der Stückliste mehr oder weniger relevant sind, können verschiedene Herangehensweisen gewählt werden. Ein Ansatz kann sein auf das firmeneigene Knowhow zurückzugreifen. Beispielsweise

könnten Service-Mitarbeiter befragt werden, welche Bauteile innerhalb der Hauptbaugruppe häufig ausfallen oder ausgetauscht werden. Dies kann jedoch zu sehr subjektiven Beurteilungen führen. Ein zweiter Ansatz ist deshalb Service-Meldungen zu untersuchen. Diese werden angelegt, wenn Kunden technische Probleme mit der Werkzeugmaschine haben. Bedingung hierfür ist eine entsprechende Pflege der Service-Meldung bzw. eine Auswertemöglichkeit von semi-strukturierten Daten. Ein Vorteil hierbei ist, dass Daten von einer Vielzahl von Maschinen vorliegen und diese objektiv ausgewertet werden können.

Neben der Anzahl der ausgewählten verfügbarkeitsrelevanten Baugruppen können weitere Informationen hinzugefügt werden. Die Summe der verbauten Teile innerhalb der Hauptbaugruppe kann bspw. ein Indiz für die Komplexität der Baugruppe sein. Weiterhin kann eine Klassifizierung der Bauteile in Ersatz- und/oder Verschleißteile (E/V) verwendet werden. Sofern die Service-Meldungen einer Baugruppe zugeordnet werden können, kann auch die Anzahl dieser Meldungen eine Inputinformation für ein mathematisches Modell sein.

Generell kann als Input zunächst jede Information verwendet werden, für die ein Zusammenhang mit der technischen Verfügbarkeit gezogen werden kann. Ein vereinfachtes Beispiel für solche Daten zeigt Tabelle 2. Die Spalten „1“, „2“ und „3“ beschreiben Eingruppierungen in verschiedene Arten von Ersatz- und Verschleißteilen:

Tabelle 2: (Beispiel-)Input Daten

Antriebe	Führungs- leisten	Rotations- Achsen	verbaute Teile	Summe E/V	1	2	3	techn. Ver- fügbarkeit
8	16	6	6452	243	3	124	89	0,990033
6	14	4	4412	380	8	105	23 3	0,995004
4	9	3	7699	838	3	245	57 6	0,996503
1	4	1	811	197	1	118	75	0,996500
6	8	4	4389	276	16	147	69	0,992424
...								

4.2 Training und Anwendung des Modells

Um Verfügbarkeitszahlen von neuen Baugruppen bzw. neuen Produktgenerationen vorherzusagen, werden die in 4.1 beschriebenen Daten verwendet, um einen Algorithmus zu trainieren. Dabei werden die Daten in vier Gruppen unterteilt. Zunächst sollten die Daten in Trainings- und Testdaten differenziert werden. Der Großteil der Daten sollte dabei zum Trai-

nieren des Modells genutzt werden. Beide Gruppen werden erneut in Ein- und Ausgabewerte gegliedert. In diesem Fall sind die Verfügbarkeitszahlen die Ausgabewerte (s. Tabelle 2: Spalte „techn. Verfügbarkeit“). Die Eingabewerte sind bspw. die aggregierte Stückliste sowie die Ersatz- und Verschleißteilkennzeichnung (Tabelle 2: Spalten „Antriebe“ bis „3“). Der Algorithmus verwendet nun die Eingabewerte, um die Parameter des Modells zu optimieren und als Ergebnis die Ausgabewerte zu liefern.

Die Testdaten werden zur Validierung genutzt. Hierzu werden dem Algorithmus die Eingabewerte zur Verfügung gestellt und der Ausgabewert wird durch den Algorithmus berechnet. Dieser kann nun mit dem realen Ausgabewert der Testdaten verglichen werden. Im Beispiel aus Tabelle 2 könnte dies wie folgt aussehen. Die ersten 100 Zeilen werden zum Trainieren des Modells, die letzten 20 Zeilen zum Validieren genutzt. Bei der Validierung wird die Spalte „techn. Verfügbarkeit“ nicht an das trainierte Modell übergeben, sondern durch das Modell berechnet. Die berechnete technische Verfügbarkeit, kann dann mit der realen technischen Verfügbarkeit verglichen werden. Weichen die Werte stark voneinander ab, kann dies verschiedene Ursachen (falscher Algorithmus, overfitting, etc.) haben und muss im Einzelfall erörtert werden.

Ist der Unterschied des vorhergesagten und des realen Ausgabewerts tolerierbar, kann das Modell zur Vorhersage von Verfügbarkeitszahlen neuer Baugruppen genutzt werden. Verfügbarkeitszahlen für nicht veränderte Baugruppen können von der alten Produktgeneration übernommen werden. Für Baugruppen, die entscheidend verändert wurden, werden dem Modell die Eingabewerte (analog der Testdaten) übergeben. Da in diesem Fall nicht zwangsläufig alle Inputdaten vorhanden sind, da beispielsweise noch keine Service-Meldungen vorliegen, sollte ein möglichst robuster Algorithmus verwendet werden, es bietet sich in diesem Fall ein Algorithmus der Klasse Entscheidungsbäume (s. Kapitel 2.2) an.

Abbildung 4 fasst die aufgezeigten Schritte nochmals zusammen:

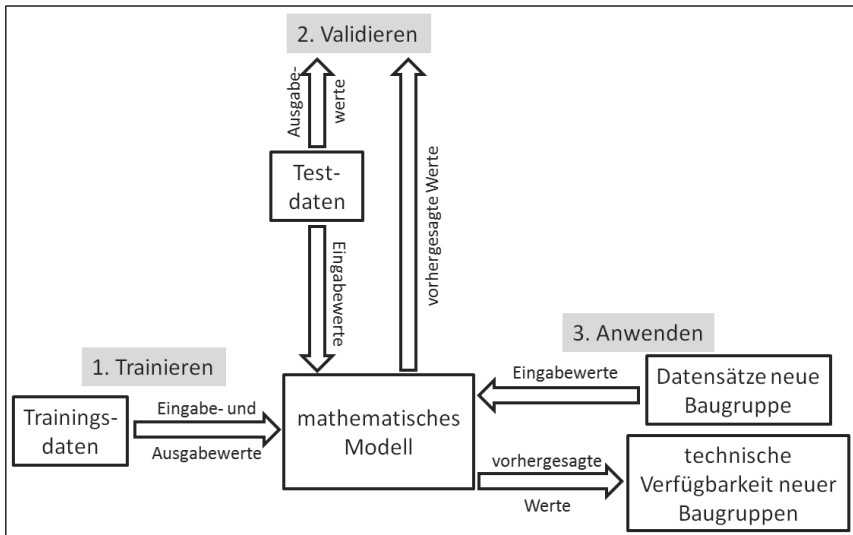


Bild 4: Vorhersage von Verfügbarkeitszahlen neuer Baugruppen

4.3 Prototypische Umsetzung bei den GROB-WERKEN

Ein Prototyp der beschriebenen Vorgehensweise wurde bei den GROB-WERKEN in Python implementiert und soll hier diskutiert werden.

Einen Auszug der Input-Daten zeigt Tabelle 2. Das Verhältnis der Trainingsdaten zu Testdaten betrug ca. 75% zu 25%. Als Algorithmus wurde der sogenannte RandomForest, aus der Klasse der Entscheidungsbäume genutzt. Wie der Name vermuten lässt, handelt es sich nicht um einen einzelnen Entscheidungsbaum, sondern um eine Vielzahl von Bäumen (Wald). Hierbei werden mehrere Modelle gleichzeitig trainiert und zu einem „Forest“ aggregiert. Zur Validierung wurde der Mittelwert der absoluten Abweichung zwischen der vorhergesagten und realen technischen Verfügbarkeit berechnet. Dieser lag mit 0,000934 in einem grenzwertig hohen Bereich, da dies bedeutet, dass der vorhergesagte Wert im Durchschnitt 0,000934 über bzw. unter dem realen Wert liegt. Die vorhergesagte technische Verfügbarkeit also beispielsweise 99,51%, die reale technische Verfügbarkeit jedoch 99,6% (oder 99,42%) betrug. Tritt diese Abweichung je veränderter Baugruppe auf kann die vorhergesagte Verfügbarkeit leicht von der realen technischen Verfügbarkeit der gesamten Werkzeugmaschine abweichen. Im Zuge steigender Datenmengen und unter zur Hilfenahme von weiteren Verfahren, wie bspw. der Kreuzvalidierung wird jedoch angenommen, dass die Qualität des Mo-

dells noch erheblich gesteigert werden kann. Diese Annahme gilt es jedoch in Zukunft noch zu untersuchen.

Ein weiterer Aspekt, der kurz erwähnt werden soll, ist die Analyse der Gewichtung der einzelnen Eingabewerte. Hierbei zeigt sich, dass die Angaben zu Ersatz- und Verschleißteilen deutlich höher gewichtet werden, als die Angaben zur Anzahl konkreter Baugruppen. So liegt die Gewichtung der Anzahl der Rotationsachsen bei 1,3%, wohingegen die Gewichtung der Summe der Ersatz- und Verschleißteile bei 12,7% liegt. Diese Diskrepanz zeigt, dass für zukünftige Modelle der Schwerpunkt der Eingabewerte vermehrt auf Informationen zu Ersatz- und Verschleißteilen gelegt werden sollte bzw. generell weitere Input-Informationen getestet werden können.

5. Zusammenfassung und Ausblick

Aufgrund des steigenden Wettbewerbsdruck sehen sich Unternehmen gezwungen, sich in den verkaufsentscheidenden Kenngrößen gegenüber der Konkurrenz abzugrenzen. Im Bereich der Werkzeugmaschinen betrifft dies unter anderem die Kennzahl der technischen Verfügbarkeit. Um diese bereits im Anschluss an die Projektierungsphase angeben zu können, wurden zwei aufeinander aufbauende Methoden erläutert. Zum einen wurde gezeigt, wie baugruppenbezogene Verfügbarkeitszahlen aus Felddaten generiert werden können. Damit ist es möglich, für eine bestehende Produktgeneration kundenindividuell Verfügbarkeitszahlen zu generieren. Zum anderen wurde eine Methodik vorgestellt, um bereits während des Produktentstehungsprozesses von neuen Produktgenerationen Verfügbarkeitszahlen zu bestimmen. Weiterhin konnte gezeigt werden, welche Kriterien von Big Data im Bereich der Werkzeugmaschine erfüllt werden.

Zur Weiterentwicklung und Validierung der vorgestellten Methode können folgende Punkte erarbeitet werden:

- Wie weit ist ein solches Modell generalisierbar und auf andere Maschinen übertragbar bzw. wo sind Grenzen, die einen Vergleich zweier Produktgenerationen unmöglich machen?
- Welche weiteren Datenquellen könnten als Eingabewerte genutzt werden, um das Modell zu optimieren?
- Macht eine Aggregation der Stückliste Sinn, oder sollte die gesamte Stückliste als Input dienen?

- Für die Validierung des Modells könnten für eine neue Produktgeneration Verfügbarkeitszahlen vorhergesagt werden und diese zu einem späteren Zeitpunkt mit Hilfe von Felddaten der neuen Produktgeneration verglichen werden.
- Gibt es weitere Anwendungsfälle für diese Vorgehensweise? Könnte beispielsweise die Anzahl der Service-Anrufe für neue Baugruppen vorhergesagt werden?

6. Literaturangaben

- [1] Lass, S; Theuer, H; Hennig, G; Schumacher, J: Modellierung intelligenter Produktionssysteme. Productivity Management 18(1), 2013, S.47-50
- [2] VDI-Richtlinie 3423: Verfügbarkeit von Maschinen und Anlagen – Begriffe, Zeiterfassung und Berechnung, 2011
- [3] Eberlin, S; Hock, B: Zuverlässigkeit und Verfügbarkeit technischer Systeme. Wiesbaden: Springer Fachmedien 2014
- [4] Fasel, D; Meier, A: Big Data Grundlagen, Systeme und Nutzungspotenziale. Wiesbaden: Springer Fachmedien 2016
- [5] Gandomi, A; Haider, M: Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management, 35, 2015, S.137-144
- [6] Prof. Dr. Alpaydin, E: Maschinelles Lernen. München: Oldenbourg Wissenschaftsverlag GmbH 2008
- [7] Mitchell, T: Machine Learning. McGraw-Hill Science/Engineering/Math 1997

Anwendung von ausgewählten multivariaten Analysemethoden des Data Mining zur Identifizierung von ausfallverbundenen Fahrverhaltensmustern auf Basis von Kundenbelastungskollektiven



Application of multivariate data mining methods on load collectives to identify failure related driving patterns

Dipl.-Ing. **Patrick Dobry**, Dr.-Ing. **Andreas Jacobi**,
Daimler

Univ. Prof. Dr.-Ing. **Bernd Bertsche**,
Institut für Maschinenelemente, Universität Stuttgart

Kurzfassung

Ausfälle spezifischer Fahrzeugkomponenten finden ihre Ursache oftmals in einem bestimmten Nutzungsverhalten des Endkunden wieder. Um dieses Nutzungsverhalten verstehen und interpretieren zu können, bedienen sich Fahrzeughersteller und Zulieferer an zeit- und ortsunabhängigen, in elektronischen Steuergeräten abgespeicherten Kundenbelastungskollektiven, welche während eines Werkstattbesuches via Diagnose ausgelesen werden können. Auf Basis derartiger Datensätze von Belastungskollektiven ausgefallener Fahrzeuge können durch geeignete Anwendung multivariater Analysemethoden des Data Mining Nutzungsmuster identifiziert und somit im Feld gefährdete Fahrzeuge frühzeitig erkannt werden. In diesem Artikel werden, in Anlehnung an den CRISP-Prozess, anhand einer ausgefallenen Fahrzeugkomponente der Einsatz zweier ausgewählter Analysemethoden vorgestellt, sowie deren anschließenden Anwendungsmöglichkeiten bezüglich präventiver Kundendienstmaßnahmen- und einer an den Kunden angepasste Erprobungsplanung/-optimierung aufgezeigt.

Abstract

The failure of certain vehicle components are often caused by the customer's specific driving behavior. To understand and interpret this behavior, OEMs and suppliers use time- and location independent load collectives which are stored in the control units and can be transferred by diagnosis during a workshop visit. Based on failed vehicles' datasets, specific usage patterns can be identified by multivariate analytical methods in order to track vehicles at risk in the field. This article considers the implementation of two analytical methods in case of a specific failed vehicle component by following the CRISP-model. The resulting use cases for preventive customer services and customer's based test planning/ optimization are shown subsequently.

1 Einleitung

Im Automobilumfeld besitzt das Streben nach einer höheren Zuverlässigkeit zur Sicherung der marktführenden Position im Premiumsegment eine äußerst große Bedeutung. Um den stetig wachsenden Ansprüchen gerecht zu werden, benötigen Hersteller ein Konzept, die große Varianz in der Kundennutzung für die Auslegung neuer Komponenten zu berücksichtigen, um die Risiken späterer Probleme im Feld minimieren zu können. Hierzu ist eine genaue Kenntnis der durch Kunden induzierten Belastung der einzelnen Fahrzeugkomponenten von Nöten. Sie wird mithilfe von im Fahrzeug gespeicherten zeit- und ortsunabhängigen Belastungskollektiven überwacht und per Diagnose dem Hersteller zur Verfügung gestellt [1]. Somit lassen sich sowohl Varianz und Streuung der gesamten Fahrzeugflotte, als auch Belastungen kleinerer Fahrzeuggruppierung, abbilden. Das hohe Nutzungspotential dieser Daten in Verbindung mit hochmodernen Steuergeräten und Datenloggern führen zu immer größeren Datenmengen, die es gilt, mit geeigneten Verfahren des Data Minings zu einem Mehrwert sowohl für den Hersteller, als auch für den Kunden zu generieren. Für diesen Zweck wurde durch ein Konsortium unter Beteiligung der Daimler AG das CRISP-Modell entwickelt, welches den prinzipiellen Prozessablauf eines Data Mining Einsatzes im Unternehmen beschreibt und für Cross Industrie Standard Process for Data Mining steht [2].

In dieser Arbeit wird am Beispiel eines ausgewählten Bauteiles, in Anlehnung an CRISP, der Einsatz zweier multivariater Analysemethoden vorgestellt, welche die Möglichkeit bieten, anhand eines Modells auf Basis von Kundenbelastungskollektiven das Ausfallverhalten der Komponenten abzubilden, um Fahrzeuge hinsichtlich ihres Ausfallrisikos klassifizieren zu können. Es erfolgt zudem eine kurze Darlegung verschiedener Einsatzmöglichkeiten der entwickelten Analysemodelle im Unternehmen.

2 Belastungskollektive

Belastungskollektive (BLK) sind Daten, die mittels Zähl- und Klassierverfahren, aus Signalmessungen verschiedener Sensoren im Fahrzeug generiert und in den Steuergeräten gespeichert werden. Zu diesen Verfahren gehören die Momentanwert-, Ereignis-, Integral-, sowie die Rainflowzählung [3]. Die Transformation der Signaldaten sorgt neben der signifikanten Reduzierung des benötigten Speicherplatzes im jeweiligen Steuergerät für eine Eliminierung des Zeit-, Ort- und Personenbezuges und ist aufgrund existierender Datenschutzgesetze verpflichtende Basis für eine Datenerhebung aus Sicht des Herstellers [1]. Gängige Belastungskollektive sind beispielsweise „Anzahl Start-Stopp Vorgänge“ oder „Motor-Verweildauer im Drehmomentintervall x“. Auch mehrdimensionale Belastungskollektive sind unter Einhaltung der nicht-Reproduzierbarkeit des zeitlichen Bezuges eine geeignete Möglichkeit, die Belastung bestimmter Komponenten mit einem geringeren Informationsverlust abzubilden.

Die Datenübertragung aus den Steuergeräten der Fahrzeuge an den Hersteller erfolgt entweder in Form von Diagnosesitzungen, die während eines Werkstattaufenthaltes durchgeführt werden, oder mittels einer kabellosen Datenübertragung via Telediagnose in festgelegten Zeitintervallen (Abbildung 1).

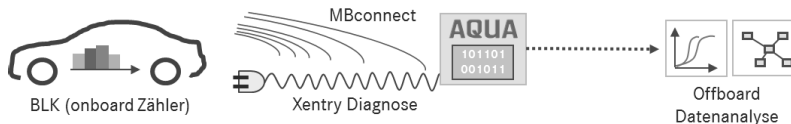


Abbildung 1: Belastungskollektive und deren Übertragung aus dem Kundenfahrzeug [4]

Die kabellose Form der Datenerhebung über Telediagnose bietet dem OEM die Möglichkeit, neben der Kenntnis auftretender Belastungen, dem Kunden neue Dienstleistungen in Form von beispielsweise präventiven Reparaturen und belastungs-abhängigen Service-Intervallen anzubieten. Somit kann die Anzahl von Fahrzeugausfällen im Feld und die damit verbundenen Kosten gesenkt und gleichzeitig die Kundenzufriedenheit erhöht werden.

3 Angewandte Multivariate Analyseverfahren des Data Mining

Da der Ausfall bestimmter Fahrzeugkomponenten im Feld meist von mehr als nur einer Einflussgröße (Eingangsvariable) bestimmt wird, kommen zur Klassifizierung einzelner Fahrzeuge multivariate Analysemethoden zum Einsatz. Es wird dabei zwischen linearen und nicht-linearen Methoden unterschieden. Die in diesem Beitrag angewandten Methoden der linearen Diskriminanzanalyse und der logistischen Regression zählen zu den linearen Verfahren [5], deren Ziel es ist, mithilfe eines linearen Modells auf Basis der verfügbaren Eingangsvariablen X_i eine Aussage über die Ausfallwahrscheinlichkeit einer bestimmten Komponente im Feld treffen zu können. Folglich werden Fahrzeuge ausschließlich einer der beiden Klassen „Ausfall“ oder „nicht Ausfall“ (Survivor) zugeordnet. Man spricht hier von einem Zwei-Klassen-Fall.

Lineare Diskriminanzanalyse LDA

Die Diskriminanzanalyse zählt zu den strukturprüfenden Verfahren, deren Methodik darauf beruht, eine Linearkombination der i Eingangsvariablen X_i zu erstellen, welche die beiden Klassen optimal voneinander separiert indem deren Abstand zueinander maximiert, sowie die Streuung innerhalb der jeweiligen Klassen minimiert wird. Jeder Eingangsvariablen X_i wird ein Diskriminanzkoeffizient b_i zugeordnet, b_0 bildet die Konstante der Linearkombination [5].

$$Y_j = b_0 + b_1X_{1j} + b_2X_{2j} + \dots + b_iX_{ij} \quad (1)$$

Diese lineare Funktion erwirkt eine bestmögliche Trennung der Verteilungen beider Gruppen.

In welchem Maß sich die beiden Gruppen tatsächlich voneinander unterscheiden, lässt sich durch das Diskriminanzkriterium Γ beschreiben.

$$\Gamma = \frac{\text{Streuung zwischen den Gruppen}}{\text{Streuung in den Gruppen}} = \frac{\sum_{g=1}^2 I_g \cdot (\bar{Y}_g - \bar{Y})^2}{\sum_{g=1}^2 \sum_{i=1}^{I_g} (Y_{gi} - \bar{Y}_g)^2} \quad (2)$$

Die Streuung zwischen den Gruppen wird anhand des quadrierten Abstands der Gruppen-Centroide \bar{Y}_g zum Gesamtmittel \bar{Y} ermittelt, während sich die Streuung innerhalb der Gruppen durch die quadrierten Abstände der Gruppenelemente zu den jeweiligen Gruppen-Centroiden ergibt. Die Modelloptimierung besteht demnach in der Findung des maximalen Diskriminanzwertes, welcher auch als Eigenwert γ des Modells bezeichnet wird. Dies geschieht mithilfe der Methode der kleinsten Quadrate. Der kritische Diskriminanzwert liegt exakt zwischen den beiden Centroiden der Gruppen und bildet dementsprechend den Grenzwert [5]. Liegt der Diskriminanzwert eines Objekts in Bezug auf den kritischen Diskriminanzwert auf der Seite des Gruppencentroiden der Gruppe A, erfolgt eine Zuordnung zur Gruppe A, andernfalls wird das jeweilige Objekt der Gruppe B zugeordnet.

Logistische Regressionsanalyse

Die logistische Regressionsanalyse basiert analog zur linearen Diskriminanzanalyse auf einem linearen Modell. Im Unterschied zur LDA liefert jedoch die logistische Regressionsanalyse direkte Wahrscheinlichkeiten p über die Gruppenzugehörigkeit der zu klassifizierenden Objekte [6]. Hierzu wird die Linearkombination der Eingangsgrößen in die logistische Verteilungsfunktion

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

eingesetzt, woraus folgende logistische Regressionsfunktion resultiert.

$$p(x_i) = \frac{1}{1 + e^{-(b_0 + b_1 X_{1j} + b_2 X_{2j} + \dots + b_L X_{Lj})}} \quad (4)$$

Da es sich in diesem Fall um eine nicht-lineare Funktion handelt, können die Parameter nicht mittels der Methode der kleinsten Quadrate geschätzt, sondern müssen mithilfe des Maximum Likelihood Verfahrens iterativ ermittelt werden. Als Gütemaß des Modells dient der negierte doppelte Log-Likelihood $-2LL$, sowie der Likelihood-Ratio-Test. Für eine Klassifizierung der Objekte wird ein Trennwert p^* festgelegt [5]. Im vorliegenden zwei-Gruppen-Fall beträgt dieser 0,5. Die Anpassung dieses Trennwertes ermöglicht, das Risiko einer Fehlzugehörigkeit zu einer bestimmten Gruppe mit einfließen zu lassen.

4 Anwendungsbeispiel und Vergleich der Modelle

Dieser Abschnitt dient der Veranschaulichung der beiden im vorherigen Abschnitt vorgestellten Analysemethoden anhand eines Ausfallbildes einer speziellen Motorkomponente einer ausgewählten Mercedes-Benz Baureihe.

Der Ablauf des Analyseprozesses orientiert sich an den sechs Teilschritten des bereits in Abschnitt 1 erwähnten CRISP-Modells. Diese beinhalten das Verstehen der Aufgabe (*Business Understanding*), das Verständnis der Daten (*Data Understanding*), die Datenvorbereitung (*Data Preparation*), der Modellbildung (*Modelling*), der Evaluation und dem Einsatz im Unternehmen (*Deployment*) [2]. Die Abbildung 1 zeigt schematisch dessen prinzipiellen Ablauf.



Abbildung 2: CRISP-Modell [2]

Sämtliche Aufbereitungs-, Berechnungs- und Modellierungsschritte wurden mithilfe der IBM Software SPSS-Modeler durchgeführt [7].

Schritt 1, das *Business Understanding*, ist im vorliegenden Fall klar definiert. Es werden Modelle gesucht, mit deren Hilfe es möglich ist, auf Basis von im Fahrzeug aufgezeichneten Kundenbelastungskollektiven und protokollierten Reparaturdaten eine Aussage über das Ausfallverhalten bestimmter risikobehafteter Komponenten treffen zu können. Die daraus gewonnenen Erkenntnisse können sowohl einen Anstoß zu möglichen Feldmaßnahmen, als auch Input für eine zukünftige, auf Kundenbelastung ausgerichtete Erprobungsplanoptimierung darstellen. Im zweiten Prozessschritt, dem *Data Understanding*, kommt es zu einer detaillierten Betrachtung der verfügbaren Daten, deren Bedeutung und der vorliegenden Qualität. Für den vorliegenden Anwendungsfall werden drei übergeordnete Informationsquellen mit folgenden Inhalten (hier standardisiert angegebenen) benötigt:

- Allgemeine Fahrzeuginformationen
 - Baureihe, Motor, Produktionsdatum, Erstzulassungsdatum, Vertriebsland
- Reparaturdaten
 - Fehlerort (Komponente), Fehlerart, Reparaturdatum, km-Stand
- Diagnosedaten
 - Diagnosedatum, BLK-Rohdaten, km-Stand

Mit dieser Kenntnis wird zu Schritt 3, der *Data Preparation*, übergegangen. In diesem Prozessschritt gilt es, die relevanten Daten anhand der festgelegten Attribute Baureihe, Motor, Produktionszeitraum, Fehlerort (Komponente) zu selektieren und durch geeignete

Verknüpfungen in eine konsistente Datentabelle zu überführen. Der schematische Ablauf der Datenzusammenführung ist in Abbildung 3 dargestellt.

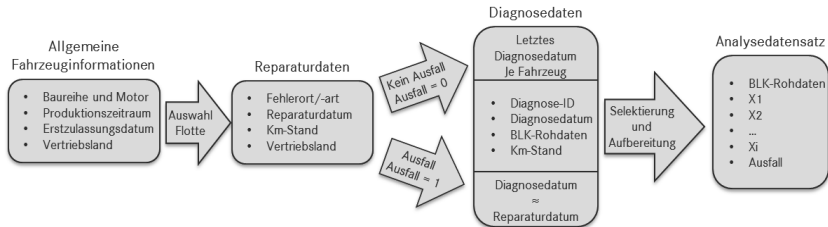


Abbildung 3: Synthese des Analysedatensatzes

Ein neues Attribut „Ausfall“ wird erstellt, welches sogleich die binäre Zielvariable des folgenden Analyseprozesses darstellt. Fahrzeuge, welche bezüglich der ausgewählten Komponente einen Ausfall vorzuweisen haben, werden mit einer 1 deklariert, nicht ausgefallene Fahrzeuge mit einer 0 (Survivor). Oftmals weisen Fahrzeuge mehr als nur eine Diagnosesitzung und somit eine Mehrzahl an Belastungskollektiv-Datensätzen auf, bei denen es gilt, den jeweils richtigen für die Modellbestimmung auszuwählen. Bei nicht ausgefallenen Fahrzeugen ist, unter der Annahme, dass es trotz der bisherigen Belastung zu keinem Ausfall der betrachteten Komponente gekommen ist, der aktuellste Datensatz zu verwenden, während bei ausgefallenen Fahrzeugen jener Datensatz relevant ist, welcher während des Werkstattaufenthaltes bzw. zeitnah zum Reparaturdatum aus den Steuergeräten ausgelesen wurde. Dieser Schritt ist von essentieller Wichtigkeit, da die Belastung nach der Reparatur keinerlei Einfluss auf das Ausfallverhalten der Komponente hat und es somit andernfalls zu einer Verfälschung des Modells kommen kann. Es resultiert letztendlich ein Analysedatensatz in Form der in Abbildung 3 dargestellten Struktur.

Durch Softwarefehler oder vom Flashen des Steuergerätes resultierende Zurücksetzung der Zähler bedingte fehlerhafte und nicht plausible BLK-Datensätze werden beseitigt-, und die Eingangsparameter auf deren Verteilungseigenschaften kontrolliert. Besonders hinsichtlich des Verfahrens der Linearen Diskriminanzanalyse besteht die Forderung, dass sämtliche Eingangsgrößen annähernd normalverteilt sind, um ein verlässliches Modell gewährleisten zu können.

Modellerstellung

Zum vierten und fünften Prozessschritt, dem *Modelling* und der *Evaluation*, wird der vorliegende Datensatz anhand einer Partition in zwei Datensätze aufgeteilt. Beim ersten Datensatz handelt es sich um das *Training-Set*, auf dessen Basis das Klassifizierungsmodell erstellt wird. Mit dem zweiten Teil, dem *Testing-Set*, wird das Modell validiert. In der Literatur wird eine Aufteilung von 70% Trainings-Data und 30% Testing-Data empfohlen [6], wovon in

den vorliegenden Analysen nicht abgewichen wurde. Die in dieser Arbeit verwendeten Analysemethoden bedienen sich der schrittweisen Hinzunahme eines jeden Parameters zum entstehenden Modell und einer damit verbundenen Überprüfung der jeweiligen Eingangsgrößen auf deren Beitrag zur Modellverbesserung. Parameter, die keinerlei positiven Einfluss haben, werden dem Modell nicht hinzugefügt [6]. Durch diesen Schritt wird ermittelt, in welchem Umfang die einzelnen Eingangsgrößen Einfluss auf die Trennung der beiden Gruppen „Ausfall“ und „Survivor“ nehmen. Man spricht in diesem Fall von einer *Predictor Importance*, welche nach vollendeter Modellerstellung anhand von parameterspezifischen Gewichtungsfaktoren dem Modell entnommen werden kann. Die Erkenntnis, welche Faktoren den größten Einfluss auf die Schädigung der jeweiligen Komponente haben, können eine wesentliche Grundlage für die zukünftige Weiter- bzw. Neuentwicklung der jeweiligen Fahrzeugkomponenten bilden.

Betrachtet wird im vorliegenden Anwendungsbeispiel eine Flotte von insgesamt 35.647 Fahrzeugen, von denen 2.123 eine Beanstandung des betrachteten Bauteils aufweisen. Aufgrund der verhältnismäßig großen Anzahl an Fahrzeugen und Ausfällen besteht die Möglichkeit, neben der Modellerstellung für die Gesamtflotte, bestimmte vorklassifizierte Gruppen separat zu analysieren um eine möglicherweise höhere Modellgüte zu erreichen. Da erfahrungsgemäß innerhalb der Vertriebsländer deutliche Unterschiede sowohl bezüglich des Kundenfahrverhaltens, als auch dem Ausfallverhalten bestehen, wurde als eine Untergruppierung das „Vertriebsland Y“ untersucht. Eine weitere Möglichkeit der Klassifizierung bildet eine Kundensegmentzuordnung auf Basis bekannter realer Belastungsprofile aus dem Feld. Je nach Nutzung kann mithilfe geeigneter Verfahren jedem Fahrzeug im Feld ein Kundensegment zugeordnet werden. Da es sich bei den in diesem Fall verwendeten Eingangsgrößen um relative Werte wie beispielsweise „Laufleistung pro Tag“ handelt, spielt die absolute Dauer bestimmter Belastungen keine Rolle. Diese Klassifizierung dient ausschließlich der Beschreibung des reinen allgemeinen Kundenfahrverhaltens. Das in diesem Beitrag ausgewählte Kundensegment wird exemplarisch als „Kundenprofil Y“ bezeichnet.

Für folgende Gruppen wurden sechs Analysemodelle (Tabelle 1) erstellt, jeweils ein Modell der Linearen Diskriminanzanalyse LDA und eines der Logistischen Regression LR.

Tabelle 1: Modellübersicht

Gruppe	Fahrzeuge	Ausfälle	Ausfallquote	LDA	LR
Gesamt	35.674	2.123	6,13%	LDA_G	LR_G
Kundenprofil X	7.627	226	3,31%	LDA_X	LR_X
Vertriebsland Y	3.815	68	1,84%	LDA_Y	LR_Y

Ergebnisse

Das Diagramm 1 zeigt die prozentualen Anteile der korrekt- und inkorrekt klassifizierten Fahrzeuge der jeweiligen Analysemodelle. Es ist ersichtlich, dass eine Vorklassifizierung generell positiven Einfluss auf die allgemeine Trefferquoten nehmen kann.

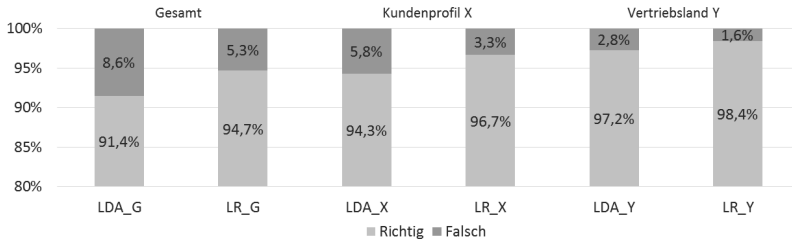


Diagramm 1: Gesamtgüte der Modelle

Einen detaillierteren Einblick liefern die Diagramme 2 und 3. Sie geben Aufschluss darüber, wie viele Fahrzeuge der jeweiligen Gruppe fälschlicherweise der anderen zugeordnet wurden. Beim Identifizieren von Ausfällen ist in den beiden Kundenprofil-Modellen eine Verschlechterung der Klassifizierungsgüte erkennbar. Dies lässt sich auf die Art der Vorklassifizierung zurückführen. Die Bestimmung eines spezifischen Kundenprofils beruht auf der Analyse des allgemeinen Fahrverhaltens. Somit wird automatisch die Streuung zwischen den beiden Gruppen Ausfall und Survivor reduziert, was eine größere Schnittmenge der beiden Verteilungen zur Folge hat, wodurch die Wahrscheinlichkeit einer fehlerhaften Zuordnung besonders innerhalb der kleinen Gruppe der ausgefallenen Fahrzeuge steigt. Die Verschlechterung der jeweiligen Gütemaße F und $-2LL$ bestätigen dies. Im Fall Vertriebsland ist dieses Phänomen aufgrund der großen Varianz nicht zu erkennen. Diese Erkenntnis zeigt, dass die Wahl der Vorklassifizierung einen signifikanten Einfluss auf die später resultierende Modellgüte linearer Verfahren nimmt und deshalb von besonderer Wichtigkeit ist.

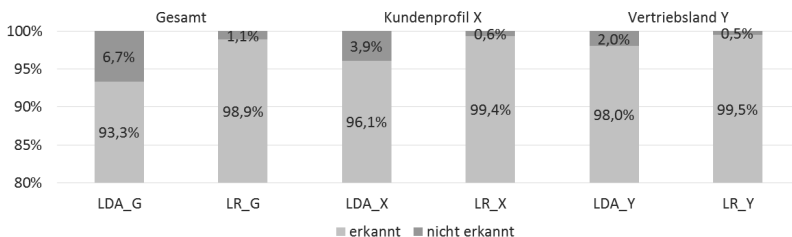


Diagramm 2: Erkannte und nicht erkannte Survivor

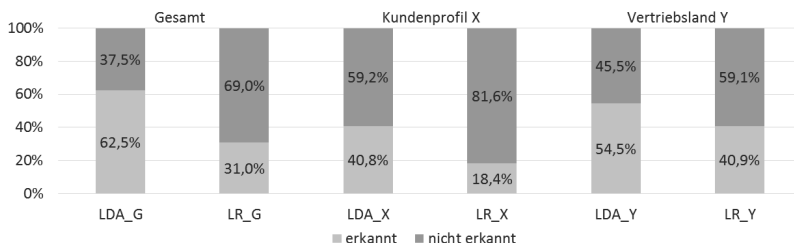


Diagramm 3: Erkannte und nicht erkannte Ausfälle

Die vorliegenden Ergebnisse lassen möglicherweise den Anschein gewisser Zweifel bezüglich der Modelltauglichkeit im Hinblick auf die Vorhersagekraft und Prognosegenauigkeit erwecken, da ein nicht zu verachtender Anteil falsch klassifiziert wurde. Jedoch birgt selbst eine Früherkennung von beispielsweise 41% der Ausfälle (LR_Y) für den Fahrzeughersteller ein enormes Potential hinsichtlich Kosteneinsparung und Kundenzufriedenheit. Dieses Potential ist nicht allein abhängig vom prozentualen Anteil der entdeckten Fehler, sondern ergänzt sich mit dem Anteil korrekt prognostizierter Zuordnungen, deren Verteilungen in den Diagrammen 4 und 5 abgebildet sind.

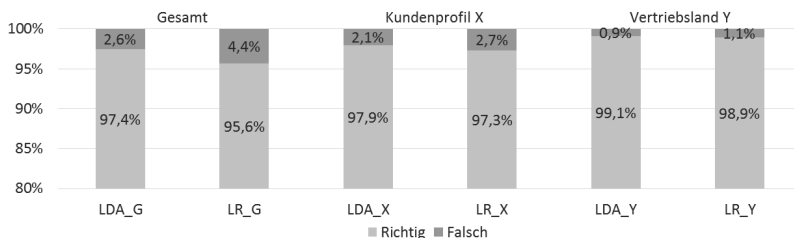


Diagramm 4: Prognostizierte Survivor

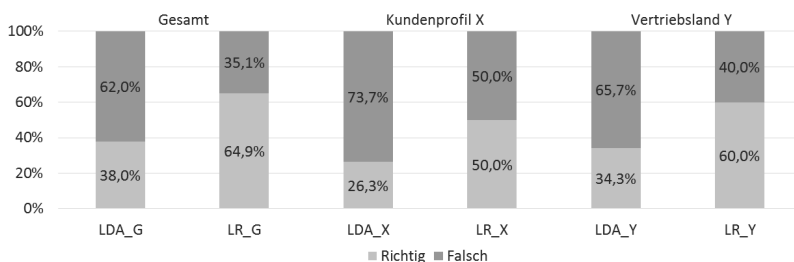


Diagramm 5: Prognostizierte Ausfälle

Es erfolgt eine Angabe, wie groß der Anteil korrekt prognostizierter Ausfälle / Nicht-Ausfälle der jeweiligen Modelle ist. Hier liefern die Modelle der logistischen Regression vielversprechende Ergebnisse. Hinsichtlich des Anteils prognostizierter Ausfälle muss betont

werden, dass es sich hierbei nicht um falsch klassifizierte Fahrzeuge handelt. Vielmehr bildet dieser Anteil eine Gruppe an Fahrzeugen im Feld, welche ein kritisches Nutzungsmuster aufweist, jedoch zum Zeitpunkt der Betrachtung als Survivor noch keinen Ausfall zu verzeichnen hat. Somit besteht für diese Gruppe weiterhin ein erhöhtes Ausfall-Risiko und verlangt nach einer gezielten Beobachtung.

Die Auswertung der *Predictor Importance* aller Modelle zeigt, dass bestimmte Einflussgrößen wie beispielsweise „geringe Geschwindigkeit“, „geringe Drehzahlen“ und „hohe Drehzahl bei hohem Drehmoment“ das Ausfallverhalten der betrachteten Komponente signifikant beeinflussen. Der jeweilige Fachbereich konnte den schädigenden Einfluss dieser Größen bestätigen. Dabei sei erwähnt dass die genannten Schadensparameter nicht zwingend die einzigen sein müssen, welche Einfluss auf die Zuverlässigkeit der Komponente nehmen. Viele Faktoren komplexer Schadensmechanismen werden in den Umfängen der Belastungskollektiv-Datenerhebung möglicherweise gar nicht erfasst. Aus diesem Grund befinden sich Hersteller im stetigen Kreislauf zwischen Schadensparameterermittlung, Definition neuer Belastungskollektive und deren Integration in die Applikation der unterschiedlichen Fahrzeugsteuergeräte.

Das Diagramm 6 zeigt eine ermittelte Verteilung des Einflusses bestimmter anonymisierter Belastungskollektive BLK1 bis BLK10. Demnach bietet in diesem Fall die schrittweise Hinzunahme der Eingangsparameter eine sehr gute Möglichkeit, schädigende Einflussfaktoren zu identifizieren.

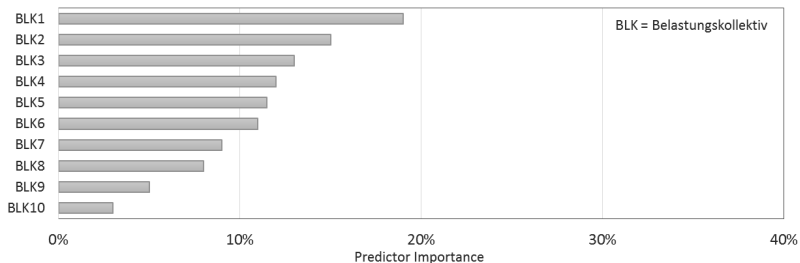


Diagramm 6: Einfluss der einzelnen Belastungskollektive auf das Modell (*Predictor Importance*)

Im letzten Prozessschritt des CRISP-Modells, dem *Deployment*, können nun die entwickelten Analysemodelle verwendet werden, um Fahrzeuge, welche neue BLK-Daten aus dem Feld liefern, hinsichtlich ihres Ausfallrisikos zu untersuchen und gegebenenfalls präventive Maßnahmen seitens des Fahrzeugherstellers einzuleiten. In welchem Umfang dies geschieht, muss unter Betrachtung der Faktoren Kosten, Kundenzufriedenheit und Aufwand festgelegt werden.

Die ermittelten Schädigungsparameter dienen zur zukünftigen Weiter- oder Neuentwicklung der betrachteten Komponente, sowie zum Abgleich der Erprobungsplanung. Geplante

Fahrprofile von Erprobungs- und Dauerlauffahrzeugen können nach einer Überführung der im Fahrzeug aufgezeichneten Zeitsignalen in das Datenformat der Belastungskollektive, unter der Verwendung der entwickelten Modelle überprüft werden, inwieweit das im Feld beobachtete kritische Belastungsmuster in der Erprobung abgebildet wird und unter welcher Wahrscheinlichkeit die Schädigung der Komponente in diesem Fall überhaupt entdeckt werden kann.

5 Fazit und Ausblick

Unter der Orientierung an den sechs Teilschritte des CRISP-Modells, wurde in diesem Beitrag der Einsatz ausgewählter Data Mining Methoden in Form zweier linearen multivariaten Analyseverfahren beschrieben, mit dem Ziel Fahrzeugausfälle auf Basis von Belastungskollektiven und Reparaturdaten prognostizieren zu können. Die Ergebnisse zeigen, dass bereits verhältnismäßig einfache lineare Verfahren vielversprechende Erkenntnisse bezüglich Ausfallmechanismus und auftretender Schadensparameter liefern. Die Art der Vorklassifizierung spielt hinsichtlich der Modellgüte dabei eine signifikante Rolle und bedarf weiterer Entwicklungen. Speziell nicht-lineare Verfahren in Form von Cluster-Analysen und neuronalen Netzen sind dafür in Betracht zu ziehen. Es ist die Kombination geeigneter Verfahren, die ein Analysemodell erfolgreich erscheinen lassen, nicht die einzelne Methode an sich.

6 Literaturverzeichnis

- [1] Köttermann, T.; Jacobi, A.; Bracke, S.: „Anwendung multivariater Methoden auf automobile Daten zur lastbasierten Zuverlässigkeitsanalyse“. Tagung Technische Zuverlässigkeit, Leonberg, 2015
- [2] Cleve, J.; Lämmel, U.: „Data Mining“. Oldenbourg Wissenschaftsverlag GmbH, S.733-S.737, 2014
- [3] Bergmeir, P.; Nitsche, C.; Nonnast, J.; Bargende, M.: „Classifying component failures of a hybrid electric vehicle fleet based on load spectrum data“ The Natural Computing Applications Forum 2015
- [4] Jacobi, A.; Luy, J.-F.: „Zuverlässigkeitssteuerung bei Mercedes-Benz Cars“. Kongress QS Excellence, Bad Nauheim, 2016
- [5] Backhaus, K.; Erichson, B.; Plinke, U.; Weiber, R.: „Multivariate Analysemethoden“ Springer-Verlag, 14. Auflage, S.216-S.314, 2016
- [6] Wendler, T.; Grötrup, S.: „Data Mining with SPSS Modeler“. Springer-Verlag, S.733-S.737, 2016
- [7] IBM. SPSS Modeler 17 Handbook. International Business Machines Corporation, 2015

Stochastische Sensitivitätsmaße für smarte Systeme

Stochastic sensitivity indices for smart systems

Steffen Ochs, M.Sc., Dr.-Ing. **Elena Maja Slomski**, Prof. Dr.-Ing. **Tobias Melz**, Fachgebiet Systemzuverlässigkeit, Adaptronik und Maschinenakustik SAM, TU Darmstadt, Darmstadt



Kurzfassung

Die Zuverlässigkeitsbewertung adaptronischer (smarter) Systeme stellt, obwohl sie in den letzten Jahren zunehmend Aufmerksamkeit erfährt, ein immer noch wenig erforschtes Gebiet dar. Durch die Einbringung zusätzlicher aktiver Komponenten in eine zuvor passive Struktur werden Wechselwirkungseffekte zwischen den aktiven und den passiven Komponenten erzeugt, die das Systemverhalten erheblich verändern können. Diese Wechselwirkungen sind einerseits erwünscht, da sie den Funktionsbereich des Systems erweitern. Andererseits erschweren sie die Beurteilung der Systemzuverlässigkeit aufgrund der Entstehung vielfältiger Systemzustände. Am Fachgebiet Systemzuverlässigkeit, Adaptronik und Maschinenakustik SAM der TU Darmstadt wird der Ansatz verfolgt, anhand stochastischer Sensitivitätsanalysen die Zuverlässigkeit smarter Systeme mit statistischen Maßen (sog. Sensitivitätsmaßen) zu bewerten. Der vorliegende Beitrag befasst sich mit der gezielten Auswahl und Verknüpfung von stochastischen Sensitivitätsanalysemethoden zur quantitativen Bewertung von smarten Struktursystemen. Besonderer Wert wird dabei auf die Darstellung und die Bewertung der stochastischen Sensitivitätsmaße gelegt. Bisherige Analysemethoden ermöglichen lediglich eine normierte Betrachtung der stochastischen Sensitivitätsmaße. Durch die Adaption von Darstellungsverfahren aus dem Bereich der qualitativen Sensitivitätsanalysen ist es gelungen, die stochastischen Sensitivitätsmaße einheitenbehaftet darzustellen. Somit können Wechselwirkungseffekte eindeutig identifiziert und der Einfluss von Haupt- und Wechselwirkungseffekten auf das Systemverhalten quantifiziert werden.

Abstract

The reliability quantification of smart systems is still an area that is barely researched, although it has received increasing attention in the past years. The integration of active components into a passive structure initiates interactions between the active and the passive

components that considerably affect the system behavior. On the one hand, these interactions are desired as they extend the functional range of the system. On the other hand, they make the quantification of the system reliability more difficult due to the variety of system conditions that occur. The research group System Reliability, Adaptive Structures, and Machine Acoustics SAM of Technische Universität Darmstadt pursues the approach of quantifying the reliability of smart systems with statistical measures (so-called sensitivity indices) using stochastic sensitivity analyses. The present paper focuses on the systematic selection and combination of qualitative and quantitative methods of sensitivity analysis to improve the description and evaluation of stochastic sensitivity indices. Previous methods allow only a standardized evaluation of the stochastic sensitivity indices. By adapting methods from the field of qualitative sensitivity analysis we have succeeded in representing the stochastic sensitivity indices by means of physical units. Thus, interaction effects can be clearly identified and the influence of main and interaction effects on the system's behavior are quantified.

1. Einleitung

Die Adaptronik ist eine Strukturtechnologie, die durch Verwendung fortschrittlicher Methoden der Strukturdynamik und Signalverarbeitung sowie unter Einbeziehung neuartiger Aktoren und Sensoren eine Überwachung und Verbesserung mechanischer Eigenschaften von Produkten ermöglicht [1]. Aufgrund ihrer Möglichkeit, sich an verändernde Randbedingungen anzupassen, werden diese Produkte auch als smarte Systeme bezeichnet [2]. Neben der Steigerung des Funktionsumfangs verursacht die Integration von aktiven Komponenten in die mechanische Struktur aber auch komplexe Wechselwirkungen zwischen den Eingangsparametern des Systems. Die sich daraus ergebende Vielfältigkeit an Systemzuständen erschwert die Auslegung, Optimierung und Zuverlässigkeitsanalyse des smarten Systems. Ein denkbare Einsatzgebiet für smarte Systeme ist potenziell überall dort, wo aktiv Strukturschwingungen reduziert werden sollen. Ein einseitig fest eingespannter Euler-Bernoulli-Kragbalken mit einem kollozierten flächigen Piezoaktor und -sensorpaar, welches über einen Positive-Position-Feedback-Regler (PPF-Regler) verbunden ist, dient in diesem Beitrag als Beispielsystem zur Demonstration des dynamischen Verhaltens eines Anbindungssystems mit aktiver Unterdrückung störender Schwingungen durch piezokeramische aktive Elemente. Die Störkraft $F(t)$, welche die Struktur zu Schwingungen $w(x, t)$ anregt, greift in einer vorgegebenen Entfernung a_F zum eingespannten Balkenende an der Balkenstruktur an (Bild 1).

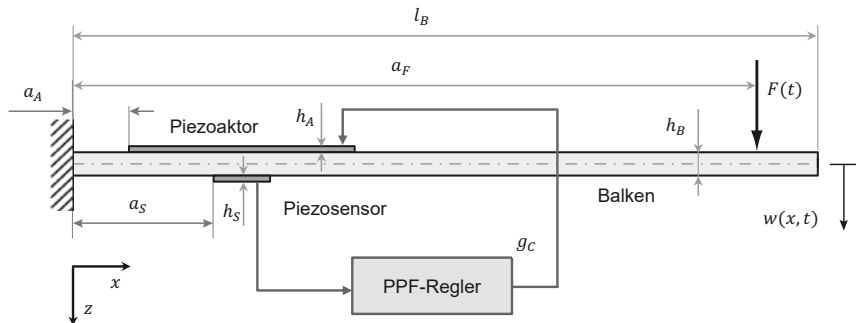


Bild 1: Schematischer Aufbau des smarten Balkensystems in Draufsicht.

Zur analytischen eindimensionalen Beschreibung des Systems werden insgesamt 22 Eingangsparameter benötigt (Tabelle 1).

Tabelle 1: Eingangsparameter für das Modell des smarten Balkensystems.

Sym.	Parameter	Wert	Einheit	Sym.	Parameter	Wert	Einheit
h_B	Balkenhöhe	3	mm	l_S	Sensorlänge	10	mm
l_B	Balkenlänge	200	mm	b_S	Sensorbreite	10	mm
b_B	Balkenbreite	40	mm	a_S	Position des Sensors	35	mm
ρ_B	Dichte des Balkens	2700	kg/m ³	ρ_P	Dichte der Piezokeramik	7800	kg/m ³
E_B	E-Modul des Balkens	70,0	GPa	E_P	E-Modul der Piezokeramik	62,1	GPa
ζ	Strukturdämpfungsverhältnis	1,5	%	$d_{31,P}$	Ladungskonstante	$-1,8 \cdot 10^{-10}$	m/V
h_A	Aktorhöhe	0,8	mm	g_C	PPF-Verstärkungsfaktor	0,9	s ²
l_A	Aktorlänge	50	mm	ω_C	PPF-Eigenkreisfrequenz	$2\pi \cdot 60$	1/s
b_A	Aktorbreite	30	mm	ς_C	PPF-Dämpfungskoeffizient	0,5	-
a_A	Position des Aktors	15	mm	C_f	Ladungsverstärkerkapazität	1	nF
h_S	Sensorhöhe	0,5	mm	a_F	Position der Kraftanregung	200	mm

Im Rahmen dieses Beitrags werden Ergebnisse zu den Sensitivitätsanalysen der Parameter Balkenhöhe h_B , Aktorhöhe h_A , Sensorhöhe h_S und Position der Kraftanregung a_F detailliert dargestellt. Die gesamte Herleitung der Bewegungsgleichungen eines Euler-Bernoulli-Balkens mit applizierten piezokeramischen Aktoren und Sensoren bei gegebenen Randbedingungen sowie die Einbindung einer Positive-Position-Feedback-Regelung sind im Detail in OCHS ET AL. [3] beschrieben.

Ein Ansatz zur systematischen Zuverlässigkeitsbewertung von adaptronischen Systemen wurde beispielsweise bereits 2006 von HASSAN [4] vorgestellt. Hierbei wird jede mögliche Fehlfunktion qualitativ über eine Fehlerbaumanalyse erfasst und es wird jeweils eine Grenzfunktion zugewiesen. Allerdings wird auf die Berechnung der Strukturzuverlässigkeit, d. h. auf die Lösung der hochdimensionalen Integralterme, verzichtet. HAN [5] empfiehlt für die Zuverlässigkeitsanalyse adaptronischer Systeme alternativ die Verwendung von globalen Sensitivitätsanalysen (GSA). Die zugehörigen Sensitivitätsmaße werden bei GSA anhand des gesamten Parameterraums quantifiziert. Somit können neben dem direkten Einfluss (Haupteffekt) eines Parameters auch dessen Wechselwirkungen analysiert werden. Der am häufigsten verwendete globale Ansatz basiert gemäß SALTELLI ET AL. [6] auf einer Varianzanalyse des Systemverhaltens in Kombination mit einem Monte-Carlo-Sampling als Grundlage zur Berechnung von quantitativen Sensitivitätsmaßen. Die Darstellungsform der stochastischen Ergebnisse ist jedoch bis heute für die kommerzielle Nutzung unzureichend. Durch das in diesem Beitrag vorgestellte Bewertungsverfahren (siehe Kapitel 3), bei dem Auswertemethoden der qualitativen und der quantitativen GSA miteinander kombiniert werden, wird die Bewertung der berechneten Sensitivitätsmaße erleichtert, wodurch ihre Anwendbarkeit im industriellen Alltag praktikabel wird.

2. Globale Sensitivitätsanalysen

Unter dem Begriff Sensitivitätsanalyse werden unterschiedliche Verfahren zur Untersuchung des Zusammenhangs zwischen den Eingangsparametern und den Ausgangsgrößen eines Systems zusammengefasst. Die Wahl der Sensitivitätsanalysemethode wird hauptsächlich durch das verfolgte Ziel der Untersuchung, das Simulationsmodell mit seinen Eingangsparametern und das betrachtete Systemverhalten (Ausgangsgrößen) bedingt. Für Ausgangsgrößen, die sich durch eine Linearkombination der Eingangsparameter gut approximieren lassen, sind konventionelle Verfahren wie die Korrelations- und Regressionsanalyse hinreichend geeignet. Der entscheidende Nachteil dieser Methoden besteht in der lokal eingeschränkten Betrachtung und der damit verbundenen fehlenden Möglichkeit, Wechselwirkungen eindeutig zu identifizieren [7]. Bei komplexeren nichtlinearen Zusammenhängen, wie sie bei smarten Systemen häufig auftreten, ist daher der Einsatz globaler Verfahren zu empfehlen. Globale Verfahren der Sensitivitätsanalyse setzen keine Einschränkung des berücksichtigten Parameterraums voraus. Die zugehörigen Sensitivitätsmaße werden somit anhand des gesamten Wertebereichs der Parameter bestimmt. Bei jeder GSA muss zur Erstellung der benötigten Daten aus der Grundgesamtheit eine Auswahl an Kombinationen der Parameterwerte, auch Realisationen genannt, getroffen werden. Die

getroffene Auswahl an Wertekombinationen wird als Stichprobe oder Sample bezeichnet. Die erstellte Stichprobe soll im Allgemeinen den vorgegebenen Parameterraum optimal ausfüllen. Der Vorgang des Auswählens wird als Stichprobenerstellung oder Sampling bezeichnet.

Ist man lediglich an einer qualitativen Bewertung, beispielsweise in Form der Gruppierung in relevante und irrelevante Parameter, interessiert, stellt ein Screeningverfahren einen geeigneten Ansatz dar. Ein Verfahren ist beispielsweise das Morris-Screening [8]. Bei der qualitativen GSA erfolgt das Sampling üblicherweise anhand eines groben Gitters, welches auf den ursprünglichen Parameterraum, auch Hyperkubus genannt, angewendet wird. Der Hyperkubus Ω^k bezeichnet den Parameterraum Ω , der von k Parametern aufgespannt wird. Die Anzahl an Realisationen wird bewusst niedrig gehalten, um eine schnelle Analyse zu gewährleisten. Die aus dieser Stichprobe berechneten Sensitivitätsmaße können nur zur qualitativen Bewertung des Systemverhaltens herangezogen werden, da durch das grobe Raster nicht alle Systemzustände erfasst werden können.

Ist das Ziel der Untersuchung hingegen eine quantitative Bewertung des Systemverhaltens, muss auf eine stochastische GSA zurückgegriffen werden. Die am häufigsten verwendete Methode ist, wie bereits zuvor erwähnt, die Varianzanalyse in Kombination mit einem Monte-Carlo-Sampling. Bei quantitativen Methoden erfolgt das Sampling meist anhand stochastisch verteilter Zufallszahlen, auch als Monte-Carlo-Simulation (MCS) bekannt. Der Stichprobenumfang ist im Vergleich zur Rastersuche um einen Faktor ≥ 50 größer, was einerseits eine quantitative Bewertung des Systemverhaltens erlaubt, andererseits aber auch eine deutlich längere Analysezeit benötigt. Die Erzeugung der Zufallszahlen erfolgt anhand von Zufallszahlengeneratoren. Diese erzeugen im Allgemeinen Zahlenwerte zwischen 0 und 1 aus dem Einheitshyperkubus Ω^k . Die Berechnung der tatsächlichen Variablenwerte erfolgt, indem die generierten Zufallszahlen als Quantilwerte betrachtet und auf den ursprünglichen Parameterraum angewendet werden. In diesem Beitrag werden Quasizufallszahlen verwendet. Diese werden im eigentlichen Sinne nicht zufällig erstellt, sondern aus deterministischen Zahlensequenzen entnommen. Dies hat den Vorteil, dass die Zahlenfolgen derart konstruiert werden können, dass die gewonnene Stichprobe möglichst gleichförmig den Raum des Hyperkubus ausfüllt und somit eine schnellere Konvergenz der Sensitivitätsmaße erreicht wird [6]. Ein schnelleres Konvergenzverhalten bedeutet zugleich eine Verringerung des Simulationsaufwands. Zur Erstellung dieser Zahlensequenzen existieren unterschiedliche Ansätze, wobei in diesem Beitrag die Sobol'-Sequenz [9], als derzeit verbreitetster Ansatz, verwendet wird.

Die qualitative und quantitative GSA unterscheidet sich folglich in der Art des Samplings. Damit einhergehend erfolgt auch die Wahl der Bewertungsmethode zur Berechnung der Sensitivitätsmaße, auf die im nächsten Kapitel genauer eingegangen wird. Eine ausführliche Beschreibung der Samplingverfahren für die GSA ist beispielsweise in SALTELLI ET AL. [7] wiederzufinden.

3. Bewertung der Sensitivitätsmaße

Bei der quantitativen GSA mittels Varianzanalyse werden die zugehörigen quantitativen Bewertungsmaße durch das Verhältnis der durch einen Parameter $i = 1$ bis k erzeugten Varianz in der Systemantwort bezogen auf die Gesamtvarianz des Systemverhaltens beschrieben. Das Sensitivitätsmaß 1. Ordnung, der Haupteffekt, beschreibt quantitativ den direkten Einfluss eines Parameters. Er berechnet sich anhand der vom Parameter X_i verursachten Varianz bezogen auf die komplette Varianz der Ausgangsgröße Y im Parameterraum X [6].

$$S_{H_i} = \frac{\text{Var}[E(Y|X_i)]}{\text{Var}[Y]} \quad (1)$$

Als relative Größe kann er Werte zwischen 0 (kein direkter Einfluss) und 1 (starker direkter Einfluss) annehmen und ermöglicht somit einen quantitativen Vergleich der Einflüsse verschiedener Eingangsparameter, bezogen auf die Gesamtvarianz.

Der Totaleffekt hingegen gibt den gesamten Einfluss eines Parameters auf die betrachtete Ausgangsgröße wieder und fasst folglich den direkten Einfluss (Sensitivität 1. Ordnung) und alle Wechselwirkungen (Sensitivitäten höherer Ordnung) des Parameters zusammen [10].

$$S_{T_i} = 1 - \frac{\text{Var}[E(Y|\mathbf{X}_{-i})]}{\text{Var}[Y]} \quad (2)$$

Der Totaleffekt ist ebenfalls eine relative Größe. Der Ausdruck \mathbf{X}_{-i} bezeichnet dabei alle Eingangsparameter ohne X_i . Da der Totaleffekt sowohl die Sensitivität 1. Ordnung als auch darüber hinaus die Sensitivitäten der Wechselwirkungen beinhaltet, ist sein Wert gleich dem des Haupteffekts oder höher. Eine große Differenz zwischen dem berechneten Haupt- und dem berechneten Totaleffekt eines Parameters weist in Konsequenz auf starke Wechselwirkungen dieses Parameters mit einem oder mehreren anderen Parametern hin.

Die analytischen Lösungen der Gleichungen (1) und (2) sind für die meisten adaptiven Systeme, wie auch für das Beispielsystem, nicht berechenbar. Aus diesem Grund müssen die Sensitivitätsmaße über Schätzverfahren bestimmt werden. HAN [5] hat gezeigt, dass bei der Sensitivitätsanalyse strukturdynamischer Systeme die Anwendung von statistischen

Verfahren der Anwendung von spektralen Verfahren vorzuziehen ist. Die dort verwendeten statistischen Schätzer, eingeführt von SOBOL' [9] für den Haupteffekt und von JANSEN [11] für den Totaleffekt, kommen auch im vorliegenden Beitrag wie folgt zur Berechnung der Sensitivitätsmaße zum Einsatz.

$$\begin{aligned} \text{Haupteffekt-Schätzer:} \quad \hat{S}_{H_i} &= \frac{\frac{1}{N-1} \cdot \sum_{r=1}^N f_{B,r} \cdot [f_{A_{B(i)},r} - f_{A,r}]}{\frac{1}{2N-1} \cdot \sum_{r=1}^{2N} (f_{AB,r} - \frac{1}{2N} \sum_{r=1}^{2N} f_{AB,r})^2} \end{aligned} \quad (3)$$

$$\begin{aligned} \text{Totaleffekt-Schätzer:} \quad \hat{S}_{T_i} &= \frac{\frac{1}{2N} \cdot \sum_{r=1}^N (f_{A,r} - f_{A_{B(i)},r})^2}{\frac{1}{2N-1} \cdot \sum_{r=1}^{2N} (f_{AB,r} - \frac{1}{2N} \sum_{r=1}^{2N} f_{AB,r})^2} \end{aligned} \quad (4)$$

Die Analyse von k Eingangsparametern erfolgt dabei durch die Durchführung und Auswertung von $N \cdot (k + 2)$ Systemsimulationen mit den Simulationswerten f_{κ} , $\kappa \in (\mathbf{A}, \mathbf{B}, \mathbf{AB}, \mathbf{A}_{B(i)})$, der betrachteten Ausgangsgröße. Für die Anwendung der Schätzverfahren werden zwei unabhängige Stichproben \mathbf{A} und \mathbf{B} der Größe $N \times k$ sowie deren Verknüpfungen $\mathbf{A}_{B(i)}$ benötigt. Diese werden durch das Einsetzen der i -ten Spalte von \mathbf{B} in die i -te Spalte der Matrix \mathbf{A} gebildet.

In der praktischen Anwendung der globalen Sensitivitätsanalyse ist die Bewertung der Sensitivitäten anhand der einheitenlosen Bewertungsmaße Haupt- und Totaleffekt nur bedingt zielführend, da direkte Maßnahmen nur schwierig abzuleiten sind. Ein weiterer Nachteil dieser relativen Größen ist die schlechte Vergleichbarkeit verschiedener Systemmodellierungsansätze, da sich die Größen jeweils nur relativ auf die Grundgesamtheit beziehen, welche z. B. beim Vergleich mit experimentellen Sensitivitätsanalysedaten nicht bekannt ist. Eine Umrechnung in einheitenbehaftete Größen ist nicht möglich.

In Anlehnung an die Vorgehensweise bei der Elementareffekt-Methode (EE-Methode) nach MORRIS [8] werden zur besseren Bewertung der Sensitivitäten daher zwei weitere Sensitivitätsmaße als Erweiterung der bisherigen stochastischen Analyse bestimmt. Ein Elementareffekt beschreibt die absolute Änderung der Ausgangsgröße bezogen auf die Änderung eines einzigen Eingangsparameters an einem beliebigen Ort im Parameterraum Ω und kann somit einheitenbehaftet dargestellt werden. Er fasst den direkten Einfluss eines Parameters i sowie seine Wechselwirkungen mit anderen Eingangsparametern am betrachtenden Ort im Parameterraum Ω zusammen. Die Schwierigkeit bei der Anwendung der EE-Methode ist, dass sie ursprünglich für die Anwendung auf ein Sampling für eine qualitative GSA erstellt wurde, bei dem die Simulationpunkte anhand eines vorgegebenen

Gitters eindeutig zueinander angeordnet sind. Hierdurch ist die Bedingung, dass bei einem Sprung zu einem anderen Simulationspunkt immer nur ein Eingangsparameter geändert wird, erfüllt. Diese eindeutige Zuordnung ist auf den ersten Blick bei einem Monte-Carlo-Sampling nicht gegeben, da die Generierung der Zufallszahlen und damit der Simulationspunkte stochastisch erfolgt. Wird das Monte-Carlo-Sampling jedoch wie bei der Verwendung der angesprochenen Schätzverfahren aufgebaut, so kann ein Elementareffekt wiederum über die Differenz der Simulationsergebnisse aus den Stichproben \mathbf{A} und $\mathbf{A}_{B(i)}$ wie folgt ermittelt werden:

$$EE_i^r = \frac{f_{\mathbf{A}_{B(i)},r} - f_{\mathbf{A},r}}{\Delta_i^r} \quad (5)$$

Δ_i^r beschreibt die auf die Differenz des oberen und unteren Grenzwertes des i -ten Parameters normierte Änderung für den r -ten Simulationsdurchlauf, während alle anderen Parameter unverändert bleiben. Die Gleichung erfüllt somit die Bedingung, dass immer nur ein Eingangsparameter verändert wird.

Durch die Berechnung des arithmetischen Mittelwerts der absoluten Elementareffekte eines Parameters i für alle Simulationsdurchläufe r wird das globale Sensitivitätsmaß μ_i^* wie folgt bestimmt:

$$\mu_i^* = \frac{1}{N} \sum_{r=1}^N |EE_i^r| \quad (6)$$

Der totale Einfluss des Parameters auf das betrachtete Systemverhalten wird somit anhand des Mittelwertes aller Elementareffekte eines Parameters i quantifiziert. Des Weiteren kann anhand der Standardabweichung σ_i der Elementareffekte eines Parameters i direkt Wechselwirkungen des Parameters mit anderen Eingangsparametern wie folgt berechnet werden:

$$\sigma_i = \sqrt{\frac{1}{N-1} \sum_{r=1}^N (EE_i^r - \overline{EE}_i)^2} \quad (7)$$

Zeigen sich deutliche Unterschiede in der Berechnung der Elementareffekte EE_i^r eines Parameters i für die unterschiedlichen Simulationsdurchläufe r , so deutet dies entweder auf starke Wechselwirkungen des Parameters mit anderen Parametern, auf ein nichtlineares Verhalten im Parameterraum oder eine Kombination von beidem hin.

Der Unterschied zwischen den ursprünglichen Sensitivitätsmaßen Haupteffekt \hat{S}_{H_i} und Totaleffekt \hat{S}_{T_i} einerseits und den neu eingeführten Sensitivitätsmaßen Elementareffekt-Mittelwert μ_i^* und -Standardabweichung σ_i bei einer quantitativen GSA von komplexen Systemen andererseits wird anhand des in Kapitel 1 eingeführten Beispielsystems im Folgenden erprobt.

4. Vergleich und Diskussion der Sensitivitätsmaße am Beispielsystem

In diesem Beitrag werden die Einflüsse von vier Eingangsparametern des Beispielsystems (siehe Tabelle 2) auf zwei Ausgangsgrößen des Systems, die erste Resonanzfrequenz Y_1 und die Amplitude Y_2 der Übertragungsfunktion zwischen Kraftanregung und Geschwindigkeit des freien Balkenendes an der ersten Resonanzfrequenz, untersucht.

Tabelle 2: In der GSA betrachtete Eingangsparameter und ihre Grenzwerte.

Symbol	Eingangsparameter	untere Grenze	obere Grenze	Einheit	Dichtefunktion
h_B	Balkenhöhe	2,55	3,45	mm	$\mathcal{U}(2,55; 3,45)$
h_A	Aktorhöhe	0,68	0,92	mm	$\mathcal{U}(0,68; 0,92)$
h_S	Sensorhöhe	0,425	0,575	mm	$\mathcal{U}(0,425; 0,575)$
a_F	Position der Kraftanregung	170	200	mm	$\mathcal{U}(170; 200)$

Der Parameterraum Ω^k dieser $k = 4$ Parameter wird über die Angabe von unteren und oberen Grenzwerten definiert. Innerhalb der Grenzwerte sind alle vier Eingangsparameter gleichverteilt. Die Grenzwerte orientieren sich an einer Variabilität der Parameter von $\pm 15\%$ zum Ursprungswert (siehe Tabelle 1), wie sie im Rahmen eines Konstruktionsprozesses auftreten kann. Die Position der Kraftanregung wird dabei nur in eine Richtung variiert, da sie in der anderen Richtung außerhalb der Abmessung des Balkens liegen würde. Es wurden insgesamt 36 000 Systemsimulationen durchgeführt, um die Konvergenz der Sensitivitätsmaße zu gewährleisten. Dies entspricht einer Stichprobengröße von $N = 6000$.

In Bild 2 sind die berechneten Sensitivitätsmaße Haupteffekt \hat{S}_{H_i} und Totaleffekt \hat{S}_{T_i} für beide Ausgangsgrößen Y_1 und Y_2 dargestellt. Zu erkennen sind die relativen Parametersensitivitäten bezogen auf die Gesamtvarianz der jeweiligen Ausgangsgröße. Es wird deutlich, dass die erste Resonanzfrequenz Y_1 im Kontext der vier betrachteten Eingangsparameter unmittelbar abhängig von der Balkenhöhe h_B ist. Dieses Verhalten ist plausibel, da die Eigenfrequenzen des Balkensystems zwar formal von den Eigenschaften der Piezokeramik (Aktorhöhe h_A und Sensorhöhe h_S) abhängig sind, aufgrund der geringen

Gesamtmasse der Piezokeramiken im Vergleich zur Balkenmasse weisen diese jedoch nur einen sehr geringfügigen Einfluss auf diese Ausgangsgröße auf. Auch die Position der Kraftanregung a_F hat keinen Einfluss auf die erste Resonanzfrequenz. Eine quantitative Aussage über die Verschiebung der ersten Resonanzfrequenz in Hz kann anhand dieser Sensitivitätsmaße nicht getroffen werden. Die Amplitude Y_2 der Übertragungsfunktion an der ersten Resonanzfrequenz ist im Vergleich zu Y_1 deutlich stärker von der Position der Kraftanregung a_F abhängig. Ebenso ist ein im Vergleich zur Balkenhöhe h_B und zur Position der Kraftanregung a_F geringer Einfluss der Parameter der aktiven Komponenten (Aktorhöhe h_A und Sensorhöhe h_S) des Systems erkennbar. Eine konstruktionsbedingte Abweichung von $\pm 15\%$ der passiven Komponente hat einen deutlich größeren Einfluss auf das Systemverhalten als die Parameter der aktiven Komponenten. Anhand dieser Sensitivitätsmaße kann jedoch keine Aussage über die Höhe der Abweichung von Y_2 in dB getroffen werden. Wechselwirkungen, erkennbar durch einen Unterschied zwischen dem Haupt- und dem Toteffekt eines Parameters, scheinen im betrachteten Versuchsraum nur eine untergeordnete Rolle zu spielen.

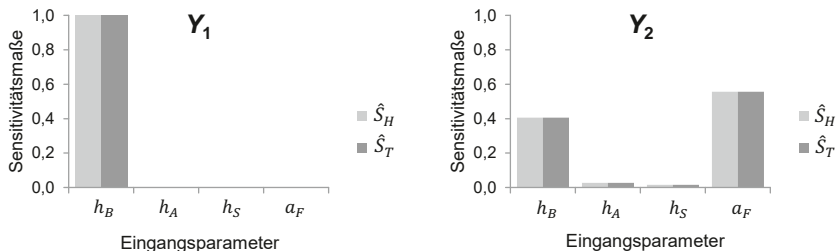


Bild 2: Haupt- und Toteffekte der vier betrachteten Eingangsparameter bezogen auf die Ausgangsgrößen Y_1 und Y_2 .

Die in Bild 3 dargestellten Sensitivitätsmaße μ_i^* und σ_i liefern qualitativ dieselben Beobachtungen wie die Sensitivitätsmaße \hat{S}_{H_i} und \hat{S}_{T_i} . Auch in diesem Betrachtungsfall wird die Ausgangsgröße Y_1 primär von der Balkenhöhe h_B und die Ausgangsgröße Y_2 von den Parametern der passiven Komponente beeinflusst. Durch die quantitative Darstellung der Effekte in der physikalischen Einheit der entsprechenden Ausgangsgröße können jedoch zusätzlich auch absolute Aussagen getroffen werden. Somit kann beispielsweise direkt abgelesen werden, dass im verwendeten Simulationsmodell eine maximale Veränderung der Balkenhöhe h_B um $\pm 15\%$ eine Verschiebung der ersten Resonanzfrequenz Y_1 um 20 Hz

hervorrufen, was einer deutlichen Änderung der Systemeigenschaften entspricht. Des Weiteren ist zu erkennen, dass die Effekte der drei weiteren Eingangsparameter h_A , h_S und a_F auf die erste Resonanzfrequenz unter 0,25 Hz liegen und somit das Systemverhalten nur marginal beeinflussen.

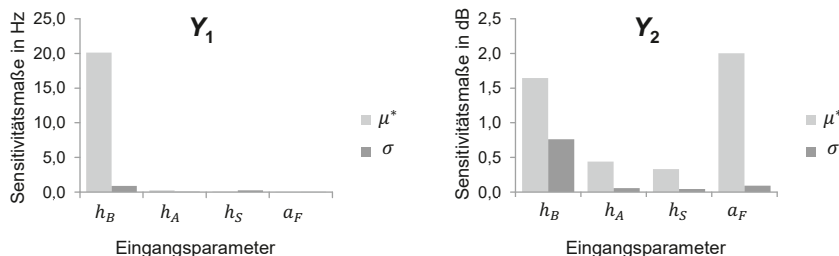


Bild 3: Elementar-Mittelwerte und -Standardabweichungen der vier betrachteten Eingangsparameter bezogen auf die Ausgangsgrößen Y_1 und Y_2 .

Für die Ausgangsgröße Y_2 , die Amplitude der Übertragungsfunktion zwischen Kraftanregung und Geschwindigkeit des freien Balkenendes an der ersten Resonanzfrequenz, kann zudem abgelesen werden, dass eine Veränderung der vier Eingangsparameter um $\pm 15\%$ nur eine Veränderung der Amplitude um max. 2 dB hervorruft. Absolut gesehen reagiert das smarte System somit robust gegenüber konstruktionsbedingten Änderungen dieser vier Eingangsparameter. Des Weiteren ist zu beobachten, dass der Einfluss der Balkenhöhe h_B über dem Parameterraum Ω streut und auf Wechselwirkungseffekte, ein nichtlineares Verhalten im Parameterraum oder eine Kombination von beidem geschlossen werden kann, während dies in der ursprünglichen Darstellung in Bild 2 nicht ablesbar ist.

5. Fazit

Durch die Anpassung und Anwendung von Bewertungsmaßen aus der qualitativen globalen Sensitivitätsanalyse auf Stichproben aus der quantitativen globalen Sensitivitätsanalyse wird die Voraussetzung für eine praktikable Anwendbarkeit von stochastischen Sensitivitätsmaßen geschaffen. Die Sensitivitätsmaße Elementareffekt-Mittelwert und Elementareffekt-Standardabweichung liefern dabei qualitativ dieselben Beobachtungen wie die bekannten Sensitivitätsmaße Haupt- und Totaleffekt. Zusätzlich ermöglicht die Bewertung der Sensitivitäten anhand der beiden neu eingeführten einheitenbehafteten Sensitivitätsmaße eine Quantifizierung der Einflüsse und lässt beispielsweise einen direkten Vergleich

verschiedener Modellierungsansätze smarter Systeme sowie eine experimentelle Validierung der Modelle mittels Messungen am Prüfstand zu.

6. Literaturangaben

- [1] Nuffer, J., Hanselka, H.: Intelligente Materialien. In: Technologieführer (2007) S. 36–45
- [2] Melz, T., Wiedemann, M. (Hrsg.): Smarte Strukturen und Systeme. Tagungsband des 4SMART-Symposiums. Oldenbourg: De Gruyter 2016
- [3] Ochs, S., Li, S., Adams, C., Melz, T.: Efficient Experimental Validation of Stochastic Sensitivity Analyses of Smart Systems. Smart Structures and Materials: Selected Papers from the 7th ECCOMAS Thematic Conference on Smart Structures and Materials. Computational Methods in Applied Sciences 43 (2017) S. 97–113
- [4] Hassan, M. H. M.: A system model for reliability assessment of smart structural systems. Structural Engineering and Mechanics 23(5) (2006) S. 455–468
- [5] Han, S.: Varianzbasierte Sensitivitätsanalyse als Beitrag zur Bewertung der Zuverlässigkeit adaptiver Struktursysteme. TU Darmstadt Diss. 2011
- [6] Saltelli, A. et al.: Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Computer Physics Communications 181(2) (2010) S. 259–270
- [7] Saltelli, A. et al.: Global Sensitivity Analysis - The Primer. Chichester, England: John Wiley & Sons 2008
- [8] Morris, M. D.: Factorial Sampling Plans for Preliminary Computational Experiments. Technometrics 33 (1991) S. 161–174
- [9] Sobol', I. M.: Sensitivity Estimates for Nonlinear Mathematical Models. Mathematical Modeling and Computational Experiment 1 (1993) S. 407–414
- [10] Homma, T., Saltelli, A.: Importance measures in global sensitivity analysis of model output. Reliability Engineering and System Safety 52/1 (1996) S. 1–17
- [11] Jansen, M. J. W.: Analysis of variance designs for model output. Computer Physics Communications 117 (1999) S. 35–43

Prognose der Verfügbarkeit von flexiblen Produktionssystemen in der Industrie 4.0



Forecast of the availability of flexible production systems in industry 4.0

Fei Long, GSaME, Prof. Dr.-Ing. **Bernd Bertsche**,
Institut für Maschinenelemente, Universität Stuttgart;
Prof. Dr.-Ing. **Peter Zeiler**, Hochschule Esslingen,
Fakultät Mechatronik und Elektrotechnik, Göppingen

Kurzfassung

Die Prognose der Verfügbarkeit von flexiblen Produktionssystemen dient dazu, das Risiko von unerwarteten Maschinenausfällen und die daraus resultierenden Verluste zu reduzieren. Die Flexibilität hat weitreichende Auswirkung auf die Gestaltung von Produktionssystemen in der Industrie 4.0 und ist nach wie vor eine Herausforderung in der Modellierung. In dieser Arbeit wird zuerst ein Konzept der flexiblen Produktion vorgestellt und anschließend mit einem High-Level-Petrinetz modelliert. Schließlich werden die Produktivität und die Verfügbarkeit der Produktionssysteme analysiert und diskutiert.

Abstract

Prediction of the availability of flexible production systems serves to reduce the risk of unexpected machine failures and the resulting losses. Flexibility has a significant impact on the design of production systems in the industry 4.0 and is still a challenge in modelling. In this paper, a concept of flexible production is presented and modelled with high-level Petri net. Finally, the productivity and availability of the product systems are analysed and discussed.

1. Einleitung

Die Produktionssysteme haben sich zunächst von der handwerklichen Produktion in die klassische industrielle Produktion am Anfang des zwanzigsten Jahrhunderts entwickelt, danach in die schlanke Produktion in der Mitte des zwanzigsten Jahrhunderts und schließlich in die wandlungsfähige Produktion in den letzten zehn Jahren [1]. Zur Erhöhung der Effektivität bei wachsender Produktvielfalt wurde die schlanke Produktion weit verbreitet. Mit der rasanten Popularisierung des Computers und Internets gewinnt die digitale Produktion [2] immer

mehr an Bedeutung. Vor kurzem wurde ein neues Konzept der Produktionssysteme vorgeschlagen: die Produktion als ein System zu betrachten [3]. Als zukünftige Fabriken werden die wandlungsfähige Produktion, vernetzte Produktion, digitale Produktion, Lernfähige Produktion und Nachhaltigkeit im Lebenszyklus des Systems Produktion in [3] vorgestellt. Diese Konzepte weisen auf die Entwicklungsrichtung der Produktionssysteme in Industrie 4.0 hin.

Mit der Entwicklung der Produktionssysteme wandelte sich die Produktionsstrategie von Massenproduktion über kundenspezifische Massenproduktion zur Massenpersonalisierung. Am Ende des achtzehnten Jahrhunderts führte die erste Massenproduktion durch Maschinen zur ersten industriellen Revolution. Im späten zwanzigsten Jahrhundert hat kundenspezifische Massenproduktion (KMP) als eine wichtige strategische Initiative viel Aufmerksamkeit von der Wissenschaft und Industrie erregt [4, 5]. In den letzten Jahren wird die Personalisierung als einer der Haupttreiber für die nächste Transformation der Weltwirtschaft betrachtet [6, 7].

Die Industrie 4.0 ist durch eine starke Individualisierung der Produkte unter den Bedingungen einer hoch flexiblen Produktion gekennzeichnet [8]. Die Individualisierung ist einer der acht wichtigsten globalen Megatrends, die direkte Auswirkungen auf die Produktion haben [3]. Aufgrund der verkürzten Produktlebenszyklen und der zunehmenden Schwankung von Markt und Nachfrage muss die Produktion eine Vielzahl von Produkten, Designs und Volumenschwankungen anbieten. Die Flexibilität hat weitreichende Auswirkungen auf die Gestaltung von Produktionssystemen in der Industrie 4.0. Deshalb ist es nötig, kundenorientierte flexible Produktion in Industrie 4.0 zu implementieren. In den letzten Jahrzehnten haben zahlreiche Studien die Flexibilität und das Design und die Messung flexibler Produktionssysteme untersucht [9-11]. Eine Definition der flexiblen Produktion wird in [12] diskutiert. In vielen Studien [11-15] wurde die Modellierung flexibler Produktionssysteme behandelt. Die Literatur zeigt, dass die flexible Produktion nach wie vor eine Herausforderung in der Modellierung aufgrund der dynamischen Struktur und komplexe Wechselwirkungen ist. Die Petrinetze wurden in umfangreichen Studien [11, 12, 15-17] zur Untersuchung der flexiblen Produktion verwendet. Sie zeichnen sich durch starke analytische und quantitative Ansätze aus. Daher wird die flexible Produktion in dieser Arbeit mit den Petrinetzen weiter untersucht.

Die Verfügbarkeit ist ein wichtiger Parameter für die objektive Charakterisierung des Leistungsgrads von Produktionssystemen. In den letzten Jahrzehnten haben viele Forscher die Verfügbarkeit von Produktionssystemen auf Basis der gesammelten umfassenden Daten untersucht [18-20]. Heutzutage werden zahlreiche Datenerfassungsgeräte in den Maschinen zur Überwachung ihrer Zustände eingebaut. Darüber hinaus ermöglichen cyber-physische Systeme und das Internet der Dienste und Dinge die Überwachung und Steuerung der ge-

samen Produktionssysteme [20-22]. Damit kann die Analyse der Produktionsprozesse vereinfacht werden. Die Vorhersagemodelle können die Entscheidungen zur Optimierung der Systeme unterstützen. Sie wurden jedoch selten untersucht, weil die Prognose stark von der Expertenerfahrung abhing und die Vorhersage zufälliger Ereignisse und Unfälle schwierig war. Durch die Modellierung der stochastischen Prozesse und die Analyse der Verfügbarkeit von Produktionssystemen können potenzielle Verzögerungen, Störungen und Ausfällen prognostiziert werden. Die Prognose der Verfügbarkeit von Produktionssystemen ermöglicht es, die potenziellen Risiken von unerwarteten Maschinenausfällen und die daraus resultierenden Verluste zu reduzieren und anschließend die Optimierung der Produktionssysteme zu unterstützen. Bosse et al. [23] schlugen ein Modell zur Vorhersage der Verfügbarkeit von IT-Service vor. Das Modell basiert auf einem System mit fester Struktur und ist für die flexible Produktion nur eingeschränkt verwendbar. Deshalb ist es nötig, ein neues Modell zu konzipieren.

In der Literatur wurden verschiedene Methoden [24-27] zur Analyse der Verfügbarkeit vorgestellt. Die Petrinetze verfügen über große Potenziale, die Produktionssysteme in der Industrie 4.0 zu modellieren. Sie wurden als eine Methode vorgeschlagen, um komplexe Systeme zu modellieren [27]. In einer früheren Arbeit [28] wurden die graphischen Modellierungsmethoden nach den speziellen Eigenschaften der Produktionssysteme in der Industrie 4.0 — Vernetzungsfähigkeit, Kommunikation, Intelligenz und Veränderungsfähigkeit — bewertet. Die erweiterten farbigen stochastischen Petrinetze (ECSPNs) [29] wurden als eine der leistungsfähigsten Methoden bewertet und werden daher in dieser Arbeit weiter untersucht.

Ziel dieser Arbeit ist es, die flexible Produktion mit ECSPNs zur Analyse der Produktivität und Verfügbarkeit der Produktionssysteme zu modellieren. In Abschnitt 2 werden die Petrinetze und ECSPNs in Kürze beschrieben. Dann wird ein Konzept der flexiblen Produktion in Abschnitt 3 vorgestellt und anschließend mit ECSPN modelliert. Die Modellierungsergebnisse werden in Abschnitt 4 erläutert. Am Ende der Arbeit wird eine Zusammenfassung gegeben.

2. Modellierungsmethode

Die Petrinetze wurden von Carl Adam Petri zur Modellierung von nebenläufigen diskreten Systemen im Jahre 1962 entwickelt [30]. Sie bestehen aus vier Grundelementen: Stelle, Transition, Kante und Marke. Die Stellen sind als Kreise dargestellt. Sie beschreiben mögliche Zustände eines Systems oder einer Komponente und können eine oder mehrere Marken enthalten. Die spezifischen Werte der Zustände werden durch die Marken beschrieben, die im Allgemeinen als ausgefüllte Kreise dargestellt sind. Die Transitionen stellen die Aktivitäten

dar, die die Werte der Zustände ändern, und können das dynamische Verhalten in einem System modellieren. Sie sind als Quadrate dargestellt. Die Kanten verbinden die Stellen und Transitionen und stellen den Fluss einer Marke durch Pfeile dar. Sie zeigen die logischen Zusammenhänge und die Strukturen der Modelle.

In den letzten Jahrzehnten wurden viele neue Petrinetze entwickelt, die jeweils eigene Eigenschaften haben. Auf Basis der farbigen Petrinetze und der erweiterten stochastischen Petrinetze entwickelte Pozsgai die erweiterten farbigen stochastischen Petrinetze (ECSPNs) [29]. Die ECSPNs können nicht nur die allgemeinen Aspekte der Produktionssysteme wie Struktur und Arbeitsablauf, sondern auch stochastische Prozesse, verschiedene Zustände der Produktionsprozesse und deren Änderungen sowie Eigenschaften wie Warteschlangen und Wächterfunktionen simulieren. Mehr Modellierungsaspekte von ECSPNs werden in [29] vorgestellt.

3. Modellierung einer flexiblen Produktion

Im Folgenden wird ein Konzept der flexiblen Produktion vorgestellt und anschließend mit ECSPNs modelliert. Zum Vergleich wird zuerst die serielle Produktion (siehe Bild 2 links) erklärt. Das folgende Bild 1 zeigt ein konventionelles Modell einer seriellen Produktion (Modell K), die aus vier Prozessen A, B, C und D besteht. Die Maschinen werden durch die Transitionen Tr3, Tr6, Tr9 und Tr12 modelliert. Die Verfügbarkeit der jeweiligen Maschine wird mit den markierten Stellen P31, P61, P91 und P121 modelliert. Weil die Reihenfolge der Prozesse vorher definiert wird, ist die Struktur der Prozesse daher fest. Deshalb ist sie nicht anpassungsfähig. Diverse Varianten eines Produkts und verschiedene Produkte können nicht oder nur mit aufwändigen Anpassungen hergestellt werden. Zusätzlich hat ein Maschinenausfall in der seriellen Produktion eine Kettenreaktion zur Folge. Ein Ausfall der ersten Maschine bewirkt, dass der ganze Produktionsprozess abbricht. Deshalb wird für jede Maschine ein Puffer benötigt, um die Kettenreaktion zu reduzieren. Um die Anpassungsfähigkeit zu ermöglichen, wird eine flexible Produktion benötigt. Das Bild 2 rechts illustriert ein Konzept der flexiblen Produktion.

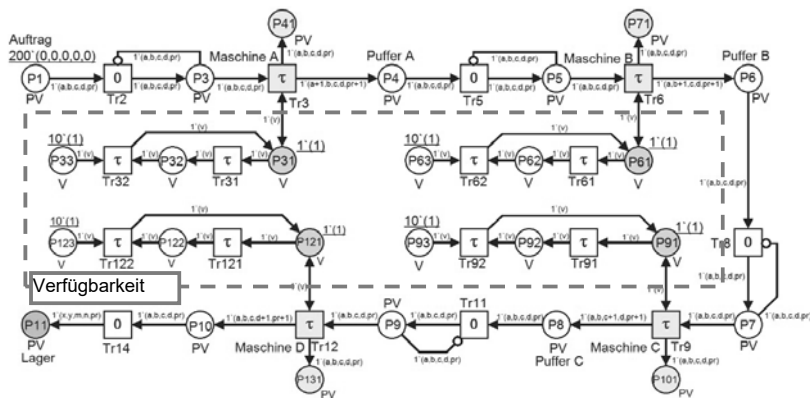


Bild 1: Konventionelles Modell der seriellen Produktion (Modell K)

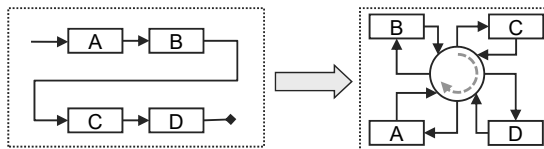


Bild 2: Konzept der flexiblen Produktion

Die Maschinen werden nicht nacheinander, sondern durch ein Verteilzentrum verbunden. Alle Maschinen werden als Individuen betrachtet und sind unabhängig von anderen. Die Reihenfolge der Prozesse soll bei der Produktion keine Rolle spielen. Dadurch können die Prozesse flexibel sein. Die Struktur der Prozesse wird in diesem Konzept abgeschafft. Somit sind die durch Strukturveränderung entstandenen Probleme nicht mehr vorhanden. Das Bild 3 zeigt ein Petrinetz-Modell der flexiblen Produktion mit Berücksichtigung der Warteschlange und Priorität (Modell FP). Die Stelle P1 beschreibt den Auftrag der zu herstellenden Produkte. Die Produkte werden mit den Marken in Form „n'(a,b,c,d,pr)“ dargestellt. Dabei bedeutet „n“ die Anzahl der Produkte. Die Buchstaben a, b, c und d weisen die entsprechenden Prozesse hin. Die Zahl „0“ bedeutet, dass der entsprechende Prozess durchzuführen ist. Die Zahl „1“ deutet hin, dass der Prozess entweder nicht notwendig ist oder schon bearbeitet wurde. Die Marke entspricht RFID auf dem Bauteil mit individuellem Arbeitsplan.

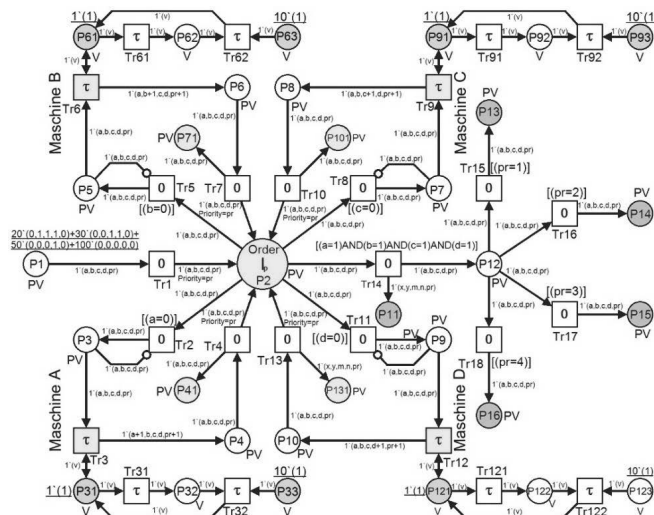


Bild 3: Modell der flexiblen Produktion mit Priorität (Modell FP)

Die Stelle P2 entspricht dem Verteilzentrum in obigem Konzept und ist ein zentraler Puffer für alle Maschinen. Sie kann als der „Supermarkt“ in heutigen Produktionssystemen betrachtet werden. Das Warteschlangenprinzip und die Priorität dienen dazu, die zu bearbeitenden Produkte nach definierter Reihenfolge zu sortieren. Die Wächterfunktionen der Transitionen Tr2, Tr5, Tr8 und Tr11 haben die Funktion, die richtigen zu bearbeitenden Prozesse eines Produkts zu erkennen. Sie können nach Bedarf angepasst werden. Die Normalkante und die Verbotskante zwischen Tr2 und P3 bewirken, dass nur jeweils ein Produkt zu Maschine A versandt wird. Während der Bearbeitung wird kein weiteres Produkt zu dieser Maschine transportiert.

Die Transition Tr3 entspricht dem Bearbeitungsprozess von Maschine A. Sie ist durch eine Lesekante mit der Stelle P31 verbunden, die die Verfügbarkeit der Maschine A zeigt. Die Elemente zwischen P31 und P33 modellieren die Lebensdauer und die Ersatzteile der Maschine. Nach der Bearbeitung wird das Produkt durch den Kantenausdruck „1'(a+1,b,c,pr+1)“ zwischen Tr3 und P4 markiert. Die Priorität wird um eins erhöht. Die Stelle P41 dient dazu, die Produktivität der Maschine A zu bewerten. Die anderen Maschinen B, C und D sind wie die Maschine A gestaltet. Mit der Wächterfunktion in Tr14 „[(a=1)AND(b=1)AND(c=1)AND(d=1)]“ können die Endprodukte aussortiert werden. Die Ele-

mente hinter der Stelle P12 dienen dazu, die Endprodukte nach Produktvarianten zu sortieren.

4. Ergebnis

Das Modell der flexiblen Produktion wird mit einer Software REALIST simuliert. Die Prozessdauer, die Lebensdauer und die Reparaturdauer werden mit zweiparametrischen Weibullverteilungen ($b=1$) beschrieben. Die Prozessdauer (Tr3, Tr6, Tr9, Tr12) und die Reparaturdauer aller Maschinen (Tr32, Tr62, Tr92, Tr122) sind zehn Stunden. Die charakteristische Lebensdauer aller Maschinen (Tr31, Tr61, Tr91, Tr121) ist 1000 Stunden. Das im Bild 3 dargestellte Modell der flexiblen Produktion kann durch Anpassung der Wächterfunktionen nach Tabelle 1 auch für eine definierte serielle Produktion verwendet werden. Zum Vergleich werden die beiden Modelle in Bild 1 und 3 mit verschiedenen Parametern simuliert. In den Modellen K1 und F1 haben alle Maschinen zehn Ersatzteile. Dagegen haben sie in den Modellen K2 und F2 keine Ersatzteile. Dabei bedeutet „K“ konventionelle serielle Produktion, und „F“ serielle Produktion mit flexiblem Modell. Die Priorität wird in beiden Modellen nicht verwendet.

Durch Variation der Ersatzteile können die Produktivität und Verfügbarkeit des ganzen Systems analysiert werden. Alle Modelle werden mit 3000 Stunden und 1000 Replikationen simuliert. Die Ergebnisse sind in Bild 4 dargestellt. Dabei entspricht die mittlere Markenbesetzung dem Mittelwert aller Replikationen. Es gibt zwei hundert Produkte, die nach der Prozessreihenfolge A, B, C und D hergestellt zu werden. Der Vertrag ist als 200' (0,0,0,0,0) dargestellt.

Tabelle 1: Wächterfunktionen der Transition in flexiblen Modellen

Modell	Tr2	Tr5	Tr8	Tr11
Flexible Produktion	$[(a=0)]$	$[(b=0)]$	$[(c=0)]$	$[(d=0)]$
Serielle Produktion	$[(a=0)]$	$[(a=1) \text{ AND } (b=0)]$	$[(a=1) \text{ AND } (b=1) \text{ AND } (c=0)]$	$[(a=1) \text{ AND } (b=1) \text{ AND } (c=1) \text{ AND } (d=0)]$

Das Bild 4.a zeigt, dass die beiden Modelle gleiche Produktivität in beiden Fällen (mit und ohne Ersatzteil) haben. Die Bilder 4.b und 4.c zeigen, dass jede Maschine gleiche Produktivität in beiden Modellen hat. Die minimale und maximale Puffergröße des konventionellen Modells kann man in Bild 4.d erkennen. Die Puffergröße A ist größer als B, und B größer als C. Es ist sehr auffällig, dass die Puffergröße A deutlich zunimmt, wenn es keine Ersatzteile gibt. Der Grund liegt darin, dass der Ausfall einer vorderen Maschine eine Kettenreaktion auf die hintere Maschine hat. Ohne Instandhaltung wird dieser Effekt besonders groß. Zusätzlich besitzt der Auftrag vom Modell K in Stelle P1 einen sehr großen Lagerraum, wie im Bild 4.e

dargestellt. Im Vergleich dazu benötigt das flexible Modell keine zusätzlichen Puffer. Das Bild 4.f zeigt, dass sich die Verfügbarkeiten der Maschinen in beiden Modellen gleich verhalten. Das konventionelle Modell kann nur für eine definierte serielle Produktion verwendet werden. Im Vergleich dazu kann das flexible Modell die Produktion verschiedener Produktvarianten modellieren. Im Folgenden werden die Anwendungen des flexiblen Modells erklärt. Vier Produktvarianten sind in unterschiedlicher Losgröße herzustellen: 20 „a“, 30 „ab“, 50 „abc“ und 100 „abcd“.

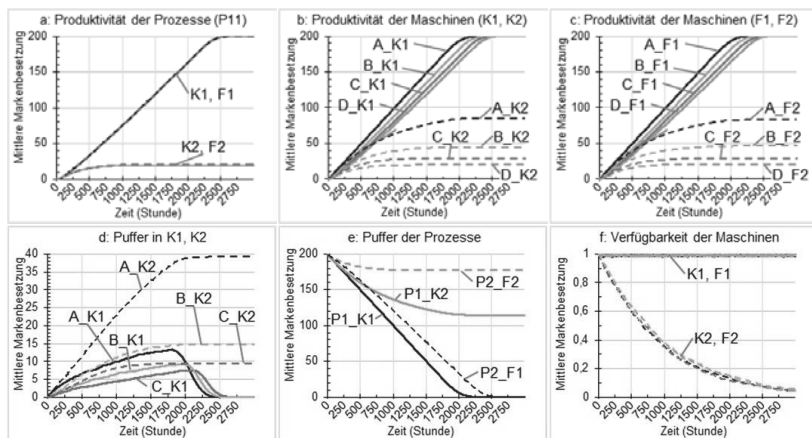


Bild 4: Ergebnisse des konventionellen und flexiblen Modells

Diese Information wird im Vertrag als „20'(0,1,1,1,0)+30'(0,0,1,1,0)+50'(0,0,0,1,0)+100'(0,0,0,0,0)“ dargestellt. Insgesamt werden acht Modelle mit verschiedenen Variablen simuliert. In F1, FS1, FP1 und FSP1 ist die Anzahl der Ersatzteile für alle Maschine zehn, und in F2, FS2, FP2 und FSP2 die Anzahl null. Dabei bedeutet „F“ flexible Produktion, „FS“ serielle Produktion mit flexiblem Modell und „P“ das Modell mit Priorität. In den FS-Modellen werden die Wächterfunktionen nach Tabelle 1 angepasst. Die Produktivität der Maschinen und die Anzahl der hergestellten Produkte werden analysiert und in Bild 5 vorgestellt. Die Produktivität des Systems wird in Stelle P11 bewertet und in Bild 5.a dargestellt. Die Produktivität von F1 ist am Anfang sehr gering. Jedoch steigt sie später sehr schnell und ist am Ende am höchsten. Der S-förmige Verlauf deutet hin, dass die Prozesse zufällig bearbeitet werden. Dagegen werden die Produkte im FS-Modell nach einer Prozessreihenfolge hergestellt. Deshalb hat FS1 eine konstante Produktivität. Die Produkte werden mit Warteschlange

und Priorität sortiert. Ein Teil der Produkte muss auf die Produkte mit höherer Priorität warten. Deshalb haben FP1 und FSP1 konstante und im Vergleich zu F1 und FS1 niedrigere Produktivität.

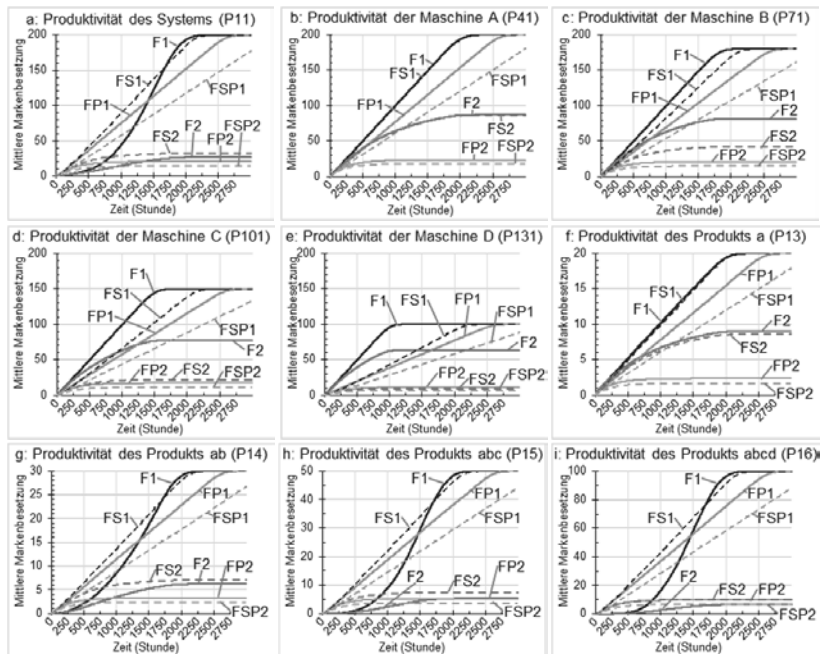


Bild 5: Produktivität der flexiblen Modelle

Wie erwartet haben die Modelle ohne Ersatzteile eine geringe Produktivität. Die Bilder 5.b bis 5.e zeigen die Produktivität der jeweiligen Maschine in allen Modellen. Alle Maschinen haben konstante Produktivität. Die Produktivität ist im Modell F1 am höchsten. Besonders auffällig ist die Produktivität der Maschine D. Maschine D muss darauf warten, bis alle anderen Prozesse der Produkte durchgeführt werden. Eine Ausnahme ist die Produktivität der Maschine A im Modell FS1. Wie im Bild 5.b gezeigt, überlappen sich die Linien F1 und FS1. Der Grund dafür ist, dass alle Produkte mit dem Prozess „a“ beginnen. Betrachtet man die Produktivität von den Modellen ohne Ersatzteil, kann man deutlich erkennen, dass F2 eine viel höhere Produktivität besitzt als die anderen Modelle. Die Bilder 5.f bis 5.i stellen die Produktivität der jeweiligen Produktvariante dar. Die Summe aller Produkte entspricht der Produktivität des

Systems. Ohne Ersatzteil kann das Model FS2 mehr Produkte als andere Modelle herstellen. In diesem Fall hat das flexible Modell keinen Vorteil. Insgesamt kann man erkennen, dass das flexible Modell höchste Produktivität hat. Besonders haben die Maschinen B, C und D in flexiblem Modell viel höhere Produktivität als in den anderen Modellen. Somit können die Maschinen früher ausgeschaltet werden oder mit weiterer Produktion beginnen.

Die Verfügbarkeit eines reparierbaren Systems hängt wesentlich von dessen Zuverlässigkeit und seiner Instandhaltung ab. Durch Zerlegung der Systemstruktur aus einer Kombination von seriellen und parallelen Strukturen ist es möglich, die Verfügbarkeit der Systeme mit fixierter Struktur zu berechnen. Das Verfahren ist jedoch für dynamische Systeme mit veränderter Struktur nicht anwendbar. Das im Bild 3 dargestellte Modell ermöglicht zwar die Modellierung der flexiblen Produktion. Die Maschinen werden jedoch als Individuen betrachtet. Weil die Struktur des Systems abgeschafft wird, ist es nur möglich, die Verfügbarkeit jeder einzelnen Maschine separat zu betrachten. Die Verfügbarkeit des Systems kann man durch Analyse der Produktivität des Systems indirekt ermitteln. In dieser Arbeit wird nur die Anzahl der Ersatzteile als Variable zur Analyse der Verfügbarkeit verwendet. Die Einflüsse der anderen Parameter auf die Verfügbarkeit werden in weiteren Forschungsarbeiten behandelt.

Die Anzahl der Ersatzteile wird nach der Tabelle 2 eingestellt. Im Model F1 haben alle Maschinen ausreichende Ersatzteile. Dagegen haben sie keine Ersatzteile im Model F2. In den Modellen F3, F4 und F5 wird nur die Anzahl der Ersatzteile von Maschine A variiert.

Die Ergebnisse sind in Bild 6 dargestellt. Das Bild 6.a zeigt die Verfügbarkeit der Maschine A. Die Modelle F2 und F3 haben den gleichen Verfügbarkeitsverlauf. Mit zunehmender Anzahl der Ersatzteile steigt die Verfügbarkeit der Maschine A. Die Verfügbarkeit der anderen Maschinen wird in Bild 6.b illustriert. Die Maschinen B, C und D haben in den Modellen F1, F3, F4 und F5 die höchste und in F2 die niedrigste Verfügbarkeit. Die Produktivität des Systems spiegelt die Verfügbarkeit des Systems wieder und wird in Bild 6.c dargestellt.

Tabelle 2: Anzahl der Ersatzteile zur Analyse der Verfügbarkeit

Modell	P33	P63	P93	P123	Modell	P33	P63	P93	P123
F1	10	10	10	10	F4	1	10	10	10
F2	0	0	0	0	F5	2	10	10	10
F3	0	10	10	10					

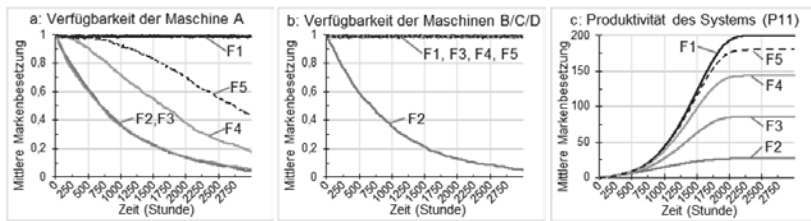


Bild 6: Verfügbarkeit und Produktivität des flexiblen Modells

Die Produktivität erhöht sich mit zunehmender Anzahl der Ersatzteile von Maschine A. Das Model F3 hat eine höhere Produktivität als F2, weil die Verfügbarkeiten der Maschinen B, C und D erhöht werden. Abhängig von der erzielten Produktivität ist es beispielsweise möglich, die Anzahl der notwendigen Ersatzteile zu bestimmen.

5. Zusammenfassung

In dieser Arbeit wird zuerst ein Konzept der flexiblen Produktion vorgestellt und anschließend mit ECSPNs modelliert. Die Produktivität und die Verfügbarkeit der Produktionssysteme werden analysiert. Die Produktivität der jeweiligen Maschine zeigt, dass das konzipierte Modell die flexible Produktion modellieren kann. Die Herstellung verschiedener Produkte mit unterschiedlicher Losgröße kann dadurch analysiert werden. Zur Analyse und Prognose der Verfügbarkeit wird die Anzahl der Ersatzteile als Parameter variiert. Dadurch kann die Produktivität und die Verfügbarkeit der Systeme bestimmt werden. Als nächster Schritt ist die Lebensdauer als Parameter mit unterschiedlichen Verteilungsfunktionen zu untersuchen. Die Flexibilität ist nur eine Basisfunktion. Zur Modellierung fortgeschrittener Eigenschaften wie Intelligenz ist es nötig, andere Methoden, z.B. maschinelles Lernen, weiter zu untersuchen.

Literatur

- [1] Rauch, E.: Konzept eines wandlungsfähigen und modularen Produktionssystems für Franchising-Modelle. Dissertation. Universität Stuttgart, 2004
- [2] Westkämper, E.; Constantinescu, C.; Lentjes, J.: Digitale Produktion. Berlin Heidelberg: Springer, 2013
- [3] Westkämper, E.; Löffler, C.: Strategien der Produktion. Berlin Heidelberg: Springer, 2016
- [4] Pine, B.J.: Mass customization: the new frontier in business competition. Boston: Harvard Business Press, 1993

- [5] Gilmore, J.H.; Pine, B.J.: Markets of one: Creating customer-unique value through mass customization. Boston Mass: Harvard Business Press, 2000
- [6] Tseng, M.M.; Jiao, R.J.; Wang, C.: Design for mass personalization. CIRP Annals-Manufacturing Technology, 2010, 59(1), S. 175-178
- [7] Jiao, J.R.; Simpson, T.W.; Siddique, Z.: Product family design and platform-based product development: a state-of-the-art review. Journal of intelligent Manufacturing, 2007, 18(1), S. 5-29
- [8] Schäfer, S.; Pinnow, C.: Industrie 4.0: Grundlagen und Anwendungen; Branchentreff der Berliner Wissenschaft und Industrie. Hrsg.: DIN Deutsches Institut für Normung e.V. Berlin Wien Zürich: Beuth, 2015
- [9] Sethi, A.K.; Sethi, S.P.: Flexibility in manufacturing: a survey. International journal of flexible manufacturing systems, 1990, 2(4), S. 289-328
- [10] Sarker B.R.; Krishnamurthy S.; Kuthethur S.G.: A survey and critical review of flexibility measures in manufacturing systems. Production Planning & Control, 1994, 5(6), S. 512-523
- [11] Saren S.K.; Tiberiu V.: Review of flexible manufacturing system based on modeling and simulation. Fascicle of Management and Technological Engineering, 2016, 24(14), S. 113-118
- [12] Zhou M.C.; Venkatesh K.: Modeling, simulation, and control of flexible manufacturing systems: a Petri net approach. Vol. 6. World Scientific, 1999
- [13] Anglani A.; Grieco A.; Pacella M. et al.: Object-oriented modeling and simulation of flexible manufacturing systems: a rule-based procedure. Simulation Modelling Practice and Theory, 2002, 10(3), S. 209-234
- [14] Tüysüz F.; Kahraman C.: Modeling a flexible manufacturing cell using stochastic Petri nets with fuzzy parameters. Expert Systems with Applications, 2010, 37(5), S. 3910-3920
- [15] Boualem M.; Cherfaou M.; Bouchentouf A.A.; Aïssani, D.: Modeling, Simulation and Performance Analysis of a Flexible Production System. European Journal of Pure and Applied Mathematics, 2015, 8(1), S. 26-49
- [16] Alanche, P.; Benzakour, K.; Dolle, F. et al.: PSI: A Petri net based simulator for flexible manufacturing systems. In Rosenberg et al. (ed.), Advances in Petri Nets, 1984, S. 1-14. Berlin Heidelberg: Springer, 1985
- [17] Narahari, Y.; Viswanadham, N.: A Petri net approach to the modelling and analysis of flexible manufacturing systems. Annals of operations research, 1985, 3(8), S. 449-472

- [18] Heisel, U.; Hammer, H.; Pritschow, G.: Influences on the availability of flexible manufacturing systems. *CIRP Annals-Manufacturing Technology*, 1992, 41(1), S. 459-462
- [19] Eymard, R.; Mercier, S.: Comparison of numerical methods for the assessment of production availability of a hybrid system. *Reliability Engineering & System Safety*, 2008, 93(1), S. 168-177
- [20] Wang, L.: Machine availability monitoring and machining process planning towards Cloud manufacturing. *CIRP Journal of Manufacturing Science and Technology*, 2013, 6(4), S. 263-273
- [21] Kagermann H.; Wahlster W.; Helbig J.: Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Securing the future of German manufacturing industry. Final report of the Industrie 4.0 Working Group. Frankfurt/Main: Forschungsunion, 2013
- [22] Bauernhansl T.; Ten Hompel M.; Vogel-Heuser B.: *Industrie 4.0 in Produktion, Automatisierung und Logistik*. Wiesbaden: Springer, 2014
- [23] Bosse, S.; Splieth, M.; Turowski, K.: Vorhersagemodell für die Verfügbarkeit von IT-Services aus Anwendungssystemlandschaften. *Wirtschaftsinformatik*, 2013, 935-949
- [24] Lie C.H.; Hwang C.L.; Tillman F.A.: Availability of maintained systems: a state-of-the-art survey. *AIIE Transactions*, 1977, 9(3), S. 247-259
- [25] Malek M.; Hoffmann G.A.; Milanovic N. et al.: *Methoden und Werkzeuge zur Verfügbarkeitsermittlung*. Humboldt-Universität zu Berlin, Bericht für Informatik, 2007
- [26] Bertsche B.: *Reliability in automotive and mechanical engineering*. Berlin, Heidelberg: Springer, 2008
- [27] Milanovic N.: *Models, methods and tools for availability assessment of it-services and business processes*. Technische Universität Berlin, Habilitationsschrift, 2010.
- [28] Long F.; Zeiler P.; Bertsche B.: Potentials of coloured Petri nets for realistic availability modelling of production systems in industry 4.0. In: *ESREL 2015 Conference*, Zürich, Switzerland. London: Taylor & Francis Group, 2015, S. 4455-4463
- [29] Pozsgai P.: *Realitätsnahe Modellierung und Analyse der operativen Zuverlässigkeitskennwerte technischer Systeme*. Dissertation. Universität Stuttgart, 2006
- [30] Petri C.A.: *Kommunikation mit Automaten*. Dissertation. Universität Bonn, 1962

Entwicklung eines Vorgehenskonzeptes zur Analyse von Fehlerfolgeketten in komplexen Systemen



Development of a procedure for the analysis of failure chains in complex systems

O. Bielefeld, M.Sc., H. Dransfeld, B.Sc., Dr.-Ing. N. Schlüter,
Bergische Universität Wuppertal, Wuppertal

Kurzfassung

Fehler, die sich auf die Qualität und Zuverlässigkeit eines Produktes auswirken, frühzeitig in der Produktentwicklung zu vermeiden, ist ein lang gehegtes Ziel und immer noch eine große Herausforderung in der Entwicklung komplexer, mechatronischer Produkte. Insbesondere die zunehmende Verschmelzung der Systemgrenzen erschwert es, Ursache-Wirkungs-Zusammenhänge von Fehlern zu erkennen. So bedarf es eines Vorgehens, welches das Zusammenspiel von mehreren Fehlern, Fehlfunktionen oder fehlerbehafteten Komponenten und deren Zusammenwirken in Fehlerfolgeketten analysieren kann.

Um solche komplexen Zusammenhänge zu untersuchen, ist eine systematische und modellbasierte Analyse unter Nutzung verschiedener, fachspezifischer Methoden erforderlich. Damit lassen sich Auswirkungen und Wechselwirkungen innerhalb und zwischen den Systemelementen anhand einer problemfokussierten Vorgehensweise transparent darstellen. Eine ganzheitliche Abbildung des zu untersuchenden Systems ist zudem notwendig, damit die im System inhärenten, potenziellen Fehlerketten entdeckt werden können.

Doch wie können Fehlerketten in einem System transparent dargestellt und identifiziert werden? Gibt es Möglichkeiten, die ermittelten Fehlerketten so zu modellieren, dass sie mit etablierten Methoden der Qualitäts- und Zuverlässigkeitswissenschaften (wie z.B. FTA, RBD,...) umfassend analysiert werden können? Diese Kernfragen bilden den wesentlichen Bestandteil des vorliegenden Beitrags.

1. Einleitung

Die gestiegene Komplexität von mechatronischen Systemen führt dazu, dass auch die Fehler und deren Folgen komplexer werden [1]. So kann ein Leistungsverlust einzelner Komponenten, für einzelne Bauteilgruppen, noch unkritisch sein, aber in Kombination für das Ge-

samtsystem diverse Funktionsausfälle bedeuten [2]. Eine entsprechende Problemfokussierung, die umfassende Wirkzusammenhänge zu einer konkreten Fehlerbeschreibung ermöglicht, ist hier ein erster Schritt, um trotz der hohen Komplexität von mechatronischen Systemen, effektiv Fehleranalysen durchführen zu können [1]. Es sind zudem bei der Entwicklung eines solchen Ansatzes zur Analyse von Fehlerfolgeketten systematische Vorgehensweisen zu erarbeiten, die die klassische Fehlerdefinition, wie bspw. aus der Qualitätsmanagement-Norm, dass ein Fehler die Nicht-Erfüllung einer Anforderung ist [3], erweitert und den Funktionsabfall einzelner Komponenten und das Scheitern von Prozessen mit betrachtet [2]. Diese umfassendere Fehlerdefinition ist zudem auf die für die Analyse verwendete Modellierung und die damit gekoppelten Methoden zu übertragen. Dementsprechend wird im Folgenden zunächst auf die Herleitung einer umfassenden und der heutigen Komplexität von mechatronischen Systemen gerecht werdenden Fehlerdefinition eingegangen, um anschließend eine darauf zugeschnittene Modellierung für das zu entwickelnde systematische Vorgehen aufzuzeigen.

2. Fehlerdefinition

In den internationalen und nationalen Normen sowie Literatur befindet sich eine Vielzahl an Fehlerdefinitionen, die zum Teil nur schwer miteinander vereinbar sind. Wie bereits zuvor erwähnt, bezieht sich die DIN EN ISO 9001:2015 bei Fehlern auf die Nicht-Erfüllung einer Anforderung [3]. Die IEC 61508 beschreibt einen Fehler als eine ungewöhnliche Bedingung, die möglicherweise dazu führen kann, dass eine Systemeinheit ihre spezifizierte Funktion nicht oder nur eingeschränkt erfüllen kann [4]. Göhner benennt für den Begriff „Fehler“ vier verschiedene Definitionen:

Fehler₁: Abweichung zwischen einem berechneten Wert und dem wahren, spezifizierten oder theoretisch richtigem Wert aufgrund eines Fehlers₂ oder einer Störung.

Fehler₂: Nichterfüllung einer Anforderungsspezifikation, Unkorrektheit.

Fehler₃: Abweichung der tatsächlichen Ausführung von der, für die Erfüllung der Anforderungsspezifikation erforderlichen, konstruktiven und fertigungstechnischen Ausführung des Systems.

Fehler₄: Menschliche Handlungen mit unerwünschtem Ergebnis, ein Irrtum oder Schnitzer. ([5], S. 324 ff.)

Die ISO 26262 benennt ebenfalls verschiedene Fehlerarten, indem sie englische Begriffe differenziert. So steht *fault* für eine abnorme Bedingung, die ein Element scheitern lassen kann, während *error* eine Abweichung zwischen dem berechneten und dem spezifizierten

Wert bezeichnet. Als *failure* im Sinne von Ausfall wird ein Aussetzen der Ausführung einer festgelegten Aufgabe verstanden [6].

Werden diese verschiedenen Fehlerbegriffe in den Kontext von Fehlerfolgeketten gesetzt, deren logischen Zusammenhänge über mehrere Elemente und Systemebenen hinweg verlaufen, so wird deutlich, dass der Systemgedanke, der die Betrachtung von Elementen und deren Relationen bzw. Auswirkungen untereinander beinhaltet, unabdingbar für eine geeignete Fehlerdefinition ist. Ebenfalls ist zu berücksichtigen, dass bereits ein Leistungsabfall einzelner Elemente, die sich auf unteren Systemebenen befinden, auf höheren Systemebenen funktionale Einbußen verursachen können. Dementsprechend wird im weiteren Verlauf unter Fehler, gemäß der ISO 26262, eine abnormale Bedingung, die ein Element scheitern lässt [6], verstanden. Ein Element kann dabei eine Komponente auf den unteren Systemebenen sein oder eine Bauteilgruppe bzw. ein Sub-System auf höheren Systemebenen, wie die Abbildung 1 skizziert.

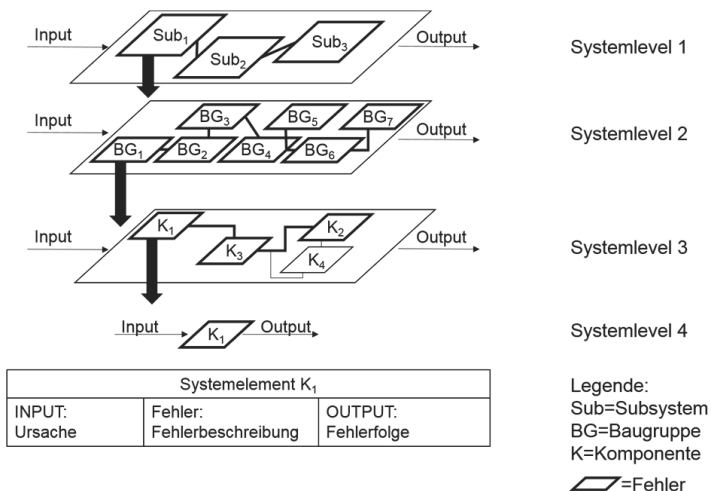


Bild 1: Fehler und Fehlerfolgen einer Komponente innerhalb eines Systems

Der Input für das fehlerhafte Element, in diesem Fall die Komponente 1, ist entsprechend ausschlaggebend bezüglich der Fehlerursache, während das Output des Elements sich auf die Fehlerfolgen auswirkt. Bei Fehlerfolgeketten, die sich aus einer Mehrzahl an Elementen zusammen setzen, können die durch Input und Output attribuierten Relationen, zwischen den Elementen, den logischen Zusammenhang herstellen.

Neben der Elementart „Komponente“ existieren jedoch noch weitere Elementarten – auch Systemsichten genannt – die Bestandteil einer Fehlerfolgekette sein können. Für eine problemfokussierte Fehleranalyse stellt sich somit die Frage, welche Arten von Elementen beschrieben werden müssen. Dies soll im Folgenden geklärt werden.

3. Modellierung von Fehlern

Bezüglich der Modellierung von Fehlern im Rahmen der Problemfokussierung gilt gemäß der Systemtheorie der Grundsatz, nur die minimal nötige Anzahl von Elementarten zu verwenden, um unnötige Komplexität zu vermeiden. Wird die hier verwendete Fehlerdefinition, die wegen des Scheiterns eines Elementes, auf Grund einer abnormalen Bedingung, verwendet, so kann jede Elementart Bestandteil einer Fehlerfolgekette sein. Eine für die Analyse von Fehlerfolgeketten geeignete Modellierung liegt somit vor, wenn das mechatronische System mit einer minimalen Anzahl an Systemsichten abgebildet wird, die ausreicht, um das Problem, mit seinen Ursache-Wirkungszusammenhängen, umfassend zu beschreiben.

Forschungsergebnisse u. a. von [7] und [1] weisen nach, dass die hier gesuchte minimale Anzahl an Sichten vier beträgt und folgende Elementarten enthält: Anforderungen, Funktionen, Komponenten und Prozesse. Diese – DeCoDe genannte – Modellierung verwendet zudem die Dekomposition von Systemen mit Hilfe von Matrizen, um die Komplexität durch die Verwendung von Systemleveln, Hierarchien und Wechselbeziehungen untereinander handhabbar zu machen [9].

Um den Zusammenhang der vier Elementarten in Bezug auf eine Fehlerfolgekette zu verdeutlichen, wird hier an einem Beispiel aus dem EU-Forschungsprojekts K-VEC (siehe Abbildung 2) eine solche Fehlerfolgekette skizziert.

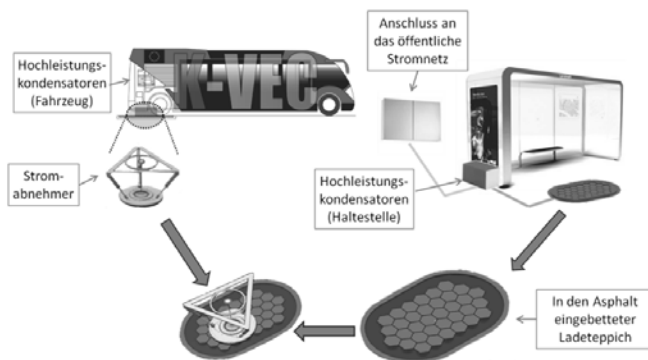


Bild 2: Schematische Darstellung des K-VEC Konzeptes in Anlehnung an [10]

Ziel des Forschungsprojektes K-VEC war die Entwicklung eines Prototyps für Busse des öffentlichen Personennahverkehrs, der einen alternativen Elektroantrieb verwendet, um hohe Energiemengen in kürzester Zeit laden zu können. Hierfür wurde ein konduktives Ladekonzept gewählt. Ein Stromabnehmer, der unter dem Fahrzeugboden angebracht ist, wird an der Haltestelle über einem Ladeteppich positioniert, welcher in den Asphalt eingelassen ist. Nachdem die notwendigen Identifikations- und Autorisierungsprozesse durchlaufen wurden, senkt sich der Stromabnehmer und verbindet sich mit den hexagonalen Platten, die durch den runden Aufbau bedeckt sind. Somit kann der Ladeprozess beginnen [11].

In der Entwicklungsphase wurde ein Wirkungszusammenhang identifiziert, der über mehrere Elemente, Elementarten und Systemlevel hinweg verläuft, wie Abbildung 3 visualisiert ist.

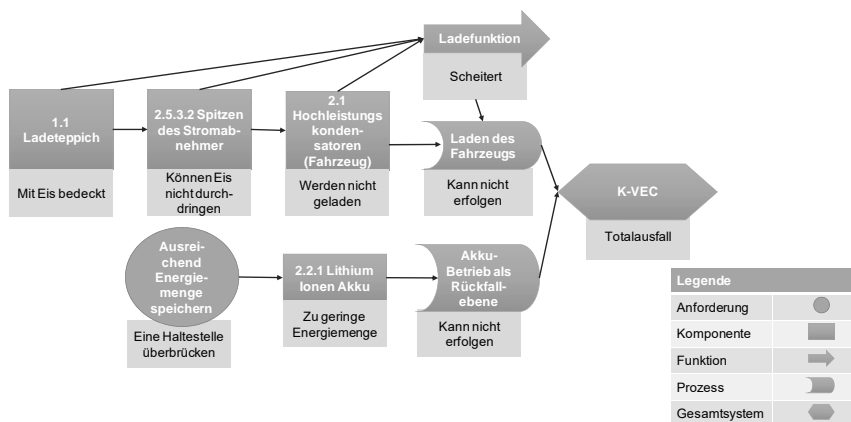


Bild 3: Beispiel einer Fehlerfolgekette beim K-VEC System

Die in Abbildung 3 dargelegte Fehlerfolgekette beschreibt, wie ein Systemausfall des K-VEC Ladekonzeptes im Feld auftreten könnte. Angenommen wird, dass eine Eisschicht auf dem Ladeteppich den Ladeprozess stören würde, indem das Eis die konduktive Verbindung zwischen den Spitzen des Stromabnehmers und den Hexagonalplatten des Ladeteppichs verhindert. Infolgedessen könnte keine Energie von den Hochleistungskondensatoren an der Haltestelle auf die des Fahrzeugs transferiert werden, womit der Prozess „Laden des Fahrzeugs“ auf Grund der abnormalen Bedingung „Eis auf Ladeteppich“ scheitert. Wäre das der Fall, würde dies nicht direkt den sofortigen Systemausfall im Sinne „Weiterfahrt des Busses nicht möglich“ bedeuten. Aufgrund der Energiereserven käme es erst zum Systemausfall, wenn zwei aufeinanderfolgende Haltestellen betroffen wären. Die direkte Folge, bezogen auf

das Gesamtsystem, ist demnach zunächst eine Minderung der Reichweite, welche mit einer Funktionseinschränkung gleichzusetzen ist [11].

Wenn zu der beschriebenen Wirkungskette eine weitere Ursache hinzukommt, wie Abbildung 3 mit der zweiten Wirkungskette aufzeigt, verändert sich aber auch die Funktionsfähigkeit des Gesamtsystems. Die Hochleistungskondensatoren sollen im Projekt K-VEC so ausgelegt werden, dass im Streckenverlauf problemlos jede zweite Haltestelle nicht angefahren werden muss. Für den Fall, dass die zweite Haltestelle in Reihe nicht angefahren werden kann, sind als Rückfallebene zusätzlich Lithium Ionen Akkus vorgesehen, welche die Weiterfahrt sichern. Kommt es nun zusätzlich zu dem Problem mit dem Eis auf einem Ladeteppich zum Ausfall des Akkus, oder aber zum Umstand, dass mehr als vier Haltestellen in Reihe Schnee und Eis auf dem Ladeteppich haben, ist die Weiterfahrt nicht möglich. Das Gesamtsystem scheitert [11].

Das Beispiel zeigt auf, dass jede Elementart von DeCoDe einen Anteil am Scheitern des Gesamtsystems mit sich bringen kann.

Allerdings ist die Modellierung an sich nicht die Lösung des Problems, sondern nur ein Baustein auf dem Weg dahin. Die Lösung bedarf der systematischen Analyse von Fehlerfolgenketten. Die hier beschriebene DeCoDe-Modellierung bietet allerdings die Ausgangsbasis für ein Vorgehenskonzept, das verschiedene Fehleranalysemethoden, je nach vorliegender Problemlösung, nutzt. Dementsprechend wird im Folgenden dargelegt, wie ein solcher modellbasierter Systems Engineering Ansatz unter Verwendung von DeCoDe und einem Vorgehenskonzept für die Fehleranalyse bei mechatronischen Systemen aussehen kann.

4. Modellbasiertes Systems Engineering für die Analyse von Fehlerfolgenketten

Das modellbasierte Systems Engineering (MBSE) wird in diversen Fachdisziplinen verwendet, um komplexe Problemstellungen zu lösen [7]. Dabei wird das Modell als Ausgangsbasis für ein je Problem zu erstellendes Vorgehenskonzept aus verschiedenen Methoden herangezogen. Die Schwachstellen solcher MBSE-Ansätze sind allerdings, dass sie fachspezifisch sind und somit eine interdisziplinäre Ausrichtung, wie es bei mechatronischen Systemen nötig ist, nicht ermöglichen [8]. Zudem erfolgt bei den klassischen MBSE-Ansätzen eine Modellierung lediglich zu Beginn des Vorgehenskonzeptes, um das Problem zu beschreiben. Erkenntnisse von [8] belegen allerdings, dass eine kontinuierliche Interaktion zwischen Modell und Vorgehenskonzept unabdingbar ist, um einen effektiven Problemlösungsprozess zu gestalten.

Um diesen Schwachstellen entgegen zu wirken, postuliert Winzer das Generic Systems Engineering (GSE), welches ein einheitliches, fachdisziplinübergreifend verwendbares Modell nutzt und dieses im Vorgehenskonzept kontinuierlich aktualisieren lässt [7], [8].

Diese Prinzipien sollen auch für die hier betrachtete Analyse von Fehlerfolgeketten berücksichtigt werden. Dies bedeutet, dass das DeCoDe-Modell von einem Vorgehenskonzept umgeben wird, das systematisch Fehlerfolgeketten identifiziert sowie analysiert und die Erkenntnisse der einzelnen Schritte wieder ins Modell überführt, so dass für die nächsten Schritte immer ein aktuelles Systemmodell als Input zur Verfügung steht (siehe Abbildung 4).

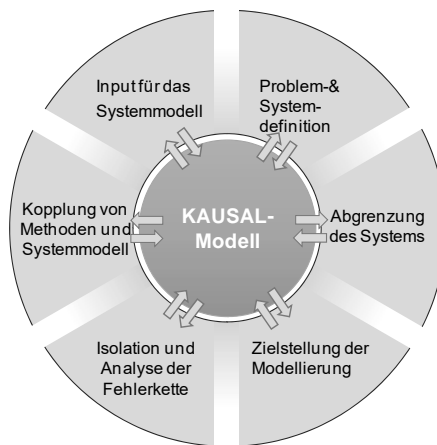


Bild 4: Modell und Vorgehenskonzept für die Analyse von Fehlerfolgeketten

Bezogen auf das oben vorgestellte Beispiel „Eis auf Ladeteppich“ bedeutet dies, dass zunächst auf Basis des identifizierten Problems das hiervon betroffene System definiert, abgegrenzt und modelliert wird, wie es MBSE-Ansätze fordern. Die oben dargestellte Fehlerkette ist im DeCoDe-basierten Systemmodell des Busses bereits latent vorhanden. Um die Auswertung der Kausalketten zu ermöglichen, ist diese Fehlerkette jedoch zunächst zu identifizieren, die beteiligten Relationen zu isolieren und in ein separates Modell (bzw. fokussiertes Modell) zu überführen, das die nötigen Informationen für den Einsatz der Fehleranalysemethoden bildet.

Hierzu kann beispielsweise die Komplexitätsmanagement-Software Looméo® (www.teseon.de) verwendet werden. Die Software ermöglicht nicht nur die Darstellung der DeCoDe-Daten, sondern auch die graphische Visualisierung der DeCoDe-Elementarten und

ihrer Relationen. Durch Fokussierung auf ein Element und seiner Relationen können im Rahmen von Workshops im interdisziplinären Entwicklerteam einzelne Ursache-Wirkungs-Zusammenhänge über verschiedene, mit dem betroffenen Element in Relation stehende andere Elemente, diskutiert werden, wie Abbildung 5 zeigt.

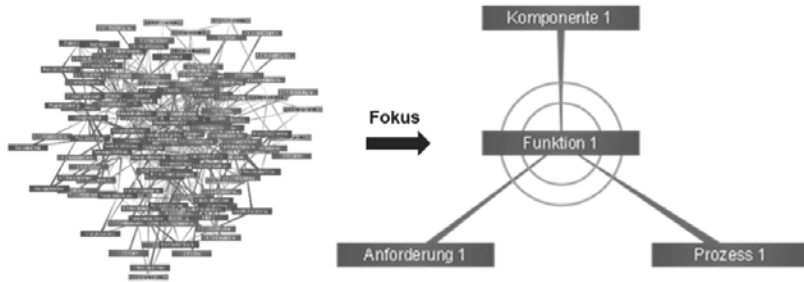


Bild 5: Exemplarische Visualisierung der Identifizierung von Ursache-Wirkungs-Zusammenhängen bei einzelnen Elementen über Fokussierung mittels der Loo-meo®-Funktion „Umgebung“ [1]

Auf Basis der Auswahl eines Systemelements und der Fokussierung mittels der Loo-meo®-Funktion „Umgebung“ wird ersichtlich, welche Elemente vom Scheitern des ausgewählten Elements direkt oder indirekt über entsprechende Relationen betroffen sind. Die so identifizierten Fehlerketten werden separat mit einer eindeutigen ID gespeichert und für die anschließende Fehleranalyse bereitgestellt. Nach Durchführung der je Systemelementart und Problemstellung am besten geeigneten Fehleranalysemethoden werden die Erkenntnisse, die im separaten Modell festgehalten sind, wieder zurück in das Gesamtmodell überführt. Bezüglich des gewählten Beispiels bedeutet dies, dass nach Isolation der Fehlerfolgekette zu bestimmen ist, welche Fehleranalysemethoden geeignet sind, um Entscheidungen bezüglich der Überarbeitung der Konstruktion zu treffen und somit eine anforderungsgerechte Gestaltung von komplexen technischen Systemen zu gewährleisten. So wurde bezüglich der Komponenten ein Reliability-Blockdiagramm verwendet. In Bezug auf die Ladefunktion und deren mögliches Scheitern kam eine Failure-Tree-Analyse zum Einsatz. Die Ergebnisse beider Methoden wurden im KAUSAL-Modell hinterlegt und waren Input für eine anschließend durchgeführte FMECA (siehe Abbildung 6).

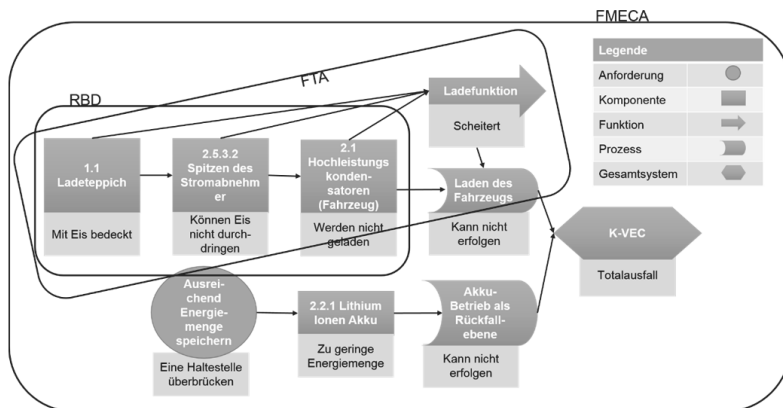


Bild 6: Durchführung verschiedener Fehleranalysemethoden bezüglich des Problems „Eis auf Ladeteppich“

Die Fehlermöglichkeits- und Einflussanalyse (FMECA) ist eine State-of-the-Art-Methode aus dem Bereich der qualitativen Zuverlässigkeitsanalyse [12]. Das Problem der Analyse ist allerdings die limitierte Darstellung von Fehlerketten hinsichtlich ihrer Länge. Innerhalb des Formblattes ist für jeden Fehler, für die potentielle Fehlerursache und -folge, lediglich eine Spalte für die Dokumentation vorgesehen. Die Verkettung von Wirkungseffekten zu Fehlern über mehrere Elemente hinweg wird dadurch vernachlässigt [13]. Doch gerade die Verkettung von unterschiedlichen, fehlerbedingten Ursachen und Wirkungen sind für die Analyse von komplexen Zusammenhängen, wie das oben beschriebene auch Beispiel zeigt, unumgänglich [6] [14] [15] [16]. Dementsprechend wurde vor der Durchführung der FMECA, im hier vorliegenden Beispiel, zunächst eine RBD und eine FTA durchgeführt, die einzelne Wirkungsketten ohne Begrenzung analysierten. Die Ergebnisse für einzelne Pfade der Fehlerfolgeketten wurden im DeCoDe-Modell hinterlegt und konnten dann für die FMECA als ein separates Element verwendet werden, wodurch die Problematik der eingeschränkten Fehlerfolgeketten behoben wurde.

Als Ergebnis der FMECA wurde im Projektteam beschlossen, den Unterboden des Busses zusätzlich mit einer Bürste auszustatten. Des Weiteren wurde die Ladekapazität der Lithium-Ionen-Akkus erhöht, so dass mehr als nur eine Haltestelle überbrückt werden kann.

5. Ausblick

Wie am K-VEC Beispiel aufgezeigt wird, führt die Komplexität heutiger technischer Systeme dazu, dass auch Fehler an Komplexität zunehmen. Solche Fehlerfolgekettten beinhalten mehrere Elementarten und verlaufen über mehrere Systemlevels. Folglich bedarf es einer Weiterentwicklung des Fehlerverständnisses.

Der erste Schritt hierzu ist die Fehlerdefinition. Nur wenn die Fehlerdefinition an sich beinhaltet, dass Fehler mehr als nur ein simples Ursache-Effekt Verhältnis aufweisen, sondern über mehrere Elemente hinweg logisch miteinander in Verbindung stehen, ist eine umfassende Betrachtung und Vermeidung von Fehlern bei komplexen, technischen Systemen möglich. Dementsprechend wurde in dieser Arbeit eine Fehlerdefinition gewählt, die unter Fehler das Scheitern eines Elements unter abnormalen Bedingungen versteht. Dabei wird ein Element nicht allein als eine Komponente verstanden. Das Element kann verschiedener Art sein: Es kann neben einer Komponente auch eine Anforderung, eine Funktion, ein Prozess oder aber eine Kette von Elementen und deren Relationen sein, die ggf. auch über Systemebenen hinweg verlaufen. Dies wurde am Beispiel des K-VEC Bussystems veranschaulicht und mit Hilfe von DeCoDE modelliert.

Um Fehlerfolgekettten umfassend zu identifizieren und zu analysieren wurde zudem ein auf dem GSE-Ansatz basierendes Vorgehenskonzept erarbeitet, das kontinuierlich mit dem Systemmodell interagiert, um für die einzelnen Schritte immer über aktuelle Informationen zu verfügen. Des Weiteren fördert der GSE-basierte Ansatz die interdisziplinäre Entwicklungsarbeit, da dieser fachdisziplinübergreifend ist.

Je nach konkret vorliegender Problemstellung der isolierten Fehlerfolgekette ermöglicht das Vorgehenskonzept, gemäß der GSE-Philosophie, die Auswahl geeigneter Fehleranalysemethoden je in der Fehlerfolgekette vorkommender Elementarten und Relationen. Das Ergebnis ist die Betrachtung und Analyse von Fehlerfolgekettten, die über eine hohe Anzahl an Systemelementen und Elementarten bestehen sowie über mehrere Systemebenen hinweg verlaufen.

Die Grundlage für eine systematische Analyse von komplexen Fehlerfolgekettten ist somit geschaffen. Allerdings ist zum heutigen Zeitpunkt noch unklar, welche Attribute für die Beschreibung der Relationen zwischen den Elementen zwingend erforderlich sind. Während die minimale Anzahl an Elementarten bereits bestimmt wurde, sind weitere Untersuchungen nötig, um die minimale Anzahl an Relationsattributen zu bestimmen, die für die umfassende Beschreibung der Problem- und Zielstellung als auch die Durchführung der verschiedenen Fehleranalysemethoden nötig sind. Hierzu sind im Rahmen des DFG-Forschungsprojektes KAUSAL weitere Analysen durchzuführen. Ziel ist es, eine auf Basis des GSE fußende

KAUSAL-Methodik zu schaffen, die eine mit einer minimalen Anzahl an Elementarten und Relationsattributen gestalteten Modellierung nutzt, um Fehleranalysemethoden verschiedener Fachdisziplinen miteinander zu koppeln. Der hierdurch entstehende Fehlerfolgekettens-Analyseworkflow soll eine qualitative, semi-qualitative und quantitative Analyse komplexer Fehlerfolgekettens in interdisziplinären Teams ermöglichen.

Literatur:

- [1] Bielefeld, O.; Dransfeld, H.; Schlüter, N.; Yazdanmadad, S.; Winzer, P.: Modellbasierte Analyse komplexer Fehlerketten zur Erhöhung der Verlässlichkeit in der Produktentwicklung. In: Christian Tschirner: Tag des Systems Engineering. (Print-on-Demand). Hg. v. Sven-Olaf Schulze und Christian Muggeo. München: Hanser, Carl, 2016, ISBN 978-3-446-45126-1, S. 163–172.
- [2] Bielefeld, O.; Dransfeld, H.; Schlüter, N.; Winzer, P.; Yazdanmadad, S. Development of an innovative Approach for complex, causally determined Failure Chains. In: Proceedings of the 19th QMOD-ICQSS Conference, International Conference on Quality and Service Science, Rome, Italy, ISBN 978-91-7623-06832, 2016, pp. 1413-1431.
- [3] Norm DIN EN ISO 9000, November 2015: Qualitätsmanagementsysteme – Grundlagen und Begriffe.
- [4] IEC 61508 (IEC/TR 61508-0:2005-10): „Funktionale Sicherheit sicherheitsbezogener elektrischer/elektronischer/programmierbarer elektronischer Systeme“.
- [5] R. Lauber, P. Göhner: Prozessautomatisierung 1, 3. Auflage, Springer, Heidelberg, 2013.
- [6] Internationale Norm ISO 26262, 2011: Road vehicles - Functional Safety.
- [7] Winzer, P.: Generic Systems Engineering. Ein methodischer Ansatz zur Komplexitätsbewältigung. 2. Auflage, Springer, Berlin Heidelberg 2016.
- [8] Winzer, P.: Generic Systems Engineering. Ein methodischer Ansatz zur Komplexitätsbewältigung. 1. Auflage, Springer, Berlin Heidelberg 2013.
- [9] Sitte, J., Winzer, P.: Demand Compliant Design. In: IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, Volume 41, No. 3, 2011.
- [10] Willing, M.; Schlüter, N.; Winzer, P.: Use case-based Reliability Assurance. In (Dahlgard-Park, S.M.; Dahlgard, J.J. Hrsg.): Entering the Experience Economy from product quality to experience quality. Proceedings of the 17th QMOD-ICQSS International Conference on Quality and Service Sciences, 2014.

- [11] Willing, M.; Winzer, P.: Fehler vermeiden heißt Fehler verstehen – Anforderungen an eine neue Methodik. In: Bracke, S.; Mamrot, M.; Winzer, P. (Hrsg.), Qualitätsmethoden im Diskurs zwischen Wissenschaft und Praxis, Bericht zur GQW-Jahrestagung 2015 in Wuppertal. Band: 2015,17. Reihe: Berichte zum Qualitätsmanagement. ISBN 978-3-8440-3351-9, Januar 2015, S.303 – 320.
- [12] Bertsche, B.; Göhner, P.; Jensen, U.; Schinköthe, W.; Wunderlich, H.-J.: Zuverlässigkeit mechatronischer Systeme – Grundlagen und Bewertung in frühen Entwicklungsphasen. Springer, Berlin Heidelberg 2009.
- [13] Verband der Automobilindustrie (VDA): Sicherung der Qualität in der Prozesslandschaft. Allgemeines, Risikoanalysen, Methoden, Vorgehensmodelle. Oberursel (Qualitätsmanagement in der Automobilindustrie, 4), 2013.
- [14] Liggesmeyer, P.; Trapp, M.: Safety. Herausforderungen und Lösungsansätze. In (Bauernhansl, T.; ten Hompel, M.; Vogel-Heuser, B. Hrsg.): Industrie 4.0 in der Produktion, Automatisierung und Logistik. Springer Vieweg Verlag, Wiesbaden 2014.
- [15] Stephan, M.: Fehlersensitive Produktgestaltung in integrierten Systemarchitekturen. IPK (Berichte aus dem Produktionstechnischen Zentrum Berlin), Berlin 1996.
- [16] Boorgest, K.: Elektronik in der Fahrzeugtechnik. Vieweg, Wiesbaden 2008.

Zuverlässigkeitsmanagement für den Anlauf von Neuproduktprojekten in der Automobilindustrie

Zuverlässigkeit bestimmen und messen



Dr. Andreas Jacobi, Christoph Jordan, Dr. Ekaterina Kaganova,
Center of Competence Reliability, Daimler AG, Sindelfingen

Kurzfassung

Um die geplante Zuverlässigkeit eines neuen Fahrzeugs und dessen kritische Komponenten zu erreichen, sind durchgängig anwendbare Messgrößen von größter Bedeutung. Dieser Beitrag beschreibt, wie die abgesicherte Zuverlässigkeit von kritischen Komponenten bestimmt wird, wenn die Erprobung unterschiedliche Erprobungsprofile mit abweichenden Raffungsfaktoren und Bauteilmusterständen enthält. Es wird untersucht, welche Änderungen in den Eingangsgrößen sich auf das Ergebnis auswirken. Die Zuverlässigkeit einer Stichprobe von Fahrzeugen wird über ein neues Berechnungsverfahren abgeschätzt, welches durchgängig in allen Erprobungsphasen eingesetzt werden kann.

Abstract

For reaching the planned reliability target of a new vehicle and its critical components universal applicable measureable quantities are of utmost importance. This article describes how the verified reliability of critical components is calculated if testing includes acceleration factors and different part specimen levels. It is investigated which variations of the input parameters affect the results. The vehicle's reliability is assessed by a new calculation method which may be universally applied throughout all testing phases.

1. Einleitung und Zielsetzung

Aus Kundensicht bleibt Zuverlässigkeit das wichtigste Kaufkriterium beim Neuwagenkauf, vor Preis und Wirtschaftlichkeit [1]. Gemessen an Anschaffungspreis und Kosten sind Autos teure Gebrauchsgüter, folglich müssen sie verlässlich funktionieren – unter verschiedenen Einsatzbedingungen ein ganzes Produktleben lang. Die gefühlte Zuverlässigkeit entscheidet mit, ob der Kunde das Auto weiterfährt oder es nach Mängeln und Versagen verkauft. Bewährt es sich aus Kundensicht, stehen die Chancen hoch, der Marke treu zu bleiben und ein weiteres Modell zu kaufen. Auch wenn es überzeugende Innovationen aufweist, erwartet der Kunde, dass es ebenso verlässlich und benutzerfreundlich ist, sowie möglichst weniger kostet. In diesem dynamischen Umfeld benötigen die Hersteller ein Konzept, wie sie die Zuverlässigkeit ihrer

Produkte sicherstellen. Es reicht nicht aus, sich auf Projektmittel, Fähigkeiten und Erfahrungen des Teams zu verlassen und zu hoffen, dass die Zuverlässigkeit erreicht wird.

Ziel des vorliegenden Beitrags ist es, einige wichtige methodische Voraussetzungen für ein wirksames Zuverlässigkeitsmanagement neuer Produkte zu beschreiben. Hierzu gehört die Definition geeigneter Ziele bezüglich Kundennutzung und Zuverlässigkeit sowie die Berechnung der durch das Erprobungsportfolio abgesicherten Zuverlässigkeit für kritische Komponenten. Für Letztere wird im Rahmen einer Sensibilitätsanalyse untersucht, wie sich Änderungen an den Eingangsparametern auf das Gesamtergebnis auswirken. Es wird ein neues Schätzverfahren zur Bestimmung der Beanstandungsquoten von Fahrzeugen und reparierbaren Systemen präsentiert.

2. Zuverlässigkeitsziele für Neuproduktprojekte

Die Qualität des vorherigen Produkts definiert die Meßlatte für die Verbesserungen des Nachfolgers. Es müssen geeignete Qualitäts- und Zuverlässigkeitsziele für das zukünftige Fahrzeug, dessen Systeme und Komponenten festgelegt werden, anhand derer in der Entwicklungs- und Anlaufphase die Zielerreichung quantitativ und qualitativ bewertet werden kann. Die Zuverlässigkeit wird durch die Kundennutzung und die Belastungen beeinflusst. Die Kunden nutzen die Fahrzeuge mehr oder weniger intensiv. Tägliche Laufleistungen und Betriebszeiten, Stadt- oder Autobahnbetrieb führen zu verschiedenen Belastungen und zu Verschleiß, gegen welche das Produkt angemessen ausgelegt werden muss.



Betrachtungsumfang	Reparierbarkeit / Auswirkung	Meßgröße Zuverlässigkeit
Gesamtsystem - Gesamtfahrzeug - Aggregate [Motor/Getriebe] - Antriebsstrang - Interieur - Exterieur - Fahrwerk	 Der Ausfall einer Komponente führt (fast) nie zum Lebensdauerende des Systems	Beanstandungsquote nach n-monatiger Servicezeit, - ausgewertet anhand einer Stichprobe - bezogen auf 100 Fhzge. - dargestellt als: complaints per hundred, abgek. cph - z.B. 12, 24, 36 „Months in Service“ abgek 12 MIS
- Komponente Beispiel: - Abgasurbolader - Lenkrad - Bauteil: - Turbinenwelle - Fehlermechanismus z.B. Bruch Welle	 Der Ausfall eines Bauteils führt (fast) immer zum Lebensdauerende der Komponente	- Ausfallwahrscheinlichkeit $F(100.000km) = x\%$ zum Teil auch bezeichnet mit $F_x = 100.000km$ - Zuverlässigkeit oder Überlebenswahrscheinlichkeit $R(100.000) = 100\% - x$

Bild 1: Messgrößen für die Zuverlässigkeit von Systemen & Komponenten

Im Entwicklungsprozess dienen Erprobungen dem Zweck, die Belastungen des kritischen Kundenbetriebs nachzubilden. Mit einem Erprobungsprogramm weist der OEM nach, dass die

geforderte Lebensdauer und Zuverlässigkeit erreicht werden und das Produkt und dessen Komponenten bzgl. der Belastbarkeit freigegeben werden können.

Die Bewährung von Neuwagen in Kundenhand kann verschiedenartig berechnet werden, siehe Bild 1. Eine große, praktische Bedeutung hat die Beanstandungsquote nach einer gewissen Servicedauer. Für eine Stichprobe, z.B. Produktionsmonat, wird die Summe der Beanstandungen ermittelt und auf den Stichprobenumfang bezogen, siehe [2]. In der betrachteten Stichprobe weisen die Fahrzeuge verschiedene Belastungen (wie Laufleistungen, Betriebszeiten etc.) während der Nutzung auf. Die OEMs nutzen die 12MiS-Beanstandungsquote als zentrale Messgröße zur Bewertung der Fahrzeugqualität im Feld. Durch Produkt- und Prozessverbesserungen in der laufenden Serie sinkt die 12MiS-Quote während der Serienproduktion. Für jede Fahrzeug-Baureihe wird über den gesamten Lifecycle (vom Job Nr.1 bis zum Auslaufen der Produktion) ein Qualitätsziel als 12 MiS-Quote festgelegt und anschließend auf dessen Teilsysteme, wie z.B. Antriebsstrang, Motor, Getriebe, Fahrwerk, Exterieur, u.a., heruntergebrochen. In der Praxis wird der Ziel- und Istwert der 12MiS-Quote in den monatlichen Q-Gremien dem Management berichtet.

Für nichtreparierbare Systeme, wie eine Fahrzeugkomponente, welche bei Ausfall ausgetauscht wird, werden Zuverlässigkeitsziele, in Form von maximal zulässigen Ausfallwahrscheinlichkeiten bezüglich einer Laufleistung, festgelegt. Ein solches Ziel könnte lauten: maximal x % Ausfälle bis 100.000km.

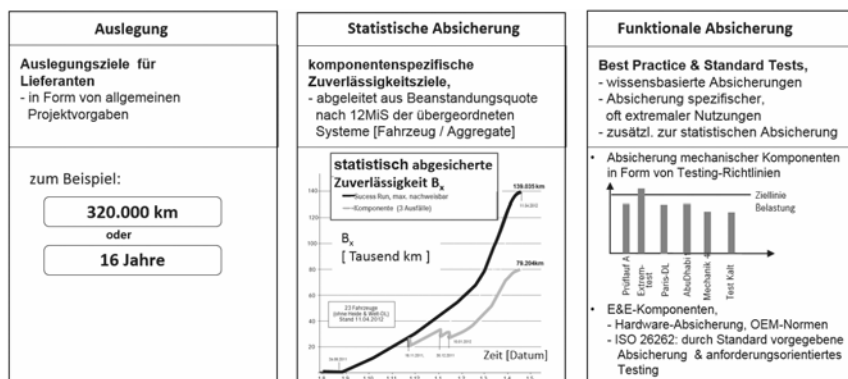


Bild 2: Ziele zur Absicherung von Komponenten

Bei der Einführung von Zuverlässigkeit im Entwicklungsprozess muss dafür gesorgt werden, dass den Entwicklungsingenieuren der Sinn und Zweck der verschiedenen Ziele hinreichend klar sind. Wie in Bild 2 dargestellt, definieren **Auslegungsziele** die beabsichtigte Nutzungsdauer und Laufleistung in Kundenhand. Alle Komponenten eines Fahrzeugs müssen diese

generell gültigen Werten übernehmen; ausgenommen sind anerkannte Verschleißteile und sicherheitskritische Komponenten mit spezifischen Vorgaben. Demgegenüber dienen **Zuverlässigkeitsziele** einer angemessenen statistischen Absicherung. Sie definieren, welche maximale Ausfallwahrscheinlichkeit eine Komponente aufweisen darf, um die Zuverlässigkeit des übergeordneten Systems bzw. die Kundenwahrnehmung nicht zu beeinträchtigen. Die dritte Säule umfasst alle weiteren Erprobungen mit spezifischen, oft extremalen Nutzungen zur **funktionalen Absicherung**.

3. Absicherung von Komponenten mittels Design Verification Plan

Die Bestimmung von Zuverlässigkeitskennzahlen und die Steuerung des Zuverlässigkeitswachstums sind für technische Systeme auf Komponentenebene möglich. In der Automobilindustrie wird der Zuverlässigkeitsprozess in der frühen Produktkonzeptphase gestartet und Komponenten mit hohem Ausfallrisiko werden identifiziert. Für die zuverlässigkeitskritischen Komponenten wird ein weibullverteiltes Ausfallverhalten angenommen, da in der Regel unterschiedliche Ursachen zum Ausfall von Komponenten führen können.

Für den Nachweis der abgesicherten Zuverlässigkeit $R(t_s)$ wird das Prior-Knowledge-Weibayes Verfahren genutzt, bei dem verschiedene Entwicklungs- oder Bauteilmusterstände der Komponente durch den Einsatz von Vorwissensfaktoren berücksichtigt werden können [3].

$$R(t_s) = \exp\left(-\frac{t_s^\beta \chi^2_{2(1+\sum_{i=1}^n \phi_i x_i); P_A}}{2 \sum_{i=1}^n \phi_i (r_i t_i)^\beta}\right) \quad \text{mit} \quad i = 1, \dots, n \quad (1)$$

t_s – geforderte Lebensdauer/Laufleistung;	r_i – Raffungsfaktor;
t_i – Ausfallzeiten oder Laufzeiten;	χ^2 – Chi-Quadrat-Verteilung;
x_i – Anzahl der Ausfälle;	P_A – Aussagewahrscheinlichkeit der
n – Anzahl der Prüfungen unterschiedlicher	Chi-Quadrat-Verteilung;
Teilestände/Stichprobenumfang;	β – Weibull-Formparameter.
ϕ_i – Vorwissensfaktor der Erprobung	

Die Zuverlässigkeit $R(t_s)$ kann damit berechnet werden, wenn Ausfälle vorliegen, d.h. $x > 0$. Wenn keine Ausfälle auftreten, also ein sogenannter „Success Run“ vorliegt, gilt als Freiheitsgrad für die Chi-Quadrat-Verteilung $2(1 + \sum_{i=1}^n \phi_i x_i) = f_i \xrightarrow{x_i=0} 2$.

In den folgenden Abschnitten werden die benötigten Parameter für die Zuverlässigkeitsberechnung bestimmt und dargestellt wie Änderungen der Eingangswerte die Berechnung der Zuverlässigkeit beeinflussen.

Wie in Kapitel 2 erläutert, definieren Autohersteller den Zielwert einer Komponente zumeist über eine Ausfallwahrscheinlichkeit $F(t_s)$ oder direkt als Zuverlässigkeitszielwert $R(t_s)$. Des

weiteren ist ein vorgeplantes Erprobungsportfolio mit einer Anzahl von Prüfungen n und den entsprechenden Laufzeiten t_i festgelegt. Um zu berechnen, welche Zuverlässigkeit sich auf Basis des aktuellen Erprobungsportfolios nachweisen lässt, werden die nach Gleichung (1) stehenden Eingangsparameter β, ϕ_i, r_i ermittelt oder abgeschätzt. Zusätzlich wird durch die Annahme „Success Run“ mit dem geplanten Erprobungsportfolio untersucht, ob das Zuverlässigkeitsziel erreicht wird oder ob zum Erreichen des Zuverlässigkeitsziels zusätzliche Prüfungen für die Komponente notwendig sind. Unter Umständen kann auch eine Überdimensionierung der Erprobung vorliegen und gegebenenfalls der Stichprobenumfang reduziert werden.

Im Folgenden wird erläutert, wie die Parameter β, ϕ_i, r_i und deren Werte in der Regel bestimmt werden. Der Weibull-Formparameter β beschreibt das Ausfallverhalten der Komponente. β wird entweder durch Erfahrungen von gleichartigen Vorgängerkomponenten oder über bereits vorhandene Erprobungsergebnisse, mit hinreichender Anzahl von Ausfällen, ermittelt. Falls keine großen Änderungen an der Komponente, deren Schadensparametern und der Umgebung vorliegen, kann der Weibull-Formparameter β direkt aus der Weibull-Verteilung des Vorgängerbauteils im Feld übernommen werden. Gegebenenfalls wird eine Wertekorrektur oder eine Schätzung durchgeführt.

Die Bestimmung der komponentenspezifischen Raffungsfaktoren r_i in der jeweiligen Erprobung i erfolgt auf Basis von Schadensparametern. Es werden von vorhandenen Erprobungs- und Felddaten für jede Komponente kritische Schadensparameter identifiziert. Anschließend wird zu jeder Erprobungsart das Kundennutzungsverhalten bzgl. dieser Schadensparameter analysiert. Als Ergebnis werden entsprechende Raffungsfaktoren je nach Alterungstyp oder -einfluss mit der Hilfe von verschiedenen physikalischen oder mathematischen Modellen (z.B. dem Arrhenius-Modell oder dem Accelerated-Failure-Time-Modell) bestimmt. Falls die Berechnung über Modelle nicht möglich ist, können Raffungsfaktoren näherungsweise mit dem Verhältnis der Anzahl der kritischen Ereignisse in der Erprobung und der Anzahl der kritischen Ereignisse beim Kunden ermittelt werden. Zudem gibt es auch die Möglichkeit, Raffungsfaktoren durch Experten schätzen zu lassen.

Die Vorwissensfaktoren ϕ_i ermöglichen die gewichtete Berücksichtigung von Erprobungsergebnissen der Komponente aus verschiedenen Entwicklungsständen. Mit den Vorwissensfaktoren wird der Ähnlichkeitsgrad eines Entwicklungsstandes in Bezug auf den Serienstand der Komponente beschrieben. Der Vorwissensfaktor ϕ_i ist im Intervall $[0,1]$ definiert. Entspricht die Erprobungscharge fast dem späteren Serienstand, so liegt der Faktor näher bei 1, bei massiven Abweichungen näher bei 0. Die Vorwissensfaktoren werden von den Bauteilverantwortlichen für die jeweilige Entwicklungsphase beurteilt, wobei Bauteiländerungen, wie

z.B. Geometrie-, Material-, Schnittstellenanpassungen, sowie Änderungen in der Montage oder Fertigung zu berücksichtigen sind.

Da mehrere Parameter für die Zuverlässigkeitsberechnung mit der Hilfe von mathematischen Modellen und vorhandenem Expertenwissen näherungsweise bestimmt werden, ist die Sensibilität der Formel (1) von besonderem Interesse.

Zum besseren Verständnis wird im Folgenden ein Beispiel aus der Praxis dargestellt. Für eine Komponente wurde ein Zuverlässigkeitsziel $R(t_s)$ von 99,7% für $t_s = 30.000\text{km}$ definiert, d.h. die Ausfallwahrscheinlichkeit $F(t_s)$ darf maximal 0,3% betragen. Der Nachweis des Zuverlässigkeitsziels soll mit einer Aussagewahrscheinlichkeit von $P_A = 90\%$ erfolgen. Zur Ermittlung des Zuverlässigkeitsnachweises wurde ein fiktives Erprobungsportfolio für die vier vorhandenen Entwicklungsphasen erstellt, s. Tabelle 1.

Tabelle 1: Erprobungsportfolio

		Anzahl Erprobungen in jeweil. Entwicklungsphase				Summe Erprobungen	Laufleistung [km]
	Hinweis: fiktive Erprobungsbezeichnungen & Werte	Entwicklungs aggregat-A	Entwicklungs aggregat-B	Nullserie	Produktions-test		
Prüfstände	Hochlast Dauerlauf	1	1	2	1	5	120.000
	Niederlast Dauerlauf		1	1	1	3	60.000
	Wechselast Dauerlauf	1		1	1	3	52.000
	Peripherie Dauerlauf		1		1	2	50.000
	Lebensdauererprobung	1	1	1		3	150.000
Fahrzeuge	Autobahn-Dauerlauf		1	2	1	4	100.000
	Stadt-Dauerlauf		2	2	1	5	40.000
	Straßen-Dauerlauf		2	2	1	5	60.000
	Taxi-Dauerlauf		1	2		3	75.000
	Welt-Dauerlauf			1	1	2	52.000
						$\Sigma = 35$	

Mit dem Erprobungsportfolio wird definiert, wie viele Komponentenprüflinge je Entwicklungsphase in bestimmten Dauerlaufprofilen getestet werden.

Die Raffungs- und Vorwissensfaktoren, die in den nachfolgenden Tabellen 2 und 3 dargestellt sind, wurden mit mathematischen Modellen und Experteneinschätzungen bestimmt. Um den Einfluss der Veränderung dieser Faktoren auf den Nachweis der Zuverlässigkeit besser untersuchen zu können, wird eine Abweichung dieser Werte von bis zu 10% angenommen. Damit ergeben sich die Minimal- und Maximalwerte der Vorwissens- und Raffungsfaktoren.

Auf Basis der vorhandenen Daten für die Vorgängerkomponente ergibt sich ein Weibull-Formparameter $\beta = 2$. Analog zur Betrachtung der Raffungs- und Vorwissensfaktoren wird eine Abweichung von 10% angenommen. Dementsprechend variiert β zwischen 1,8 und 2,2.

Tabelle 2: Raffungsfaktoren

	Hinweis: fiktive Erprobungsbezeichnungen & Werte	Raffungs-faktor	min Raffungs-faktor	max Raffungs-faktor
Prüfstände	Hochlast Dauerlauf	2,2	1,98	2,42
	Niederlast Dauerlauf	1,0	0,90	1,10
	Wechsellast Dauerlauf	1,7	1,53	1,87
	Peripherie Dauerlauf	1,5	1,35	1,65
	Lebensdauererprobung	3,0	2,70	3,30
Fahrzeuge	Autobahndauerlauf	1,2	1,08	1,32
	Stadt-Dauerlauf	2,0	1,80	2,20
	Straßen-Dauerlauf	1,2	1,08	1,32
	Taxidauerlauf	2,0	1,80	2,20
	Weltdauerlauf	1,0	0,90	1,10

Tabelle 3: Vorwissensfaktoren

	Entwicklungs-aggregat-A	Entwicklungs-aggregat-B	Nullserie	Produktions-test
Vorwissensfaktor	0,025	0,05	0,075	0,09
min. Vorwissensfaktor	0,225	0,45	0,675	0,81
max. Vorwissensfaktor	0,275	0,55	0,825	0,99

Unter Verwendung der Formel (1) mit dem vorhandenen Erprobungsportfolio, den abgeleiteten Werten der Vorwissens- und Raffungsfaktoren und des Weibull-Formparameters lässt sich in Success Run eine Zuverlässigkeit von $R(t_s = 30.000km) = 99,7\%$ nachweisen, wenn man eine mögliche Abweichung ausschließt. Somit wird das definierte Zuverlässigkeitsziel erreicht.

Da in der Realität, wie oben bereits erläutert, mit der Abweichung der abgeleiteten Werte zu rechnen ist, wird nun eine Sensibilitätsanalyse des Zuverlässigkeitswerts in Abhängigkeit von jeweils geänderten Input-Parametern durchgeführt. Wechselwirkungen werden im ersten Schritt nicht berücksichtigt. Auf der nächsten Seite sind in Tabelle 4 die Werte für die durch das Erprobungsportfolio abgesicherte Zuverlässigkeit unter Veränderung der Einzelparameter dargestellt. Dabei wurde unter Verwendung von (1) jeweils eine Maximierung und eine Minimierung der Zielfunktion vorgenommen, um im Rahmen der Abweichung Maximum und Minimum der erreichten Zuverlässigkeit zu ermitteln. Hierbei ergibt sich jeweils das Maximum oder Minimum des Input-Parameters als Lösung eines entsprechenden Linear Programming Problems (LPP). Wenn beispielsweise der tatsächliche Wert der Vorwissensfaktoren in allen Entwicklungsphasen jeweils 10% unter dem ursprünglich abgeleiteten Wert liegt und gleichzeitig die Raffungsfaktoren und Weibull-Parameter unverändert bleiben, beträgt die erreichte Zuverlässigkeit nur 99,66%. Damit kann das Zuverlässigkeitsziel nicht nachgewiesen werden.

In der Praxis kann es auch zu einer gleichzeitigen Abweichung mehrerer Input-Parameter für die Zuverlässigkeitsberechnung kommen. Zur Ermittlung dieses Einflusses wurden alle drei Input-Parameter gleichzeitig verändert und jeweils wieder eine Maximierung und eine Minimierung des Funktionswertes als LPP vorgenommen.

$$\min_{r, \phi, \beta} R(t_s) = \exp \left(- \frac{t_s^\beta \chi^2_{2(1 + \sum_{i=1}^n \phi_i x_i); P_A}}{2 \sum_{i=1}^n \phi_i (r_i t_i)^\beta} \right), \quad (2)$$

$$\left\{ \begin{array}{l} \phi_{i, \min} < \phi_i < \phi_{i, \max} \\ r_{i, \min} < r_i < r_{i, \max} \\ 1,8 < \beta < 2,2 \end{array} \right\}$$

In der nachfolgenden Tabelle 5 sind die LPP-Lösungen als zwei Szenarien dargestellt, welche bei gleichzeitigen Änderungen von mehreren Eingangsgrößen zu den maximalen Abweichungen vom ursprünglichen Zuverlässigkeitsnachweis führen. Damit repräsentieren sie die „worst-case“ Szenarien.

Tabelle 5: Nachgewiesene Zuverlässigkeit

	ursprüngliche Werte	LPP-Lösungen für veränderbare Vorwissensfaktoren		LPP-Lösungen für veränderbare Raffungsfaktoren		LPP-Lösungen für veränderbare Beta	
		min. Vorwissensfaktor	max. Vorwissensfaktor	min. Raffungsfaktor	max. Raffungsfaktor	min. Beta	max. Beta
nachgewiesene Zuverlässigkeit	99,70%	99,66%	99,72%	99,63%	99,75%	99,53%	99,81%
nachgewiesene Ausfallwahrscheinlichkeit	0,30%	0,34%	0,28%	0,37%	0,25%	0,47%	0,19%

Tabelle 4: Nachgewiesene Zuverlässigkeit

	ursprüngliche Werte	MIN LPP-Lösung: min. Vorwissensfaktor, min. Raff.faktor, min. β	MAX LPP-Lösung: max Vorwissensfaktor, max Rafffaktor, max. β
nachgewiesene Zuverlässigkeit	99,70%	99,38%	99,86%
nachgewiesene Ausfallwahrscheinlichkeit	0,30%	0,62%	0,14%

Schliesst man Wechselwirkungen aus, kann man direkt den Einfluss veränderter Input-Parameter auf die Zuverlässigkeit erkennen. So wurde deutlich, dass der Einfluss der Raffungsfaktoren größer als der Einfluss der Vorwissensfaktoren ist. Allerdings ist der Einfluss des Weibull-Formparameters auf die nachgewiesene Zuverlässigkeit am stärksten. Bei einer Verringerung des β -Werts um 10%, sinkt der Zuverlässigkeitswert um 0,17%. Dieser Effekt verstärkt sich bei einer gleichzeitigen Änderung von mehreren Eingangsgrößen: Bei der MIN LPP-Lösung von (2), sinkt der Zuverlässigkeitswert gar um 0,32%. Zusammenfassend gilt,

dass die Berechnung der Zuverlässigkeit über Formel (1) sehr sensibel auf Veränderungen der Input-Parameter reagiert und folglich bei diesen eine hohe Datenqualität benötigt wird.

4. Zielerreichung auf Systemebene

Neben der detaillierten Komponentenbetrachtung ist es gleichzeitig auch wichtig, den Blick auf das Gesamtsystem und die Zielerreichung bezüglich der 12MiS-Beanstandungsquote zu richten. Kurz nach dem Projektbeginn definierte Key Performance Indicators (KPIs) beschreiben den planmäßigen Fortschritt in den jeweiligen Entwicklungsphasen hinsichtlich der Zuverlässigkeit. Der stufenweise gestaffelte Zielflug stellt sicher, dass bereits in frühen Erprobungsphasen das Zuverlässigkeitswachstum gemessen wird und das Management, wenn gefordert, rechtzeitig Maßnahmen ergreifen kann. Bei Mercedes-Benz Cars wird diese Methodik als Predicted Reliability Approach (PRA) zusammengefasst und sorgt für eine Zuverlässigkeitssteuerung und –objektivierung im Entwicklungsprozess.

Grundlage zur Messung des Zuverlässigkeitsfortschritts bilden die im Kapitel 2 beschriebenen 12MiS-Werte. In der Praxis ergeben sich bei der näherungsweisen Berechnung der 12MiS-Werte einige Herausforderungen, für die geeignete Lösungen benötigt werden:

- **Betriebszeit & Laufleistung:** Dauerlaffahrzeuge haben das Ziel, möglichst viel Laufleistung in möglichst kurzer Zeit zu sammeln. Zur näherungsweisen Berechnung der Beanstandungsquote müssen die Laufleistungen der Erprobungsfahrzeuge in kundenäquivalente Laufleistungen transformiert werden.
- **Erprobungsprofile & Raffung:** Die Erprobung besteht aus unterschiedlichsten Profilen und daraus resultieren verschiedenartige Belastungen für die Komponenten. Infolgedessen ist es nicht möglich, einen Raffungsfaktor auf Fahrzeugebene in die Berechnung einfließen zu lassen. Für die Berechnung des Zuverlässigkeitswachstums werden extreme Dauerlaufarten ausgeschlossen. Den Fokus bilden sogenannte kundennahe Dauerläufe, welche näherungsweise eine kundenäquivalente Belastung abbilden können.
- **Dokumentation von Beanstandungen:** Die Erprobungsbereiche sind darauf fokussiert, möglichst streng und umfassend alle während der Erprobung entdeckten Mängel und Beanstandungen zu dokumentieren, damit erforderliche Optimierungen schnell eingeleitet werden können. Im Vergleich zu den Feldbeanstandungen, die durch den Kunden bemerkt und bei Werkstattbesuch dokumentiert und abgearbeitet werden, besteht eine Diskrepanz. Daher müssen alle während der Erprobung entdeckten Beanstandungen aus Kundensicht analysiert und klassifiziert werden. Daraus können sich Fehler ergeben, die aus

Entwicklungssicht durchaus wichtig sind, aber für eine Prognose der späteren Beanstandungsquote nicht herangezogen werden.

- **Formparameter der Weibullverteilung:** Das charakteristische Ausfallverhalten einer Komponente wird durch den Weibull-Formparameter β beschrieben. Durch die Überlappung unterschiedlicher Ausfallmechanismen bei einem System wird in der Literatur immer wieder auf näherungsweise zufallsbasiertes Ausfallverhalten für komplexe Systeme hingewiesen. Diese Annahme wird durch Simulationsmodelle bestärkt.

Die klassische Berechnung der Beanstandungsquote lautet

$$cph_{12MiS} = \frac{\sum_{i=1}^n x_i}{\sum_{i=1}^n LL_i^b} \cdot \emptyset LL_{12MiS} \cdot 100 \quad (3)$$

mit den Größen

cph_{12MiS} - Schätzung der 12 MiS Beanstandungsquote „complaints per hundred“,

x_i - die in der Erprobung i aufgetretenen Ausfälle,

LL_i - die zurückgelegten Laufleistung der Erprobung i und

$\emptyset LL_{12MiS}$ - durchschnittliche Jahreslaufleistung im Feld.

Mit dem zentralen Konfidenzintervall für den Schätzer des Erwartungswerts $\hat{\lambda}$ der Ausfälle aus der Poisson-Verteilung ergibt sich [4]:

$$\sum x_i = \lambda \mapsto \lambda = \frac{\chi^2_{2(x+1); P_A}}{2} \quad (4).$$

Damit wird aus (3) mit (4)

$$cph_{12MiS} = \frac{\chi^2_{2(x+1); P_A}}{2 \cdot \sum_{i=1}^n LL_i^b} \cdot \emptyset LL_{12MiS}^\beta \cdot 100 \quad (5)$$

wobei

β - Formparameter der Weibullverteilung

x - Anzahl der Beanstandungen in der Erprobung

P_A - Aussagewahrscheinlichkeit

LL_i - Laufleistung der Dauerlauffahrzeuge.

Zur Bewertung des Projektfortschritts der Erprobung benötigen die Projektbeteiligten eine klare Vorgabe welche Steigerung der Zuverlässigkeit für das gesamte System, z.B. ein Fahrzeug oder einen Motor, erzielt wurde. Diese Steigerung wird durch den Rückgang der Beanstandungsquote über der Zeit dokumentiert. Aus vorangehenden Projekten kann diese Kurve berechnet werden. Nach Ermittlung des Fehlerabstellgradientens wird dieser auf das neue 12MiS Ziel umgelegt. Dadurch ergeben sich Zielfluggeschwindigkeiten in den jeweiligen

Entwicklungschargen gegen welche die Erprobung des Neuproduktprojektes gemessen werden kann.

Die linke Darstellung im nachfolgenden Bild 3 zeigt den Verlauf der Schätzung der 12 MiS-Quote mit den gepunkteten dargestellten 5% bzw. 95%-Grenzen des Konfidenzbereiches in den Entwicklungsphasen eines Neuproduktprojekts nach Gleichung 5. In der rechten Darstellung sind die Zielkorridore für den „Anflug“ der Beanstandungsquote eingezeichnet, welche aus den Beanstandungsdaten von zwei Vorgängerprojekten gewonnen wurden.

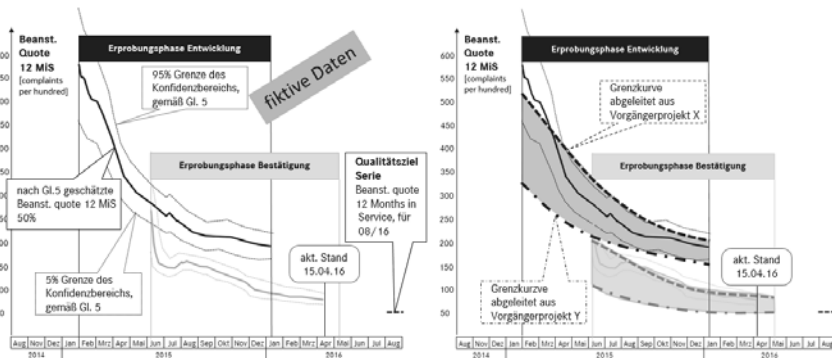


Bild 3: Verlauf der 12MiS-Beanstandungsquote in den Entwicklungsphasen eines Projekts

5. Fazit und Ausblick

Wie der Artikel zeigt, ist ein gesteuerter Zielflug für ein komplexes System nicht das Resultat einer einzigen Methode, sondern die konsequente Anwendung und Verbindung von zuverlässigkeitstechnischen Werkzeugen für die jeweils gegebenen Herausforderungen im Entwicklungsprozess.

Dabei ist es von elementarer Bedeutung zum Projektbeginn, die zu erreichenden Ziele und Messgrößen genau zu definieren. Nur so lässt sich zu einem späteren Zeitpunkt auch eine eindeutige Bewertung vornehmen, ob man sich im Plan befindet oder nicht. Ebenfalls wird empfohlen, das Gesamtsystem in reparierbare und nicht reparierbare Umfänge einzuteilen, da für die jeweiligen Betrachtungen unterschiedliche Berechnungsmethoden gewählt werden müssen. Wie die durchgeführte Sensibilitätsanalyse für die Komponentenabsicherung zeigt, beeinflussen die Eingangsparameter und deren Prämissen in starkem Maße die abgesicherte Zuverlässigkeit. Deswegen sollten die beteiligten Experten nach dem Anlauf die verwendeten Eingangsparameter und deren Prämissen anhand der vorliegenden Istdaten kritisch überprüfen. In gleicher Weise sollten nach dem Anlauf die Eingangsparameter für die

vorgestellte Schätzung der 12 MiS-Beanstandungsquote anhand der vorliegenden Ergebnisse überprüft werden. Auf diese Weise kann für nachfolgende Projekte eine höhere Ergebnisqualität der Absicherungsrechnung sichergestellt werden.

Die hier gezeigten Methoden bilden die Basis für eine ganzheitliche Zuverlässigkeitsabsicherung im Entwicklungsprozess, sowohl auf Komponenten- als auch Systemebene. Bei Mercedes-Benz Cars kommen diese Vorgehensweisen im Zuge des Zuverlässigkeitsprozesses bei allen Neuproduktprojekten flächendeckend zur Anwendung.

6. Literaturverzeichnis

- [1] DAT-Report 2016. Jährlicher Bericht der Deutschen Automobil-Treuhand GmbH.
- [2] Jordan, C.; Jacobi, A.; Köttermann, T.: Zuverlässigkeitsprognose (12 Months in Service) im Anlauf eines Neuproduktprojektes. Schneller die Qualitätssituation bewerten. 27.Fachtagung Technische Zuverlässigkeit 2015, Leonberg bei Stuttgart, 20.& 21.05.2015. Verein deutscher Ingenieure, VDI-Berichte, Bd. 2307, Düsseldorf: VDI-Verlag. 2015. S. 257-268.
- [3] Breidbach, J.; Müller, S.; Haller, S.; Jacobi, A.; Grabert, M.; Kempe, S.: Durchtrainiert, Ganzheitliche Zuverlässigkeitsabsicherung. Qualität & Zuverlässigkeit, München: Carl Hanser Verlag, QZ Jahrgang 60 (2015), 6, S.32-35.
- [4] Sachs, L: Angewandte Statistik. Berlin: Springer Verlag, 2006

Nachweis der funktionalen Sicherheit für eine Abschalt- und Erdungsautomatik (AEA) mittels Bahnspezifischer Normen EN 50126, -28, -29.



Dipl.-Ing. **O. El Adlouni**, RAMS- Systemdesign,
Rail Power Systems GmbH, Offenbach;
Doz. Dr. rer. nat. habil. **H. Schäbe**, Principal Assessor RAMS,
TÜV Rheinland Inter Traffic GmbH, Köln

Kurzfassung

Der Beitrag untersucht die Anwendung der CENELEC-Normen EN 50126 / EN 50128 / EN 50129 für den Nachweis der funktionalen Sicherheit. Dabei geht primär darum, die Kundenanforderungen DB Fernverkehr zu analysieren und nachweislich unter Berücksichtigung der gesetzlichen und normativen Vorgaben umzusetzen.

Nachdem alle Sicherheitsanforderungen der neuartigen Abschalt- und Erdungsautomatik (AEA) bekannt wurden, wurde ein Sicherheitsplan in Abstimmung mit dem Auftraggeber erstellt. In diesem Sicherheitsplan wurde eine Auswahl von zeitlich festgelegten Maßnahmen und Aktivitäten (Verantwortlichkeit, Dokumentstruktur, Verfahren) dem SIL-Level entsprechend, festgelegt. Die Schlussfolgerung ist, dass die hier gewählten Bahnnormen und deren Methodik erfolgreich zur funktionalen Sicherheit von Bahnstromanlagen angewendet wurden.

1 Einführung

Im Bereich der Bahntechnik existieren drei wesentliche Normen zur Zuverlässigkeits- und Sicherheitstechnik. Die EN 50126 für den RAMS -Prozess in Eisenbahnsystemen, die EN 50128 für Software sowie die EN 50129 für „Telekommunikationstechnik, Signaltechnik und Datenverarbeitungssysteme – Sicherheitsrelevante elektronische Systeme für Signaltechnik“, wie es in der EN 50129 heißt.

Auch wenn nun die funktionale Sicherheit von Systemen außerhalb der Signaltechnik betrachtet wird, muss zwangsläufig eine der oben genannten Normen ausgewählt werden. Diesem Problem widmen sich die Autoren in der vorliegenden Arbeit.

Der Ausgangspunkt war eine neuartige Abschalt- und Erdungsautomatik (AEA) für eine Wagenhalle, die ein elektronisches System in Form einer speicherprogrammierbaren Steuerung benutzt. Gegenüber bisher gebauten Anlagen basierend auf Relais-technik, wird für den Neubau der Halle Berlin Rummelsburg eine softwarebasierte Lösung entwickelt. Dabei wird ein besonderes Augenmerk auf den Umgang mit den neuen Kundenanforderungen unter Berücksichtigung der umfangreichen Nahtstellen gelegt.

Die durch Rail Power Systems (RPS) zu errichtende AEA-Anlage hat als oberstes Ziel, das sichere Arbeiten auf und an Fahrzeugen bei ausgeschalteter und geerdeter Oberleitung zu gewährleisten. Für die Konstruktion der AEA-Anlage sind strenge Sicherheitsanforderungen zum Schutz von Instandhaltungspersonal einzuhalten. Ein Teil dieser Sicherheit ist die funktionale Sicherheit, welche durch die korrekte Funktion der Steuerungseinrichtung bestimmt wird. Damit kommt ganz zwangsläufig die Frage der funktionalen Sicherheit ins Spiel und demnach ist eine geeignete Norm auszuwählen.

In dieser Arbeit wird dargestellt, wie die Auswahl einer geeigneten Norm in Abstimmung mit dem Auftraggeber erfolgte und welche Konsequenzen dies beim Hersteller hatte, insbesondere wie der RAMS –Prozess beim Hersteller gestaltet wurde. Für die Begutachtung wurde der TÜV-Rheinland seitens des Auftraggebers beauftragt.

2 Anforderungen an die Abschalt- und Erdungsautomatik

Bei der hier vorgestellten technischen Lösung handelt es sich um eine Abschalt- und Erdungsautomatik (AEA) für die Große Wagenhalle Rummelsburg. Diese Wagenhalle wird für Instandhaltungsaufgaben an Lokomotiven und IEC-Triebzügen genutzt und verfügt über eine Oberleitungsdeckenstromschiene (OSS). Wenn nun am Dach der Fahrzeuge Arbeiten durchgeführt werden, so muss die Dacharbeitsbühne betreten werden. Dadurch können Personen in die Nähe der Oberleitung gelangen und es besteht die Gefahr eines Stromschlages. Daher muss die Oberleitung sicher abgeschaltet und geerdet werden. Hierfür wird die AEA benötigt, um die 5 Sicherheitsregeln gemäß DIN VDE 0105 einzuhalten.

Anforderungen an Erdungsanlagen haben eine typische Struktur und bestimmte typische Elemente. Dafür wurde eine Anforderungsliste erstellt, deren Struktur folgende Elemente beinhaltet:

- Eindeutige Identifizierungsnummer,
- Festlegung der AEA-Teilsysteme und der Fremdsysteme,
- Art und Beschreibung der Anforderung,
- Gliederung der Anforderungen in z.B. übergeordnete, funktional, betrieblich, Umweltbedingungen, RAMS, Qualität,
- Herkunft (Kunde, Gesetze und Normen, Behörde Auflage, interne),

- Status (Verständlichkeit, Vollständigkeit, Messbarkeit und Widerspruchsfreiheit)
- Verantwortlichkeit (Projektleiter, Produktmanager, Designer, Gutachter, RAMS, Tester),
- Methodik der Nachweise (Berechnung, Simulation, Messung, Zertifikate),
- Kommentare (Begründungen, Abweichungen in der Methodik,...) und
- Historie für die Verwaltung.

Nachfolgend sind die wesentlichen Kundenanforderungen mit den jeweiligen Sicherheitsaspekten zusammengefasst:

1. Die Personensicherheit wird durch ein Schlüsselsystem gewährleistet. Als Mitarbeiterschlüssel werden Medien verwendet, die entweder elektrisch oder per Funksignal gelesen werden. Eine fehlerhafte Lesung der Schlüsselnummer bzw. Übertragung zur Steuerung ist soweit wie möglich auszuschließen. Die Verfahren, die für den Betrieb eines solchen Schlüsselsystems angewendet werden, müssen zwischen Betreiber, unabhängiger Prüfstelle und Lieferanten abgestimmt werden.
2. Mehrere Arbeitsstände müssen zusammengefasst werden können, wenn ein Fahrzeug sich gleichzeitig in zwei oder mehr Arbeitsständen befindet. Die Detektion von Fahrzeugen, die die Trennstellen überbrücken (Überstandbereich), erfolgt mittels Lichtschranken.
3. Signale, welche für die Personensicherheit von Bedeutung sind, müssen fehlersicher übertragen werden.
4. Als Kommunikationsprotokoll zwischen der AEA und der Mehrspannungsversorgung (MSV) soll Profisafe in Verbindung mit Profinet zum Einsatz kommen.
5. Schnittstellen werden über Profibus mit einem DP/DP-Koppler zum Betriebsführungsrechner (BFR) und mittels PN/PN-Koppler zur Dacharbeitsbühne (DAB) ausgeführt.
6. Die funktionale Sicherheit der AEA wird gemäß DIN EN ISO 13849 festgelegt und verifiziert. Als notwendiger Performance Level (PL) wurde PL d ermittelt.
7. Die funktionale Sicherheit muss durch eine unabhängige Prüfstelle beurteilt werden.

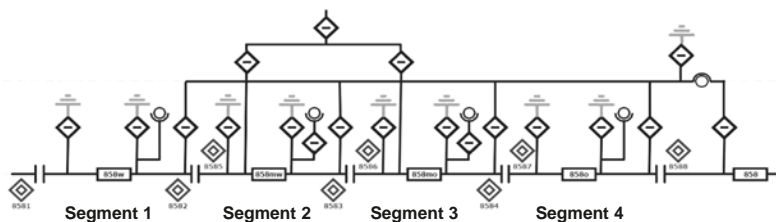
Man sieht, dass in den Kundenanforderungen die EN ISO 13849 genannt wurde. Bei näherer Analyse stellt sich jedoch heraus, dass diese Norm nicht auf Bahnanlagen und Teile davon anwendbar ist.

Diese Anforderungen wurden bei RPS daher nicht buchstabengetreu umgesetzt, sondern entsprechend weiter entwickelt.

Wenn bisher bei RPS derartige Technik im Wesentlichen in Hardware entwickelt wurde, so hat man sich entschlossen, eine Software Lösung vorzusehen. Hierfür wurde eine bereits zertifizierte speicherprogrammierbare Steuerung verwendet, sodass man bei der Hardware auf eine bereits zertifizierte Komponente zurückgreifen konnte und zudem bei der Softwareentwicklung auch zertifizierte Tools nutzen konnte.

3 AEA-Funktionen als Teil des Bahnsystems

Die Abschalt- und Erdungsanlage stellt einen Bestandteil des Bahnsystems dar. In der großen Wagenhalle befinden sich Deckenstromschienen. Die Deckenstromschiene ist pro Gleis in jeweils vier Oberleitungs-Segmente aufgeteilt (Abbildung 1). Diese Segmente lassen sich einzeln mit 15 kV/16,7 Hz beschalten bzw. ausschalten und erden. Es lassen sich auch mehrere benachbarte Segmente desselben Gleises, zu einem Arbeitsbereich zusammenführen.



Legende:

Trennstelle	EI 6 Signal	Erdungsschalter
Speiseschalter	Spannungswandler	

Abbildung 1: Oberleitungssegmente mit den primärtechnischen Komponenten (Quelle RPS)

Die Hauptfunktionen der AEA sind neben der automatischen Durchführung der 5 Sicherheitsregeln sowie die Enterdung der Oberleitung auch z.B. die Ansteuerung der EI-6 Signale (Halt für Fahrzeuge mit gehobenen Stromabnehmer) und DAB-Türen. Die fünf Sicherheitsregeln und ihre Umsetzung werden nachfolgend nochmals genannt.

Tabelle 1: Die Fünf Sicherheitsregeln nach DIN VDE 0105

Sicherheitsregel	Umsetzung durch AEA
1 – Freischalten	Öffnen der Fahrleitungstrenner
2 – Gegen Wiedereinschalten sichern	Verriegeln der Fahrleitungstrenner, Sperren des Betätigungsmechanismus, Unwirksammachen der Hilfsenergie
3 – Spannungsfreiheit feststellen	Messung mittels Spannungswandler
4 – Erden und Kurzschließen	Einlegen der Erdungsschalter (2stufig), sichere Rückmeldung über Endkontakt am Schaltmesser
5 – Benachbarte, unter Spannung stehende Teile abdecken oder abschränken	Fahrleitungsspannung steht außerhalb des Segmentes hinter Fahrleitungstrenner an. Verbindungstüren sind geschlossen und überwacht. EI6-Signale werden angesteuert.

Das unten stehende Bild zeigt den Aufbau der Steuerung innerhalb eines Gleises. Die fehlersichere CPU (S7 -1516F) befindet sich im Zentralschrank und ist angebunden mit den vier Bedienstellen über einen LWL-Ring mit Profisafe Protokoll.

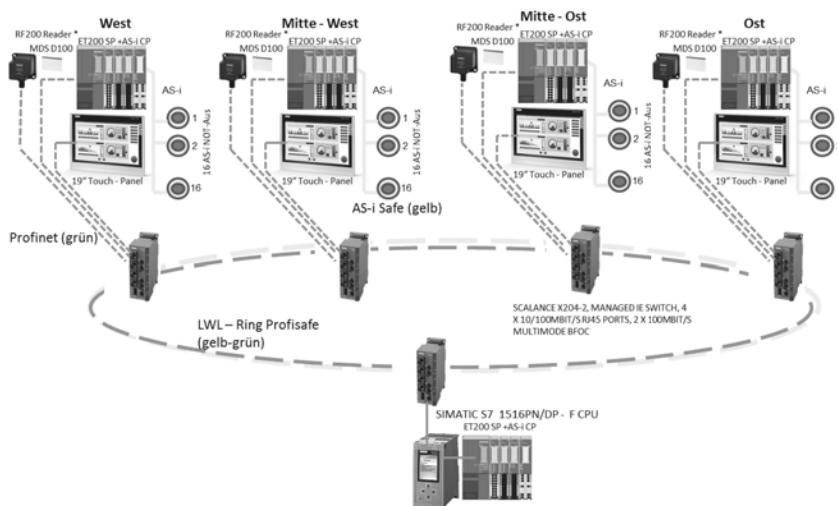


Abbildung 2: Sekundärtechnische Steuerkomponenten (Quelle RPS)

In der Abbildung 3 wird der Bedienprozess der AEA im Zusammenhang mit der Dacharbeitsbühne (DAB) dargestellt. Nach erfolgreicher Anmeldung an der Bedienstelle kann der Mitarbeiter die Zugangstür mit seinem Transponder passieren und somit gelangt er zur DAB.

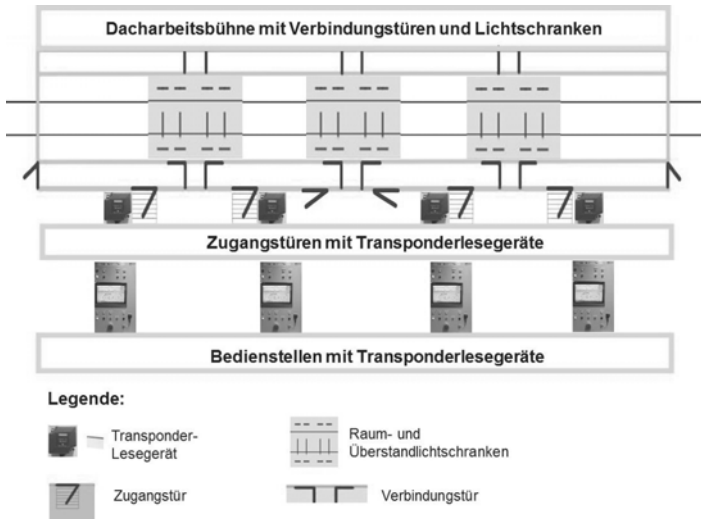


Abbildung 3: Bedienprozess der AEA (Quelle RPS)

Ein wesentlicher Aspekt für die Entwicklung der AEA, ist es die bereits identifizierten Nahtstellen zu realisieren. Dabei sind die frühzeitige Abstimmung und die verbindliche Festlegung von Zuständigkeiten, Leistungen und Terminen vor allem zwischen den Fremdsystemen (Betriebsführungsrechner BFR, Dacharbeitsbühne DAB, Zentralschaltstelle Zes) von Relevanz. Die folgende Abbildung 4 zeigt die projektspezifischen übergeordneten, internen und externen (Fremdsysteme) Nahtstellen. Für die Prozesssimulation wurde eine Demonstrationsanlage eingerichtet, um bestimmte Softwarefunktionen zusammen mit den Nahtstellen zu testen.

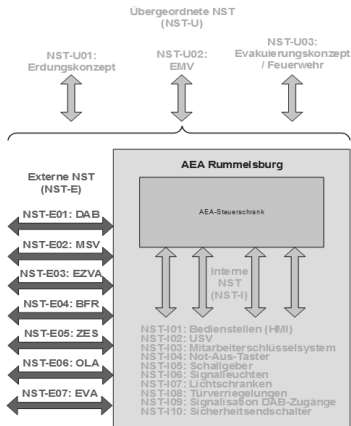


Abbildung 4: Nahtstellen der AEA (Quelle RPS)

Es wurde ein Gefahrenprotokoll aufgestellt. Hierin wurden die folgenden Sicherheitsfunktionen definiert, welche bei Anforderung die Anlage in den sicheren Zustand (Notfallerrdung) überführen:

- Notfallerrdung, ausgelöst durch Betätigung eines Not-Halt-Tasters
- Notfallerrdung, ausgelöst durch unerlaubtes Öffnen einer DAB-Zugangstür
- Notfallerrdung, ausgelöst durch unerlaubtes Öffnen einer DAB-Verbindungstür
- Notfallerrdung, ausgelöst durch Freiraumlichtschranke
- Notfallerrdung bei niedriger USV-Kapazität.

Bevor die Dacharbeitsbühne zugänglich gemacht wird, muss die Oberleitung freigeschaltet und geerdet sein und gegen Wiedereinschalten gesichert sein. Letzteres bedeutet, dass die Spannung nur wieder zugeschaltet werden darf, wenn folgende Bedingungen erfüllt sind:

- Keine Person ist mehr angemeldet.
- Kein MSV-Betrieb.
- Die Verbindungstüren sind ordnungsgemäß geschlossen.
- Der Kran befindet sich in Parkposition.
- Es liegt keine Störung im Betrieb vor.
- PIN ist eingegeben.

Diese Funktionen und Bedingungen sind nun in einer Einheit zu implementieren, die diese Logik umsetzt und implementiert. Damit liegen Anforderungen an die funktionale Sicherheit vor.

4 Begründung für die Auswahl der Normen

Es war zunächst wichtig, die relevanten Normen zu berücksichtigen. Da die große Wagenhalle Teil des Bahnsystems ist, sind damit die ISO 13849 als auch die IEC 61508 nicht anwendbar, da

diese die Bahntechnik explizit oder implizit ausschließen. Aus diesem Gründen wurde zunächst die allgemeine RAMS-Norm für das Gebiet der Bahntechnik, das ist die EN 50126 gewählt. Um nun die Aspekte der funktionalen Sicherheit möglichst gut abzudecken, wurden hierfür die EN 50129 sowie die EN 50128 ausgewählt. Streng genommen sind diese Normen nur für Signaltechnik, Telekommunikation und Datenverarbeitung anwendbar. Man kann nun den relevanten informationstechnischen Teil der Abschalt- und Erdungsautomatik als Datenverarbeitungsanlage auffassen und damit die EN 50128 / 50129 hinzuziehen. Dies wurde gemeinsam mit dem Auftraggeber so entschieden. Es wurden dabei die jeweils neuesten Versionen dieser Normen benutzt, selbst wenn die Verwendung der älteren Version der EN 50128 von 2003 noch zulässig gewesen wäre.

Die IEC 61508 ist die Grundnorm der funktionalen Sicherheit. Diese ist jedoch nicht anwendbar, wenn es branchenspezifische Normen gibt. A priori ist die IEC 61508 damit eigentlich nicht auf Bahnsysteme anwendbar, hierfür gibt es das Normenwerk EN 50126, EN 50128, EN 50129. Insbesondere die EN 50128 und EN 50129 sind streng genommen nur anzuwenden auf

- Telekommunikationstechnik, Signaltechnik und Datenverarbeitungssysteme (EN 50129) bzw.
- Telekommunikationstechnik, Signaltechnik und Datenverarbeitungssysteme – Software für Eisenbahnsteuerungs- und Überwachungssysteme (EN 50128)

Die Steuerungssysteme einer Abschalt- und Erdungsanlage können im weitesten Sinne als Eisenbahnsteuerungs- und Überwachungssystem aufgefasst werden. Dies wurde so getan und damit die Anwendung der EN 50129 und EN 50128 motiviert.

5 Normative Grundlagen zur Sicherheitsbetrachtung auf der Basis der EN 50128 und EN 50129

Auf Grundlage der EN 50129 wird das Sanduhr-Modell angewendet. Nach diesem Modell werden die Sicherheitsaktivitäten zwischen Betreiber / Auftraggeber (betriebliche Parameter wie z.B. Schalthäufigkeit, Signalisationsart) einerseits und Hersteller (System-Parameter wie z.B. Architektur, Art der Diagnose, Kommunikation) andererseits aufgeteilt. Dabei wurde durch RPS in Abstimmung mit dem Auftraggeber

- die Identifikation und Analyse der Gefährdungen sowie Weiterverfolgung (Vorläufige Gefährdungsanalyse / Hazard-Log),
- die Auswahl eines Kriteriums der Risikoakzeptanz sowie
- die Risikoanalyse (Festlegung der SIL-Einstufung) durchgeführt.

Zu Beginn des Lebenszyklus wurde damit durch RPS eine Risikoanalyse (PHA) durchgeführt. Die PHA wurde von Anfang an im Format eines Hazard Logs gemäß EN 50126 Kapitel 6.3.3.3 erstellt und dabei wurden alle Erfahrungen der vorhergehenden Projekte berücksichtigt.

Für die Risikoanalyse wurde der Risikograph nach EN 61508-5 verwendet. Dies wurde motiviert durch die Parallelen zur Anlagentechnik, jedoch auch dadurch, dass die im TSIP / SIRF ermittelten SILs bzw. SAS sich nur auf Fahrzeuge beziehen. Im Ergebnis wurde SIL 2 ermittelt. Das nachfolgende Bild zeigt den Risikographen sowie die Wahl der Parameter.

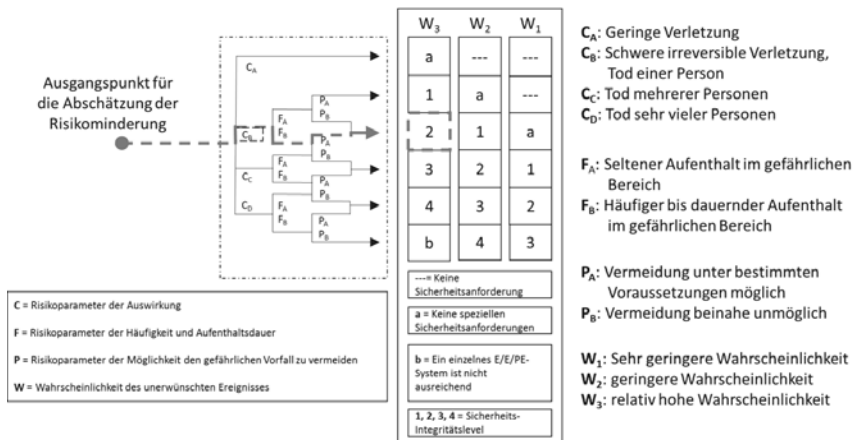


Abbildung 5: Herleitung der SIL-Einstufung anhand des Risikographen nach [4]

Die Wahl war dabei wie folgt angeregt: Ein Unfall wird jeweils eine einzelne Person betreffen, die dabei zu Tode kommen kann (CB), bedingt durch die hohe Spannung von 15 kV. Es hält sich häufig bis dauernd jemand auf der Dacharbeitsbühne auf (FB). Die Vermeidung des Unglücksfalles ist möglich, indem man Sicherheitsabstand zur Oberleitung hält (PA). Zudem gibt es kaum eine Möglichkeit, dass der Unfall durch andere Faktoren nicht eintritt (W3).

Es gibt eine Zuordnung des SIL (Sicherheitsanforderungsstufe) zu THR-Werte, d.h. den maximal zulässigen Raten gefährlicher Ausfälle der Sicherheitsfunktion.

Tabelle 2: Zuordnung von SIL zu THR nach [3]

Sicherheitsanforderungsstufe (SIL)	Tolerierbare gefährdungsrate (THR) pro Stunde und pro Funktion
4	$10^{-9} \leq \text{THR} < 10^{-8}$
3	$10^{-8} \leq \text{THR} < 10^{-7}$
2	$10^{-7} \leq \text{THR} < 10^{-6}$
1	$10^{-6} \leq \text{THR} < 10^{-5}$

Im nächsten Schritt erfolgte die Erstellung eines Sicherheitsplans nach 50126 unter Berücksichtigung der Vorgaben aus EN 50128 und EN 50129. Dies gestaltete sich zunächst etwas schwierig, was dadurch bedingt war, dass es sich um die überhaupt erste Anwendung dieser Normen auf eine AEA-Anlage handelte. In der zweiten Version entstand jedoch ein Plan, mit dem man nicht nur guten Gewissens an die Arbeit gehen konnte, sondern der auch die notwendigen planerischen Grundlagen für das Projekt legte.

Der erste rein technische Ansatzpunkt war das Erstellen einer Anforderungsliste. Diese Liste stellte das Systemanforderungsdokument als auch die Sicherheitsanforderungsspezifikation dar. Alle Anforderungen des Kunden als auch zusätzliche Anforderungen von RPS wurden hier aufgenommen. Die Sicherheitsanforderungen wurden in der Anforderungsliste separat gekennzeichnet. Diese Liste gewährleistete gleichzeitig die Nachverfolgbarkeit der Anforderungen – es wurde jeweils angegeben, wo und wie die Anforderungen im System umgesetzt wurden.

Insgesamt wurde die folgende Dokumentenstruktur für den SIL-2 Nachweis erstellt.

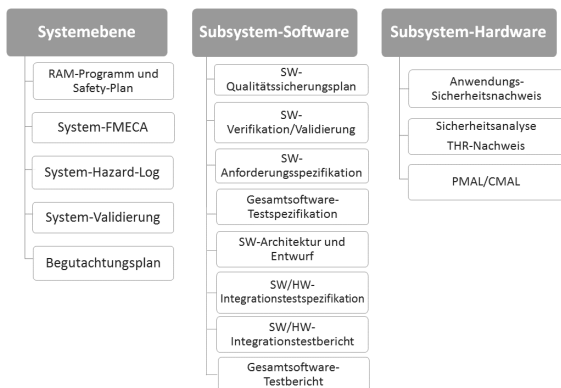


Abbildung 6: Dokumentstruktur für den SIL2- Nachweis (Quelle RPS)

Einige durch die Normen geforderte Dokumente sind im konkreten Anwendungsfall nicht relevant oder durch die SPS (TIA-Portal V13) abgedeckt. Dies wurde begründet, wie z.B. für den SW-Wartungsplan, SW-Konfigurationsmanagementplan und SW-Quellcodedokumentation. Ebenso sind die Verifikationsberichte des Validierers auch Bestandteil der Dokumentation. Das Zusammenfassen von Dokumenten wurde ebenfalls aufgezeigt und mit dem Gutachter abgestimmt.

6 Technische Lösung

Die technische Lösung von RPS basiert auf den folgenden wichtigen Punkten.

1. Für die Realisierung der Personensicherheit und Türfreigabe werden Transpondern an Stelle eines Schlüsselsystems verwendet. Auf Grund der Flexibilität einer derartigen Lösung ist man nicht gezwungen an derselben Stelle auszuchecken, an der man eingeecheckt hat. Zudem kann man die Transponder an Personen koppeln.
2. Die Daten, die auf die Transponder geschrieben werden, müssen nach EN 50159 abgesichert werden. Dabei wird folgender Übertragungsweg betrachtet: Zentrale, Anmeldestelle, Transponder, abmeldestelle, Zentrale. Durch die zusätzliche Eingabe der PIN am Bedientableau ist eine zweistufige Authentifizierung realisiert worden und so ist ein sicheres An- und Abmeldeverfahren gestaltet worden.
3. Schließlich wurden verschiedene SPS-Hersteller aufgefordert, Angebote zu unterbreiten. Die Wahl fiel letztendlich auf ein Modell, das gemäß SIL 3 nach IEC 61508 zertifiziert war. Damit war zum einen eine SPS gewählt, die höheren Sicherheitsanforderungen genügte, zum anderen musste begründet werden, weshalb sich diese SPS in einem Umfeld der EN 50128/ EN 50129 einsetzen ließ. Dies wurde im spezifischen Sicherheitsnachweis begründet, wobei sowohl die Parallele zur Anlagentechnik, als auch die Position der SPS in einer Halle entfernt vom Gleis eine Rolle spielten. Dadurch konnten eine Reihe verschärfter Umweltauflagen, die in der Bahntechnik unter bestimmten Bedingungen auftreten, entfallen.

Das nachfolgende Bild zeigt den prinzipiellen Aufbau des Systems. Dabei beachtet man die typische Aufteilung in Sensor - Verarbeitung – Aktor.



Abbildung 7: Allgemeine Darstellung der Komponenten der gesamten Steuerungskette (Quelle RPS)

Durch die Verwendung einer zertifizierten speicherprogrammierbaren Steuerung ergaben sich Vorteile. Eine Reihe von Elementen des Softwareprozesses wurden durch ein zertifiziertes Tool, das zusammen mit der SPS geliefert wurde, abgedeckt. Dies vereinfachte den SW-Entwicklungsprozess. Dabei ging es insbesondere um Techniken und Maßnahmen, die durch die SPS abgedeckt wurden wie z.B. typisierte Programmiersprache, Trennung zwischen fehlersicherer Anwendung und Standardanwendung, Fehlerkennung und Diagnose, defensive Programmierung usw. Trotzdem wurde die Softwareentwicklung jedoch gemäß EN 50128 durchgeführt.

7 Sicherheitsnachweis, Verifikation und Validierung

Nach EN 50129 besteht eine Sicherheitsanforderung immer aus einer funktionalen Sicherheitsanforderung (Überführung der Anlage in einen sicheren Zustand) sowie einer qualitativen SIL- Stufung und einer quantitativen Ausfallrate (THR).

Um diese zu gewährleisten, sind folgende Maßnahmen seitens der Norm gefordert:

- Maßnahmen zum Management und zur Beurteilung der funktionalen Sicherheit
- Maßnahmen gegen systematische Fehler
- Maßnahmen gegen zufällige Fehler.

Die Verifizierung wurde einerseits durch ein Review der relevanten Dokumente abgedeckt, andererseits durch Nachvollziehbarkeit. Hierfür wurde die Anforderungsliste benutzt um nachzuweisen, wie die einzelnen Anforderungen umgesetzt wurden und wie diese getestet oder auf andere Art und Weise nachgewiesen wurde.

Auf Grund der relativ einfachen Struktur des Systems ließ sich dies mit vertretbarem Aufwand abwickeln. Hierbei war vorteilhaft, dass die gesamte Logik in der SPS implementiert war.

Die Validierung erfolgte durch die Abnahme durch einen EBA-anerkannten Prüfer, der ohnehin für Starkstromanlagen erforderlich war und umfangreiche Tests durchführte. Damit entstand gewissermaßen als „Abfallprodukt“ der Validierungstest der funktionalen Sicherheit durch eine unabhängige Person.

Umwelttests mussten nur in geringem Maße nachgewiesen werden, da die Eignung der Schaltkomponenten ohnehin für den geeigneten Zweck nachgewiesen waren und die SPS die Umweltbedingungen ohnehin schon erfüllte. Dabei wirkte sich günstig aus, dass die Anlage

- a) in einer Halle aufgebaut wurde und
- b) sich keine relevanten Steuerungselemente im 3m Bereich des Gleises befanden.

Die Maßnahmen gegen systematische Fehler konnten gemäß den Tabellen des Anhangs A der Normen EN 50129 bzw. EN 50128 nachgewiesen werden. Es wurden eine Reihe von wichtigen Sicherheitsprinzipien in der Schaltung angewendet:

- Auswahl von Fail-Safe Komponenten
- Ruhestromprinzip
- Redundanz der Stromversorgung für die SPS
- Redundante Abfrage der Stellungsmeldung für den Erdungsschalter
- Sicherheitsrelais, Sicherheitsendschalter
- Zweihandbedienung sicherheitskritischer Befehle
- Gegenseitige Überwachung der Signale (Fremdsysteme) anhand von Lebenszeichen
- Überdimensionierung der SPS: SIL 3 statt SIL 2
- Trennung der Module, welche sicherheitsrelevante Daten bearbeiten von denen keine Sicherheitsrelevanz haben (F und S Teile von CPU)
- Zweipolige Ansteuerung (Erhöhung der Meldesicherheit)
- Sichere Datenübertragung (Profisafe Protokoll)

Daneben wurden Prinzipien für die elektrische Sicherheit angewendet, die indirekt Einfluss auf die Funktionale Sicherheit haben:

- Überspannungsschutz vom Trafo, Motorantrieb des Erdungsschalters,...
- Isolationswächter (Überwachung des Isolationswiderstandes im Niederspannungsnetz)
- Leitungsschutzschalter für EI 6, Mastschalter, ZSS (Schutz Leitungen vor Beschädigungen durch Erwärmung infolge zu hohen Stroms)
- Trenntrafo mit verstärkter Isolierung
- Überdimensionierung der USV (Autonomiezeit pro Gleis mindestens 60 min, Überwachung der Akkumulator-Kapazität bzw. Tiefentladung)
- Schutzbeschaltung mit RC-Glied,
- Überwachung der AC- und DC- Ströme mit Hilfe eines digitales Relais.

Der Nachweis, dass hinreichende Maßnahmen gegen zufällige Fehler implementiert wurden, erfolgte mit Zuverlässigkeits-Blockdiagrammen (RBD). Grundlage für die Erstellung der Blockdiagramme waren die durch RPS erstellten FMECAs. Diese wurde erweitert um die normativen Anforderungen zu erfüllen um folgende Angaben:

- Unterteilung der Ausfallraten in sicher und gefährbringend, erkannte oder unerkannte Ausfälle
- Festlegung des Prüfintervalls T1.

Das Ziel der Sicherheitsanalyse ist, für die identifizierten sicherheitsrelevanten Funktionen die Hardware Komponenten zu kategorisieren und die zugehörigen THR bzw. λ_d (Rate gefährlicher Ausfälle) zu ermitteln.

Für die Berechnung der Blockdiagramme wurden folgende Annahmen getroffen:

- alle nachgelagerten Fehler in den Fremdsystemen, welche die funktionale Sicherheit der Anlage beeinflussen, sind nicht Bestandteil dieser Analyse,
- Anwendung von Herstellerdaten für die Ausfallraten der Komponenten,
- die definierten MTTR-Werten (Erwartungswert der Reparaturzeiten) basieren aus Erfahrungen, die RPS Mitarbeitern aus ähnlichen Anlagen gesammelt haben. In diesen Werten sind keine logistischen Zeitelemente enthalten,
- das Proof Test Intervall wurde auf ein Jahr (8760 h) festgelegt,
- als Wert für die Schalthäufigkeit der elektromechanischen Komponenten wurden 20 Schaltungen / Tag angenommen,
- als β Wert (Anteil der hardwarebedingten Ausfälle in Folge gemeinsamer Ursache) wurde der Wert aus der Norm angenommen $\beta_d = 1\%$.

Nachfolgend wird ein Beispiel der erstellten Zuverlässigkeits-Blockdiagramme für die Sicherheitsfunktion „ Notfallerdung ausgelöst durch unerlaubtes Öffnen einer DAB-Verbindungstür“ gegeben.

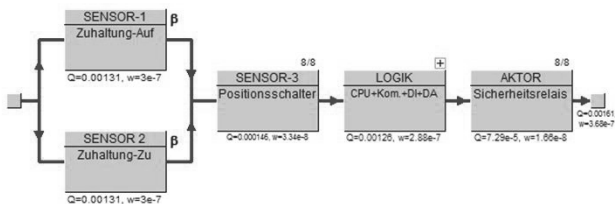




Abbildung 8: RBD für die gesamte Kette inklusiv das mit + gekennzeichneten Modul (Logik)

Für die Berechnungen wurde das Tool Reliability Workbench 10.2 der Firma Isograph verwendet.

Die Überlebenswahrscheinlichkeit der gesamten Kette lässt sich wie folgt ermitteln:

$$R_{\text{Kette}}(t) = \left[1 - \prod_{k=1}^n (1 - R_i(t)) \right] * \left[\prod_{l=1}^m R_l(t) \right].$$

Im Falle der Exponentialverteilung errechnet sich der Erwartungswert aus

$$E(T) = \frac{1}{\sum_{i=1}^m \lambda_i} \text{ für die Reihenschaltung mit } (m = 4 \text{ oder } 8) \text{ und}$$

$$E(T) = \frac{1}{\lambda} \sum_{k=1}^n \frac{1}{k} \text{ für die redundante Anordnung mit } (n = 2).$$

Unter Berücksichtigung des β -Faktors ist der THR-Wert für die gesamte Kette durch

$$\lambda_d = \beta_d \frac{2\lambda_d}{3} + \sum_{i=1}^m \lambda_{d,i}$$

gegeben.

8 Begutachtung

Die Begutachtung wurde gemäß dem Standard-Prozess des TÜV Rheinland für funktionale Sicherheit an Bahnanlagen durchgeführt.

Es wurden daher zum einen alle relevanten Dokumente begutachtet, insbesondere der Sicherheitsplan, das Gefahrenprotokoll und natürlich der Sicherheitsnachweis.

Dabei wurden zunächst Kommentare zu Dokumenten abgegeben, die dann umgesetzt wurden.

Diese Kommentare wurden in Kurzprüfberichten zu einem oder mehreren Dokumenten aufgenommen.

Daneben wurden zwei Audits durchgeführt. Zum einen ein Audit zum Qualitätssystem und zu den Sicherheitsprozessen, zum anderen ein Audit zu den Softwareprozessen. Dabei wurden die Zeitpunkte jeweils so gewählt, dass einerseits die jeweiligen Prozesse schon liefen, andererseits aber auch genug Zeit war um gegebenenfalls Korrekturen einzubringen.

Abschließend erstellte der Gutachter ein Gutachten in dem bescheinigt werden konnte, dass keine sicherheitsrelevanten Mängel vorhanden waren.

9 Zusammenfassung

In dieser Veröffentlichung wurde dargestellt, wie eine Abschalt- und Erdungsautomatik (AEA) als Teil des Bahnsystems nach Prinzipien der funktionalen Sicherheit entwickelt, implementiert und begutachtet wurde. Dies war eine erstmalige Anwendung der Normen EN 50128 und EN 50129 auf die AEA-Anlage. Es hat sich gezeigt, dass diese Bahnnormen als Normen zur funktionalen Sicherheit von Bahnstromanlagen erfolgreich angewendet werden können. Es haben sich keine Probleme bei der Anwendung dieser Normen gezeigt, begonnen von der PHA bis hin zum Sicherheitsnachweis und auch eine Begutachtung nach diesen Normen hat nicht zu Problemen, z.B. hinsichtlich der Auslegung geführt.

Referenzen

- [1] EN 50126 Spezifikation und Nachweis der Zuverlässigkeit, Verfügbarkeit, Instandhaltbarkeit und Sicherheit (RAMS), 1999
- [2] EN 50128: Bahnanwendungen – Telekommunikationstechnik, Signaltechnik und Datenverarbeitungssysteme – Software für Eisenbahnsteuerungs- und Überwachungssysteme, 2011
- [3] EN 50129: Bahnanwendungen - Telekommunikationstechnik, Signaltechnik und Datenverarbeitungssysteme, Sicherheitsrelevante elektronische Systeme für Signaltechnik, 2003
- [4] IEC 61508-5: Funktionale Sicherheit sicherheitsbezogener elektrischer/elektronischer/programmierbarer elektronischer Systeme, Teil 5; Beispiele zur Ermittlung der Stufe der Sicherheitsintegrität, 2011

Absicherung der Gesamtzuverlässigkeit eines Getriebes unter Berücksichtigung von Vorkenntnissen



Anwendung und Validierung

Dipl.-Ing. **Michael Bartholdt**, Prof. Dr.-Ing. **Bernd Bertsche**,
Institut für Maschinenelemente, Universität Stuttgart, Stuttgart;
Dipl.-Ing. **Martin Bollmann**, ZF Friedrichshafen AG, Friedrichshafen;
Prof. Dr.-Ing. **Peter Zeiler**, Hochschule Esslingen,
Fakultät Mechatronik und Elektrotechnik

Kurzfassung

Für das 6-Gang-Automatikgetriebe EcoLife der ZF Friedrichshafen AG wird ausgehend von den während der Produktentwicklung durchgeführten Berechnungen und Erprobungen auf verschiedenen Systemebenen die Gesamtzuverlässigkeit prognostiziert und mit dem Ausfallverhalten im Feld verglichen. Es wird zum einen dargestellt, wie diese unterschiedlichen Informationsquellen miteinander verknüpft werden können, um die Zuverlässigkeitswerte einer betrachteten Komponente zu präzisieren. Zum anderen, wie diese wiederum mit Zuverlässigkeitskennwerten anderer Komponenten zur Aussage auf einer höheren Ebene, bis hin zur Systemebene, verbunden werden können.

1. Motivation

Die Absicherung der Zuverlässigkeit eines neuen Getriebes erfolgt in der Produktentwicklung über Berechnungen und Erprobungen. Die Berechnungen werden meistens für Komponenten des Getriebes durchgeführt und dienen der betriebs- oder dauerfesten Auslegung. Die Erprobungstests erfolgen üblicherweise auf unterschiedlichen Systemebenen. Beispielsweise werden Tests auf Komponenten-, Baugruppen- oder Gesamtsystemebene durchgeführt. Die Tests können darüber hinaus unterschiedliche Zielsetzungen haben und als sogenannte „Success-Run-Tests“ (SR) oder „End-of-Life-Tests“ (EoL) konzipiert sein.

Die zentralen Fragestellungen sind,

- wie sich diese unterschiedlichen Informationsquellen nutzen lassen, um daraus eine Gesamtaussage zur abgesicherten Zuverlässigkeit des Getriebes abzuleiten, und
- ob die während der Produktentwicklung prognostizierte abgesicherte Zuverlässigkeit durch das Verhalten im Feld bestätigt wird.

2. Prinzipielles Vorgehen

Vorkenntnisse aus Berechnung und Simulation sowie aus Versuchen werden zur Aussage über die aktuelle Zuverlässigkeit auf Komponenten-, Subsystem- und Systemebene berücksichtigt. Die Verknüpfung verschiedener Vorinformationen innerhalb eines Teilsystems (Komponente, Baugruppe o.ä.) erfolgt durch einen Bayes-Ansatz [1] basierend auf Betaverteilungen wie in [2] (S. 39-42) beschrieben. Zur Nutzung von Zuverlässigkeitswerten verschiedener Komponenten als Vorinformation zu einem höheren Systemniveau kommt die Approximation über (Beta-)Verteilungsmomente [3] zum Einsatz, wie z.B. in [4] (S. 95-96) beschrieben.

Das Zuverlässigkeitsmodell, d.h. die zuverlässigkeitstheoretische Systemstruktur mit allen relevanten bzw. zu betrachtenden Subsystemen und Komponenten bildet dabei den Ausgangspunkt. Die a-priori-Zuverlässigkeitsverteilung bildet die wesentliche Grundlage für die Anwendung des Satzes von Bayes, der sie mit weiteren Informationen aus zusätzlichen, aktuellen Tests, zur a-posteriori-Zuverlässigkeitsverteilung verknüpft. Eine Validierung des Vorgehens entsteht aus dem Abgleich der errechneten, theoretischen Zuverlässigkeit mit der Feldzuverlässigkeit auf verschiedenen Systemebenen.

In den folgenden Kapiteln wird als theoretische Grundlage das Vorgehen zur Verknüpfung von Vorkenntnissen nach dem Bayes-Ansatz bzw. der Momentenmethode erläutert.

3. Grundlagen zur Verknüpfung von Vorkenntnissen mit aktuellen Testergebnissen

Vorinformationen aus Berechnungen/ Simulationen oder vorangegangenen Versuchen konstruktiv ähnlicher Produkte bzw. des gleichen Produkts werden zur Beurteilung der Zuverlässigkeit ergänzend zu aktuellen Tests berücksichtigt. Somit lässt sich die Aussagewahrscheinlichkeit erhöhen oder der für den Zuverlässigkeitsnachweis nötige Stichprobenumfang reduzieren.

Der Satz von Bayes verknüpft Vorkenntnisse in Form der a-priori-Dichte mit dem Ausfallverhalten des aktuelleren Produktstands zur a-posteriori-Dichte. Sie stellt eine verbesserte Schätzung des tatsächlichen Werts der Zufallsgröße „Zuverlässigkeit“ dar.

Im Rahmen dieser Arbeit stammt die a-priori-Dichte aus Berechnungen, die zusätzlichen Informationen aus Tests.

Die aus dem Bayes-Theorem resultierende a-posteriori-Dichte lässt sich mathematisch wie folgt angeben, vgl. z.B. [5] (S. 274):

$$f(R|E_X) = \frac{P(E_X|R)f(R)}{\int_0^1 P(E_X|R)f(R)dR} \quad (1)$$

mit $f(R) \triangleq$ a-priori-Dichte, $P(E_x|R) \triangleq$ bedingte Wahrscheinlichkeit, $f(R|E_x) \triangleq$ a-posteriori-Dichte, $R \triangleq$ Zuverlässigkeit.

Die bedingte Wahrscheinlichkeit wird typischerweise als binomialverteilt beschrieben [2] (S. 37):

$$P(E_x|R) = \binom{n}{x} R^{L_v b_r b_p n - x} (1 - R)^x \quad (2)$$

mit $n \triangleq$ Stichprobenumfang, $x \triangleq$ Anzahl der Ausfälle, $b \triangleq$ Weibull-Formparameter bei Feldbedingungen, $b_p \triangleq$ Weibull-Formparameter bei Prüfbedingungen, $L_v =$ Lebensdauerverhältnis, $r =$ Raffungsfaktor.

Betrachtet wird die durch eine Binomialverteilung beschriebene bedingte Wahrscheinlichkeit (vgl. (2)) und die a-priori-Dichte der Zuverlässigkeit in Form der Betaverteilung (vgl. (3), mit A_0 und B_0 als die verteilungsbeschreibenden Parameter mit „0“ als Kennzeichnung der Vorinformation). Die Methode wie hier dargestellt und angewandt überträgt das Vorwissen aus der Komponentenerprobung in vollem Umfang. Ein teilweiser Übertrag ist möglich, wie z.B. in [2] (S. 75) vorgeschlagen.

Aus der a-priori-Dichte der Form

$$f(R) = \frac{1}{\beta(A_0, B_0)} R^{A_0-1} (1 - R)^{B_0-1} \quad (3)$$

wird die a-posteriori-Dichte der folgenden Form:

$$f(R|E_x) = \frac{R^{A_0-1+n-x} (1 - R)^{B_0-1+x}}{\beta(A_0 + n - x, B_0 + x)} \quad (4)$$

Dies entspricht der Betaverteilung mit den verteilungsbeschreibenden Parametern:

$$A = (A_0 + n - x); B = (B_0 + x) \quad (5)$$

Die Integration von Gleichung (4) liefert die Aussagewahrscheinlichkeit P_A .

Vorkenntnisse aus Berechnung resultieren in einer theoretischen Stichprobenreduktion Δn_{ber} . Sie berechnet sich unter der Annahme einer Stufenfunktion für $f(R)$ mit Stufensprung bei $R(t_s)$ und eines Success-Run-Äquivalents der Berechnung (kein Ausfall) wie folgt:

$$\Delta n_{ber} = \frac{\ln \left(\frac{R_{soll}(t_s)^2}{R_{ber}(t_s)} \right)}{\ln(R_{soll}(t_s))} \quad (6)$$

Hierbei wird angenommen, dass R_{ber} mit einer Aussagewahrscheinlichkeit von $P_A = 50\%$ vorliegt. Die hier angewandte Berechnung von Δn_{ber} ersetzt die in [2] (S. 69-70) vorgeschlagene numerische Rechenweise. Es handelt sich weiterhin um eine mit Dr. Krolo abgestimmte Alternative. Die für diesen Ansatz zu Grunde gelegte Stufenfunktion als a-priori-Dichte $f(R)$ ist in Bild 1 dargestellt. Sie weist einen Stufensprung bei $R_{soll}(t_s)$ auf. Die berechnete Zuverlässigkeit zum Zeitpunkt der Solllebensdauer t_s ($R_{ber}(t_s)$) muss größer als das Quadrat

der Soll-Zuverlässigkeit zum gleichen Zeitpunkt ($R_{soll}(t_s)^2$) sein, da sonst eine negative Stichprobenreduktion errechnet wird. In der Praxis kann i.d.R. davon ausgegangen werden, dass $R_{ber} > R_{soll}$ ist.

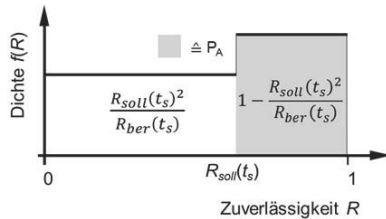


Bild 1: Zuverlässigkeitsdichte der a-priori-Verteilung zur Berücksichtigung von Vorkenntnissen aus Berechnungsergebnissen

Besteht das Vorwissen aus Berechnungsergebnissen und bestehen die aktuellen Tests aus i Success Runs (SR) oder End-of-Life-Tests (EoL) mit Lebensdauerverhältnissen (L_v), Raffungsfaktoren (r) und im Falle der EoL-Tests mit x Ausfällen, ändert sich (5) in (7). Hier sind keine Vorkenntnisse aus Tests enthalten, weswegen A_0 und B_0 entfallen.

$$A = \sum_i (L v_i \cdot r_i)^{b_i} \cdot n_i + \Delta n_{ber} - x; \quad B = 1 + x \quad (7)$$

Dies sind die verteilungsbeschreibenden Parameter der a-posteriori-(Beta-)Verteilung der Zuverlässigkeitsdichte. Aus ihr bestimmt sich der Zuverlässigkeitswert bei gegebener Aussagewahrscheinlichkeit in dieser Arbeit.

4. Grundlagen zur Verknüpfung von Vorkenntnissen zur höheren Systemebene

Vorkenntnisse einer untergeordneten Systemebene werden zur Zuverlässigkeitsverteilung der übergeordneten Systemebene verknüpft, vgl. Bild 2. Hier geschieht das über die Approximation über Verteilungsmomente der Betaverteilung.

Die a-priori-Systemverteilung auf dieser nun höheren Systemebene kann als Vorkenntnis wie in Kapitel 3 beschrieben genutzt und somit mit Testergebnissen aktueller Tests verknüpft werden.

Über Gleichungen (8) werden die Beta-Verteilungsparameter der Zuverlässigkeitsverteilung der übergeordneten Systemebene berechnet. Es handelt sich um einen approximativen Ansatz. Die sich ergebende Verteilung stellt die a-priori-Dichte dar, falls aktuelle Testergebnisse auf dieser Systemebene vorliegen und beide Informationen verknüpft werden sollen. Dies geschieht z.B. nach [2]. Der Erwartungswert entspricht hierbei dem ersten Moment der Systemverteilung, vgl. Gleichung (9). Die Methode wie hier dargestellt und angewandt

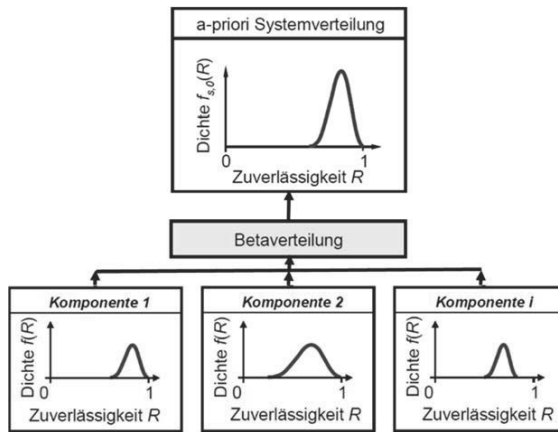


Bild 2: Verknüpfung von Zuverlässigkeitsdichten auf Komponentenebene zur a-priori-Systemverteilung (nach [4] S.93, gekürzt)

überträgt das Vorwissen aus der Komponentenerprobung in vollem Umfang. Ein teilweiser Übertrag ist möglich, wie z.B. durch in [4] (S. 96) vorgeschlagen.

$$A = (1 - E(R)) \cdot \frac{E(R)^2}{\text{Var}(R)} - E(R); \quad B = A \cdot \frac{1 - E(R)}{E(R)} \quad (8)$$

Mit dem Erwartungswert der a-priori-Systemverteilung basierend auf A_i und B_i der Komponentenverteilungen:

$$E_s(R) = \prod_{i=1}^J \frac{A_i}{A_i + B_i} \quad (9)$$

Die Varianz $\text{Var}_s(R)$ der System-Verteilung errechnet sich wie folgt:

$$\text{Var}_s(R) = \prod_{i=1}^J \frac{A_i \cdot (A_i + 1)}{(A_i + B_i) \cdot (A_i + B_i + 1)} - E_s^2(R) \quad (10)$$

Weitere Methoden zur Bildung der (a-priori-)Systemverteilung über Zuverlässigkeitsverteilungen der beinhalteten Komponenten existieren. Erwähnt seien die unter Umständen exakte Mellin-Transformation, sowie die approximative Bestimmung über Monte-Carlo-Simulationen. Ein Vergleich der Methoden wird in [6] erbracht.

5. Zusammengeführtes Vorgehen

Bild 3 zeigt das Schema der angewandten Vorgehensweise zur Berücksichtigung von Vorkenntnissen innerhalb einer Komponente (vgl. Kapitel 3) bzw. zur Verknüpfung der so gefundenen a-posteriori-Verteilungen verschiedener Komponenten zur höheren Systemebene (vgl. Kapitel 4). Die aus der Verknüpfung von Vorkenntnissen zur höheren Systemebene

entstandene Beta-Verteilung entspricht der a-priori-Verteilung eben dieser höheren Systemebene. Diese wird, verknüpft mit Informationen aus aktuellen Testergebnissen, über die in Kapitel 3 dargestellte Methode mittels des Satzes von Bayes, zur a-posteriori-Systemverteilung, die um diesen Informationsgehalt präzisiert ist, vgl. Bild 3.

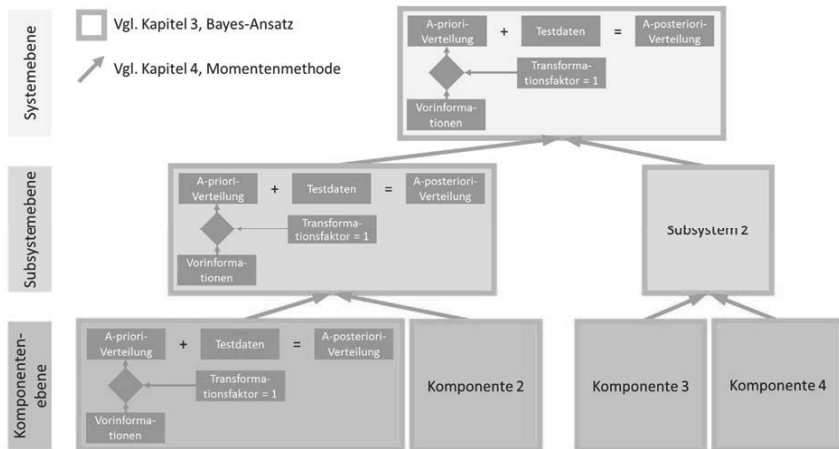


Bild 3: Angewandte Vorgehensweise zur Berücksichtigung von Vorkenntnissen

6. Das Beispielgetriebe EcoLife

Die reduzierte Systemstruktur, die modellhaft das Gesamtsystem EcoLife zuverlässigkeitstechnisch beschreibt, ist in Bild 4 zusammen mit dem Abbild des Getriebes dargestellt. Die in dieser Arbeit zuverlässigkeitstechnisch betrachteten 16 Komponenten und Subsysteme sind umrahmt. Für sie werden Vorinformationen aus Berechnung berücksichtigt.

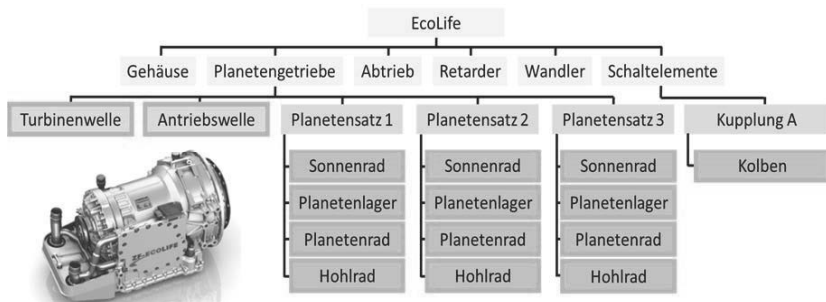


Bild 4: Reduzierte Systemstruktur des Beispielgetriebes EcoLife

Es liegen des Weiteren folgende Annahmen zu Grunde:

- Die Systemstruktur ist rein seriell.
- Die Zuverlässigkeitsdichten werden durch Beta-Verteilungsdichten beschrieben.
- Die Zuverlässigkeit der in dieser Arbeit nicht explizit betrachteten Elemente der Systemstruktur wird zu $R(t_s) = 1$ (100%) angenommen.
- Jeder Planetensatz wurde mehrmals ausfallfrei getestet und deren Komponentenbeanspruchungen zu individuellen Lebensdauerverhältnissen ausgewertet. Die Unabhängigkeit der Tests auf Subsystemebene (Planetensatz) und Komponentenebene wird vorausgesetzt. Ergebnisse der Tests fließen auf zwei Systemebenen mit ein.

7. Berechnung der a-posteriori-Zuverlässigkeit beispielhaft anhand Planetensatz 1

Zunächst werden Vorkenntnisse aus Berechnungen zu den einzelnen Komponenten, die der Baugruppe Planetensatz 1 (vgl. Bild 4) zugeordnet sind, über den Satz von Bayes mit aktuellen Testergebnissen verknüpft (vgl. Kapitel 3). Dies erfolgt exemplarisch anhand der Komponente „Planetenenlager“. Anschließend werden die jeweilig ermittelten a-posteriori-Zuverlässigkeitsdichten aller 4 Komponenten (Sonnenrad, Planetenenlager, Planetenrad, Hohlrad) über die Approximation über Verteilungsmomente der Betaverteilung (vgl. Kapitel 4) zur (a-priori-) Zuverlässigkeitsdichte der höheren Systemebene des Planetensatzes verknüpft. Prämissen und Ergebnisse der beispielhaften Berechnung sind in Tabelle 1 zusammengefasst. Sensible Daten sind durch # unkenntlich gemacht.

Tabelle 1: Berechnungsprämissen und -ergebnisse Planetenenlager, Zusammenfassung

t_s	1 [-]	b	1,5 [-]
$R_{soll}(t_s)$	#%	t_0	0,44 [-]
$P_A(R_{soll}(t_s))$	50%	T	# [-]
t_{ber}	8,72 [-]	$R_{ber}(t_s)$	#%
$R(t_{ber})$	#%	Δn_{ber}	1,98 Stück

Die Zuverlässigkeit bei Solllebensdauer $R_{ber}(t_s)$ wird wie folgt berechnet [5] (S.43):

$$R_{ber}(t_s) = (1 - F(t_{ber}))^{\left(\frac{t_s - t_0}{t_{ber} - t_0}\right)^b} \quad (11)$$

Die theoretische Stichprobenreduktion Δn_{ber} berechnet sich nach Gleichung (6).

Tabelle 2 fasst die Testergebnisse für die Komponente Planetenenlager 1 zusammen, also die Informationen der bedingten Wahrscheinlichkeit. Der verglichen zur Berechnung größere Wert für Formparameter b der Weibull-Verteilung soll berücksichtigen, dass sich bei der

Beschreibung einer tatsächlichen (z.B. empirischen) Ausfalldichte mit beobachteter ausfallfreie Zeit t_0 unterschiedliche b -Werte für eine 2- bzw. 3-parametrige Weibull-Verteilung ergeben. Der b -Wert für die 2-parametrige Weibull-Verteilung tendiert größer zu sein. Im Test wurde ZF-seitig von einer 2-parametrigen Weibull-Verteilung ausgegangen.

Tabelle 2: Testergebnisse Planetenlager 1, Zusammenfassung, 3 Success Runs (SR)

n	Pro SR 1 Stück	L_v1	1,18 [-]
b	2 [-]	L_v2	2,41 [-]
r	1 [-]	L_v3	1,69 [-]

Unter Anwendung der Gleichungen (8) werden die beschreibenden Parameter der Beta-Verteilung berechnet. Die so entstandene a-posteriori-Zuverlässigkeitsdichte des Planetenlagers sowie die weiteren in Planetensatz 1 gemäß Systemstruktur (vgl. Bild 4) berücksichtigten Komponenten sind in Bild 5 (dünn, gestrichelt) dargestellt. Jene des Sonnenrads und des Hohlrads sind zufällig identisch.

Ebenso in Bild 5 dargestellt ist die Zuverlässigkeitsdichte des Planetensatzes 1 (fett, gestrichelt). Die beschreibenden Parameter wurden nach Gleichungen (8)-(10) berechnet. Diese Zuverlässigkeitsdichte wird als a-priori-Dichte verstanden (vgl. Kapitel 5), repräsentiert durch die Summanden A_0 und B_0 , und mit den vorliegenden Testergebnissen verknüpft. Aus Gleichung (7) wird dann Gleichung (12).

$$A = A_0 + \sum_i (Lv_i \cdot r_i)^{b_i} \cdot n_i + \Delta n_{ber} - x ; \quad B = B_0 + x$$

(12)

Liegen, wie hier der Fall, keine Berechnungsergebnisse auf Planetensatz-Niveau vor, die als Vorkenntnisse berücksichtigt werden können, wird $\Delta n_{ber} = 0$.

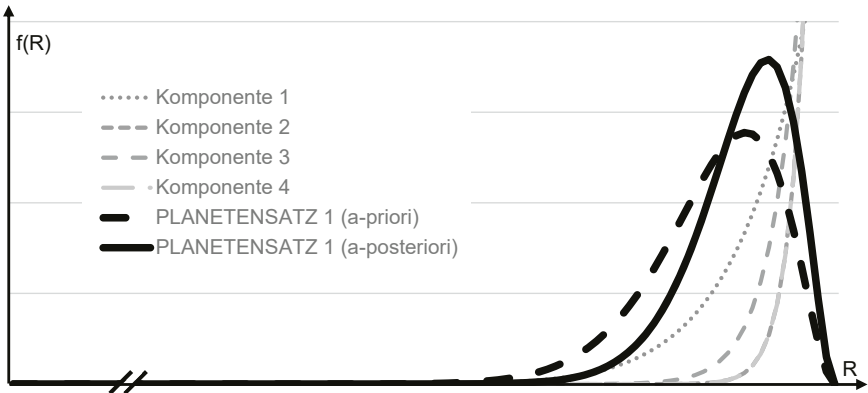


Bild 5: Zuverlässigkeitsdichten von Planetensatz 1

Planetensatz 1 wurde in 3 Tests mit verschiedenen Lebensdauerverhältnissen ohne Ausfall getestet. Die Berücksichtigung dieser Testergebnisse gemäß dem Ansatz beschrieben in Kapitel 3 bzw. nach Gleichung (12) führt zu einer verbesserten Zuverlässigkeitsaussage bzw. bei gleicher Aussagewahrscheinlichkeit zur Prognose einer höheren Zuverlässigkeit. Bild 5 stellt die so berechnete a-posteriori-Dichte (d.h. mit Berücksichtigung der Testergebnisse) des Planetensatzes 1 dar (fett, durchgezogen).

8. Die Zuverlässigkeitsdichten weiterer Komponenten, Baugruppen bzw. Subsysteme

Die Zuverlässigkeitsdichten der weiteren Einheiten gemäß Systemstruktur werden ähnlich dem oben genannten Vorgehen bestimmt. Auf z.T. abweichende Prämissen wird im Folgenden eingegangen.

Turbinenwelle: Es wird eine gleichverteilte Zuverlässigkeitsdichte als Vorinformation berücksichtigt. Das entspricht mathematisch $\Delta n_{ber}=1$ (vgl. [2] S.70). Aus den Testergebnissen werden 6 EoL-Tests und ein SR auf verschiedenen Belastungsniveaus berücksichtigt. Die Testergebnisse werden auf die äquivalente Lebensdauer bei Feld-Belastung unter Anwendung des Inverse-Power-Law bezogen und somit Lebensdauerverhältnisse berechnet. Als Formparameter wird ein für Wellen typischer Wert angenommen.

Antriebswelle: Die Berechnungsergebnisse werden mit Ergebnissen aus 5 SR-Tests verknüpft. Die Belastungen im Test variieren. Die erreichten Lastwechselzahlen werden auf ihr Äquivalent bei Feld-Belastung bezogen um die Lebensdauerverhältnisse zu erhalten. Der Testcharakter (SR oder EoL) spielt für die Berücksichtigung des „Lebensdauer“-Verhältnisses formell keine Rolle. Ausfällen wird über den Summand x Rechnung getragen, vgl. Gleichung (12). Über eine Amplitudentransformation nach [7] (S. 185) werden verschiedene Kollektivanteile auf eine rein wechselnde Belastung reduziert.

Planetensätze 2 und 3: Auf gleiche Weise, wie in Kapitel 7 beispielhaft die Zuverlässigkeitsdichte des Planetensatzes 1 berechnet wird, geschieht dies nun für die Planetensätze 2 und 3. Wieder werden die a-priori-Dichten aus der Berechnung als Vorkenntnis mit je drei EoL-Testergebnissen zur a-posteriori-Dichte verknüpft. Diese a-posteriori-Dichten werden wieder über die Approximation durch die Betaverteilung zur a-priori-Dichte auf höherem Systemniveau.

Kolben A: Berechnungsergebnisse beziehen sich auf die als am kritischsten identifizierte Stelle. Dabei werden verschiedene Oberflächenrauheiten berücksichtigt. Testergebnisse liegen für zwei verschiedene Belastungen (Drücke) aus je drei SR vor. Auf die jeweils ertragbare Lastwechselzahl wird wieder über den Ansatz des Inverse-Power-Law geschlossen. Der Exponent des Inverse-Power-Law wird über ähnliche Produkte geschätzt.

Da die restlichen Komponenten des Subsystems „Schaltelemente“ nicht Teil der Untersuchung dieses Projekts sind (d.h. mit $R=1$ angenommen wurden) (vgl. Kapitel 6), entspricht die rechnerische Zuverlässigkeitsdichte von Kolben A der des übergeordneten Subsystems „Kupplung A“ und diese wiederum jener des Subsystems „Schaltelemente“.

Bild 6 illustriert u.a. die Zuverlässigkeitsdichten der Subsysteme Turbinen- und Antriebswelle, Planetensätze 1-3 und Kupplung A.

9. Zuverlässigkeitsdichte des Gesamtsystems EcoLife

Die Dichten der Planetensätze 1-3 ergeben zusammen mit der Dichte der Turbinenwelle und der Antriebswelle die Zuverlässigkeitsdichte des Planetengetriebes (vgl. Systemstruktur, Bild 4). Diese wiederum verknüpft mit jener von Kupplung A (identisch mit der Dichte für „Schaltelemente“) ergeben die Zuverlässigkeitsdichte für das Gesamtsystem EcoLife. Die Dichte der Kupplung A wird durch einen sehr hohen A- und einen vergleichsweise kleinen B-Parameter beschrieben (ähnlich wie auch die Dichten von Turbinenwelle und Antriebswelle). Das impliziert eine hohe Zuverlässigkeit. Kombiniert mit dem Planetengetriebe resultiert die Dichte des Gesamtsystems. Die Dichte des Planetengetriebes und des Gesamtsystems EcoLife stimmen nach dem angesetzten Modell fast überein, zumal sehr hohe Zuverlässigkeiten der Antriebs- und Turbinenwelle prognostiziert wurden, vgl. Bild 6.

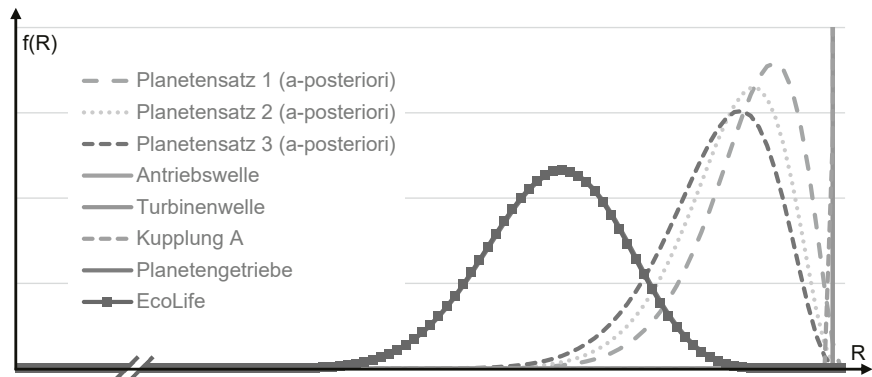


Bild 6: Zuverlässigkeitsdichten der Subsysteme sowie des Gesamtsystems EcoLife

10. Validierung: Vergleich prognostizierter mit beobachteter Zuverlässigkeit

Die Prognosegüte der Zuverlässigkeit wird retrospektiv über den Vergleich mit dem aus Felddaten ermittelten Zuverlässigkeitswert bestimmt. Dokumentierte Feldausfälle bilden die Grundlage für die Ausfallwahrscheinlichkeitsanalyse der betrachteten Einheiten.

Suspensionen (Überlebende) werden gemäß der angegebenen Produktionsvolumina berücksichtigt. Aus den erfassten Beanstandungen wurde eine Laufleistungsverteilung ermittelt, die den Suspensionen zu Grunde liegt. Die errechneten Prognosewerte der Zuverlässigkeit R bei $t_s = 1$ mit einer Aussagewahrscheinlichkeit von $P_A = 50\%$ wurden mit den im Feld beobachteten Werten verglichen.

Für Kolben A bzw. Kupplung A ergibt dem Modell gemäß sich gute Konformität zwischen vorausgesagter und der im Feld beobachteten Zuverlässigkeit. Im Falle des Planetengetriebes weicht die beobachtete Zuverlässigkeit im Feld gemäß Auswertung gegenüber der Prognose um etwa die Hälfte in Richtung höherer Zuverlässigkeit ab, was teilweise auf den Effekt der Serienschaltung aller Komponenten zurückzuführen sein könnte. Die Zuverlässigkeit der Schaltelemente wurde durch die Prognose leicht überschätzt. Ein Hauptgrund könnte in der starken Vereinfachung des Zuverlässigkeitsmodells des Teilsystems Schaltelemente liegen. Die beobachtete Gesamtzuverlässigkeit des Gesamtsystems EcoLife verhält sich zur prognostizierten ähnlich wie dies beim Planetengetriebe der Fall ist.

11. Zusammenfassung und Ausblick

Die angewandte Methode erscheint tendenziell konservativ, selbst bei voller Übernahme der Vorkenntnisse (hier: Transformationsfaktoren $\Phi=1$). Das dargestellte Beispiel der Zuverlässigkeitsprognose zeigt, dass hiernach die Prognose trotz Integration von Berechnungsergebnissen und Testergebnissen konservativer ist als die im Feld beobachtete Zuverlässigkeit. Als Methodenkombination eignet sich das dargelegte Vorgehen zur quantitativen, transparenten Abschätzung dennoch. Die prinzipielle Einsatzfähigkeit und der Nutzen der Methoden (vgl. Kapitel 3 und 4) bei isolierter und kombinierter Anwendung wurden aufgezeigt. Eine erweiterte Betrachtung der Übertragbarkeit der Prognose auf das Feldverhalten könnte die Einordnung der teilweise methodenbedingt eher konservativen Prognose unterstützen, z.B. über faktorielle Anpassungen, mit dem Ziel, dem Zuverlässigkeits-Nachweis beizutragen.

Ergebnisse aus EoL-Tests wirken sich methodenbedingt tendenziell negativ auf die a-posteriori Zuverlässigkeitsdichte aus, da die Anzahl der Ausfälle den Betaverteilungsparameter A verkleinert, B hingegen vergrößert (vgl. Gleichung (8) bzw. (12)). Bei ausschließlich SR-Tests würde A nicht gemindert werden und B minimal bleiben. Die Berücksichtigung weiterer Testergebnisse würde die Zuverlässigkeitsprognose ergänzen und präzisieren.

Raffungsfaktoren sind in das dargelegte Vorgehen zu integrieren, um Tests bei einer anderen als der Feldbelastung zu betrachten. Transformationsfaktoren würden einer nur teilweisen

Übertragbarkeit der Vorinformationen Rechnung tragen. Die Auswirkung alternativer Annahmen zur prinzipiellen Verteilung der Zuverlässigkeitsdichte aus Berechnungsergebnissen ist zu untersuchen.

Die dargelegte Struktur des Vorgehens zur Zuverlässigkeitsprognose auf verschiedenen Systemebenen bei Verknüpfung mehrerer verschiedener Subsystem-Zuverlässigkeiten diversen Ursprungs (hier: Berechnung + Test) ist praktikabel und eignet sich für eine wiederholte Aktualisierung der Zuverlässigkeitsaussage. Hierzu muss die a-posteriori Dichte wieder als a-priori-Dichte verstanden werden und mit neuen Test- bzw. Berechnungsergebnissen verknüpft werden. Die Genauigkeit der Zuverlässigkeitsaussage wächst als Folge.

Literatur

- [1] Bayes, T. u. Price, R.: An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. Philosophical Transactions of the Royal Society of London 53 (1763) 0, S. 370–418
- [2] Krolo, A.: Planung von Zuverlässigkeitstests mit weitreichender Berücksichtigung von Vorkenntnissen, Dissertation, Universität Stuttgart. Stuttgart 2004
- [3] Thompson, W. E. u. Haynes, R. D.: On the reliability, availability and bayes confidence intervals for multicomponent systems. Naval Research Logistics Quarterly 27 (1980) 3, S. 345–358
- [4] Maisch, M.: Zuverlässigkeitsorientiertes Erprobungskonzept für Nutzfahrzeuggetriebe unter Berücksichtigung von Betriebsdaten, Dissertation, Universität Stuttgart. Stuttgart 2007
- [5] Bertsche, B. u. Lechner, G.: Zuverlässigkeit im Fahrzeug- und Maschinenbau. Ermittlung von Bauteil- und System-Zuverlässigkeiten. VDI. Berlin [u.a.]: Springer 2004
- [6] Zeiler, P. u. Bertsche, B.: Component reliability allocation and demonstration test planning based on system reliability confidence limit. In: Podofilini, L., Sudret, B., Stojadinovic, B., Zio, E. u. Kröger, W. (Hrsg.): Safety and Reliability of Complex Engineered Systems. CRC Press 2015, S. 1805–1811
- [7] Haibach, E.: Betriebsfestigkeit. Verfahren und Daten zur Bauteilberechnung. VDI-Buch. Berlin: Springer 2006

Berücksichtigung sich ändernder Steuergeräteparameter auf entstehende Beanspruchungen für Zuverlässigkeitstests durch Monte-Carlo Simulation



Consideration of changing control unit parameters for resulting loads by Monte-Carlo simulation

Dr.-Ing. **Achim Romer**, M.Sc. **Jan-Michael Veith**,
Dr. Ing. h.c. F. Porsche AG, Weissach

Kurzfassung

Moderne Fahrzeuge verfügen über eine Vielzahl von Steuergeräten, welche notwendig sind, um Anforderungen an Verbrauch, Fahrleistung oder -komfort zu erfüllen. Dabei haben sowohl die Bedatung, wie auch deren Regelstrategien, einen großen Einfluss auf die sich ergebenden mechanischen Belastungen.

Der Beitrag vermittelt anhand konkreter Beispiele (Allradverteilung und Quersperre) quantitativ die sich durch Änderung des Datenstands ergebenden Lasten, sowie deren Auftretenswahrscheinlichkeit. Ferner wird ein Ansatz zur Zuverlässigkeitsbestimmung aufgezeigt. Kern des Ansatzes ist dabei, Umgebungs- und Fahrermodelle innerhalb von Simulationen per Monte-Carlo Methodik zu variieren, um schadenskritische Abläufe zu identifizieren und so verschiedene Bedatungen und Regelstrategien zu bewerten. So ist es möglich, die notwendigen Erprobungen zur Zuverlässigkeitsbestimmung zu verschlanken.

Abstract

Modern vehicles contain a variety of control units which are necessary in order to meet the requirements for fuel consumption, driving performance or travelling comfort. Data sets and control strategies of these control units have a big influence on resulting mechanical loads.

This paper gives a quantitative overview of the resulting loads caused by changes of data sets regarding specific examples (all-wheel distribution and axle-differential lock) and their probabilities of occurrence. Furthermore, an approach to evaluate a reliability analysis will be shown. The core of the approach is to vary the environmental and driver models within a simulation by Monte-Carlo method to identify critical courses and thus evaluate data sets and control strategies. This way, the extent of vehicle testing necessary for reliability evaluation can be reduced.

1. Einleitung

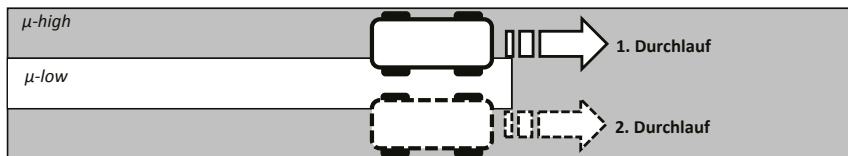
Durch die fortschreitende Digitalisierung verkürzen sich auch in der Automobilbranche die Produktlebenszyklen immer weiter. Dieser Umstand erzwingt eine Verkürzung der Entwicklungszeiten eines Produkts. Im Gegensatz hierzu steht die steigende Komplexität moderner Fahrzeugantriebsstränge. Neben der Kombination von Verbrennungsmotoren und elektrischen Antrieben führen auch immer leistungsfähigere Regelsysteme (beispielsweise zur Steigerung der Fahrdynamik oder des Komforts) zu einer steigenden Anzahl von Komponenten, sowie der Zunahme der Leistungsfähigkeit der einzelnen Bauteil-Steuergeräte. Dies führt zunehmend dazu, dass die Bedatung der Steuergeräte beziehungsweise die hinterlegte Regelstrategie maßgeblich die auftretenden mechanischen Belastungen bestimmt. Allerdings entwickeln sich Bedatung beziehungsweise Regelstrategie über den Produkt-Entstehungsprozess kontinuierlich weiter, was die genaue Betrachtung der sich daraufhin einstellenden Belastungen erfordert. Ein Ansatz hierzu ist die Anwendung einer Monte-Carlo-Simulation, um den Einfluss sich ändernder Steuergeräteparameter auf entstehende Beanspruchungen bewerten zu können.

2. Getriebeerprobung am Beispiel des Manövers „Überfahrt μ -Split“

Die Erprobung eines Getriebes kann in verschiedene Versuchsarten eingeteilt werden [1]. So werden bei der Funktionserprobung Grundfunktionalitäten, wie die Ölversorgung, sichergestellt. Bei der Dauererprobung werden Lastläufe gefahren, beispielsweise in Form einer Beanspruchungs-Zeit-Funktion aus zuvor berechneten Lastkollektiven. Eine Sonderform stellen Manöver dar, welche durch unsachgemäßen Gebrauch des Fahrzeugs Belastungen in Höhe der Kurzzeitfestigkeit einzelner Bauteile erzeugen können [2] – im Folgenden „Missbrauchsmanöver“ genannt.

Eine Form dieser Missbrauchsmanöver basiert auf unterschiedlichen Reibwerten zwischen Reifen und Straße je Rad beziehungsweise Achse. Diese Manöver stellen das Anfahren auf teilweise vereisten oder verschmutzten Fahrbahnuntergründen dar, wobei μ -high einen hohen Reibwert (μ circa 1) zwischen Straße und Reifen bezeichnet, μ -low einen niedrigen (μ circa 0,1).

Das Manöver „Überfahrt μ -Split“ zielt auf einen maximalen Energieeintrag in die Achsgetriebe ab, kombiniert mit einer hohen Stoßbelastung des gesamten Antriebstrangs. Wie Bild 1 verdeutlicht, steht das Fahrzeug zu Beginn des Manövers jeweils mit beiden Rädern einer Fahrzeugseite auf μ -high, mit den anderen beiden auf μ -low.

Bild 1: Manöver "Überfahrt μ -Split"

Das Fahrzeug wird beschleunigt, bei angetriebenen Rädern ohne Quersperre im Achsgetriebe stellt sich eine Differenzdrehzahl zwischen linker und rechter Radseite ein. Beim Übergang des schlupfenden Rads von μ -low auf μ -high wird diese Differenzdrehzahl (schlagartig) wieder abgebaut. Das Manöver erhält seinen Missbrauchscharakter durch die Vollast-Beschleunigung in dieser Situation. Bild 2 zeigt beispielhaft die Verläufe der Drehmomente und Drehzahlen in Abhängigkeit der Fahrpedalstellung.

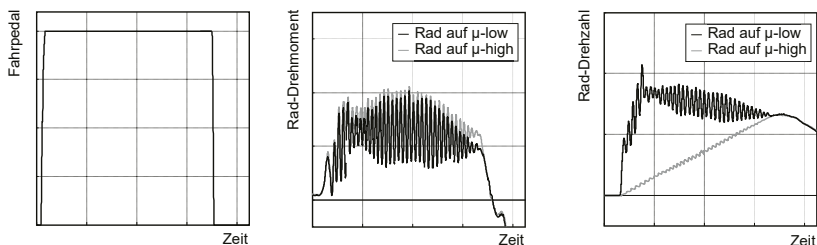


Bild 2: Verlauf von Fahrpedal, Rad-Drehmoment und Rad-Drehzahl

3. Monte-Carlo-Simulation zur Bestimmung der Streuung von Versuchsergebnissen

Wiederholt durchgeführte (reale) Manöver führen zu stochastisch verteilten Versuchsergebnissen, im Gegensatz zu meist deterministischen Simulationsergebnissen. Aufgrund minimaler Variation der Versuchsrandbedingungen streut beispielsweise das maximal auftretende Rad-Drehmoment während eines Manöverdurchgangs um einen bestimmten Wert. Zur Beschreibung der Streuung gibt es verschiedene Verteilungen, wie die Gleichverteilung oder die Normalverteilung. Die Auswertung der Messreihen ergab häufig asymmetrisch verteilte Messgrößen, welche durch eine Weibullverteilung besser beschrieben werden können [3].

Allgemein erzeugt bei der Monte-Carlo-Simulation (MCS) ein Zufallsgenerator gleichverteilte Zahlen (stochastisch unabhängig) im Intervall $[0,1]$ [4]. Gemäß dem schwachen Gesetz großer Zahlen strebt die relative Häufigkeit eines gleichverteilten Ereignisses gegen seinen

Erwartungswert. Die absoluten Häufigkeiten konvergieren jedoch nicht [5]. Eine Monte-Carlo-Simulation arbeitet daher mit einer großen Anzahl an Simulationsdurchläufen, bei denen unbekannte Parameter zufällig vorgegeben werden, bis ein Abbruchkriterium erreicht wird. Anwendung findet die MCS bei Problemstellungen, welche sehr komplex und meist nur noch numerisch zu lösen sind [6].

Soll beispielsweise das Maximum einer Funktion innerhalb eines Intervalls $I = [a, b]$ gefunden werden, welche nicht differenzierbar ist, so bietet sich der Einsatz einer MCS an. Für zufällige Werte $x \in [a, b]$ wird der Funktionswert $f(x)$ berechnet und zwischengespeichert. Mit steigender Anzahl an Berechnungen nähert sich der gefundene Maximalwert dem echten Maximum an. Um gute Ergebnisse zu erzielen, sind jedoch sehr viele Berechnungsschritte notwendig.

Eine Möglichkeit zur Steigerung der Effizienz sind mehrstufige MCS. Hier werden in der ersten Stufe Zufallszahlen aus dem gesamten Intervall gezogen. Nach einer festgelegten Anzahl an Durchläufen wird das Intervall auf den Bereich eingeeengt, in dem die höchsten Werte gefunden wurden. In der nächsten Stufe werden dann Zufallszahlen aus dem eingeeengten Intervall gezogen. In weiteren Schritten wird das Intervall schrittweise eingeeengt, siehe Bild 3.

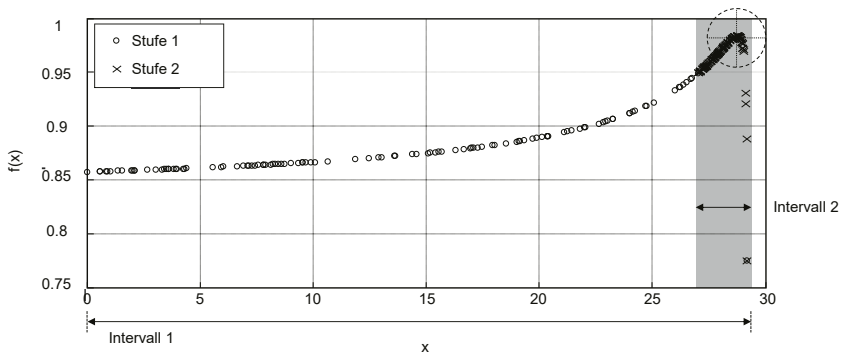


Bild 3: Suche des Maximums mit einer zweistufigen MCS

Durch eine zufällige Parametrisierung eines Simulationsmodells wird eine Streuung des Simulationsergebnisses erzeugt, wobei die Streuung der Eingangsparameter einer zuvor definierten Verteilung folgt. Eine große Herausforderung hierbei ist die Wahl der Verteilungsparameter für die Eingangsgrößen der MCS [7]. Bei internen Versuchen zeigte sich, dass kleine Abweichungen hier zu einer großen Verfälschung der Ergebnisse führen. Ein Ansatz zur Bestimmung dieser Verteilungsparameter aus Messdaten wird in Kapitel 4 gezeigt.

4. Bestimmung auftretender Belastungen und deren Streuung durch sich ändernde Steuergeräteparameter

Im Folgenden wird das gesamte Vorgehen zur Bestimmung der auftretenden Belastungen aufgezeigt. Das Hauptaugenmerk liegt dabei auf einer Methode zur Bestimmung der Verteilungsparameter für die Eingangsgrößen der MCS durch eine zweite überlagerte MCS, siehe „Optimierung“ in Bild 4.



Bild 4: Vorgehen zur Bestimmung auftretender Belastungen und deren Streuung durch sich ändernde Steuergeräteparameter

4.1. Aufbau des Simulationsmodells

Der Antriebstrang des Fahrzeugs wird als Fünf-Massen-Schwinger abgebildet, wobei die Massen über eine Drehsteifigkeit und eine geschwindigkeitsproportionale Dämpfung interagieren, wie in Bild 5 dargestellt.

Der Reifen-Straße-Kontakt wird über ein Pacejka-Reifenmodell dargestellt: In Abhängigkeit des Schlupfs und des Reibkontakts Reifen-Straße kann der Reifen Kräfte übertragen, dadurch wird das Gesamtfahrzeug (als Massenpunkt modelliert) beschleunigt.

Des Weiteren ist die im Vergleichsfahrzeug vorhandene elektronische Differentialsperre im Modell abgebildet. Ein PID-Regler regelt die Differenz aus Fahrzeuggeschwindigkeit und Radgeschwindigkeit zu Null, wenn die Differenzgeschwindigkeit einen Schwellwert übersteigt. Das Motordrehmoment wird aus Fahrpedal und Motordrehzahl berechnet, wobei ein Zweipunktregler bei Erreichen der maximal zulässigen Motordrehzahl das Motordrehmoment ausblendet. Um ein reales Ansprechverhalten des Motors abzubilden, verzögern PT1-Glieder den Aufbau des Drehmoments.

Der Reibwert μ zwischen Straße und Reifen wird für jedes Rad einzeln vorgegeben. Gemäß dem Versuchsaufbau bleibt der Reibwert des μ -high-Rads konstant, während das andere Rad zum Zeitpunkt t -Sprung von $\mu = \mu$ -low auf $\mu = \mu$ -high ansteigt.

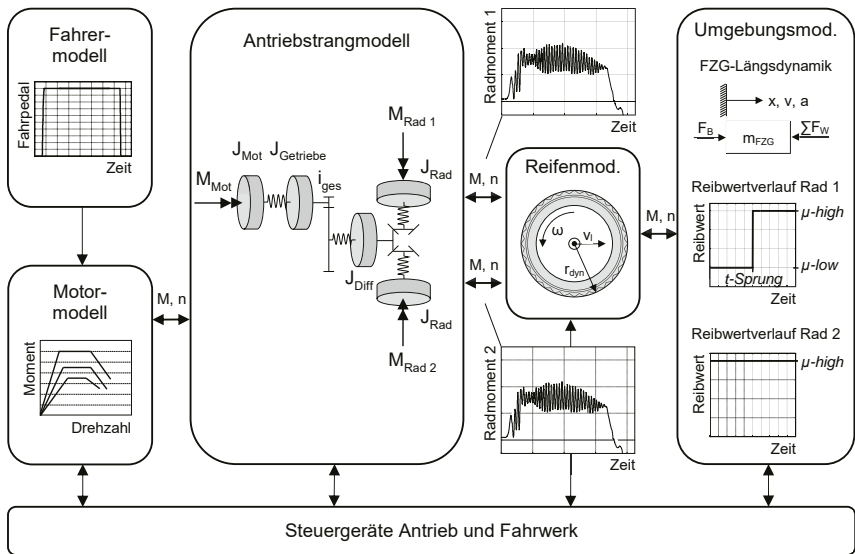


Bild 5: Aufbau des Simulationsmodells

4.2. Bestimmung realitätsnaher Verteilungsparameter für die Eingangsgrößen der Monte-Carlo-Simulation

Ziel der Variation der Simulationsparameter ist eine Modellparametrisierung, welche bei vielfacher Simulation eines Manövers Belastungen mit vergleichbarer Streuung analog zu Realmessungen erzeugen. Stellvertretend für auftretende Belastungen wird nachfolgend das maximal auftretende Rad-Drehmoment während eines Simulationsdurchgangs betrachtet.

Bild 6 zeigt den Ablauf der Simulation: In einem Durchgang werden die Parameter μ -low, μ -high und t -Sprung zufällig erzeugt, wobei diese Parameter weibullverteilt vorliegen. Anschließend wird das maximale Rad-Drehmoment M -max dieses Durchgangs berechnet. Für feste Weibull-Parameter b_x und T_x wird das Simulationsmodell n mal berechnet, wobei jeder Durchgang, aufgrund zufällig gewählter Parameter μ -low, μ -high und t -Sprung, andere maximale Rad-Drehmomente M -max erzeugen. Diesen Drehmomenten wird wiederum eine

Verteilung angenähert, welche mit der Verteilung der maximal auftretenden Rad-Drehmomente aus der vorliegenden Fahrzeugmessung verglichen werden kann.

Im nächsten Schritt wird für jeden Weibull-Parameter b_x und T_x der drei zu variierenden Parametern ein neuer Wert bestimmt, dieser wird zufällig (gleichverteilt) zwischen zwei zuvor definierten Grenzwerten ausgewählt. Nun werden wieder n Simulationsdurchgänge durchgeführt und eine Verteilung der maximal auftretenden Rad-Drehmomente berechnet. Dieser Schritt wird m mal wiederholt, wobei durch den Vergleich der Verteilung zwischen Simulation und Messung die Grenzwerte der Weibull-Parameter fortlaufend optimiert werden.

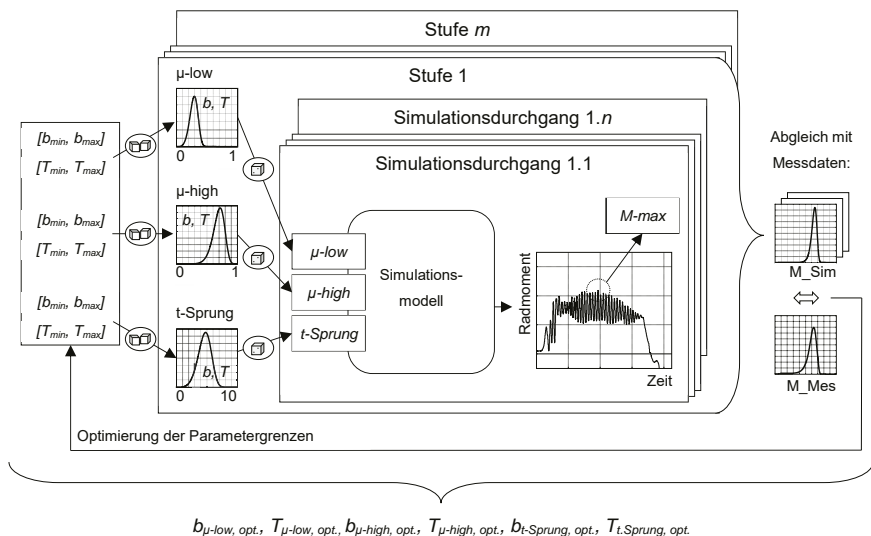


Bild 6: Ablauf der Simulation

Ergebnis der Optimierung ist eine Parametrisierung b_{opt} und T_{opt} für jeden der Parameter μ -low, μ -high und t -Sprung, so dass die wiederholte Simulation des Manövers maximale Rad-Drehmomente M -max mit einer Verteilung erzeugt, welche sich der Verteilung der Rad-Drehmomente aus den gemessenen Versuchsdaten annähert. Bild 7 zeigt das beste Simulationsergebnis, welches nach rund 150.000 Simulationsdurchgängen erzielt wurde.

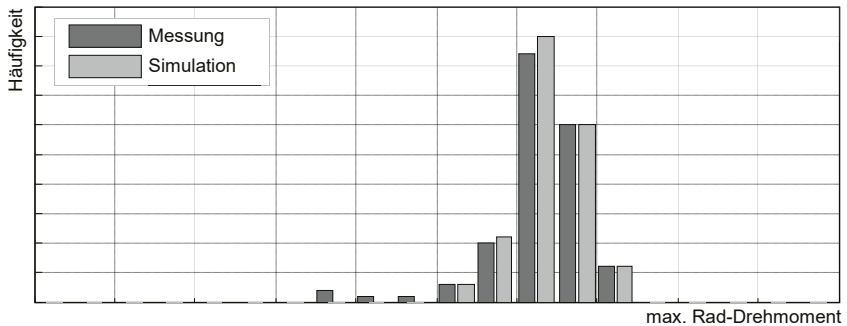


Bild 7: Häufigkeiten der max. auftretenden Rad-Drehmomente

Mit jedem Optimierungsschritt verkleinern sich dabei die Intervalle, aus denen die Weibull-Parameter b_x und T_x gezogen werden. Bild 8 zeigt den Verlauf der Ober- und Untergrenzen eines Parameterpaars und deren Mittelwerte über die ausgeführten Optimierungsschritte.

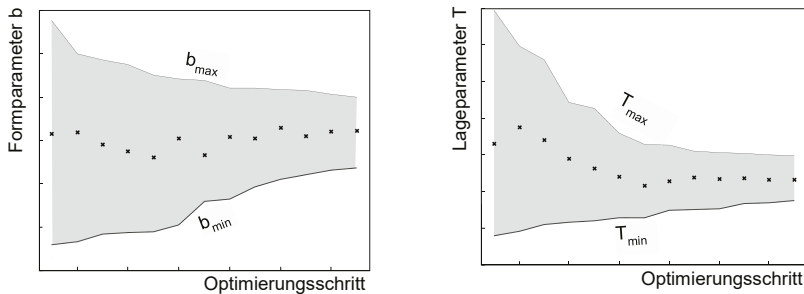


Bild 8: Optimierung der Verteilungsparameter

Durch die Optimierung der Parametergrenzen nähert sich die Streuung der maximal auftretenden Rad-Drehmomente der Streuung der Versuchsergebnisse an. In Bild 9 sind die über mehrere Simulationsdurchläufe gemittelten relativen Häufigkeiten des Rad-Drehmomentes über der Anzahl an Optimierungsschritten dargestellt. Zu Beginn der Optimierung streuen vereinzelte Werte stark nach oben, erst gegen Ende der Optimierung treten diese Extremwerte nicht mehr auf.

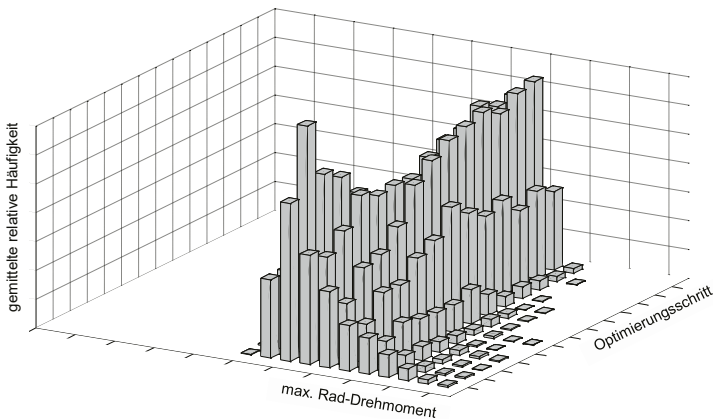


Bild 9: Gemittelte relative Häufigkeiten je Optimierungsschritt

4.3. Berücksichtigung sich ändernder Steuergeräteparameter

Im Entwicklungsprozess eines Fahrzeugs werden Regelstrategien und somit Bedatungen von Steuergeräten kontinuierlich weiterentwickelt. Im vorliegenden Simulationsmodell werden nun beispielhaft veränderte Datenstände einzelner Steuergeräte eingebunden: das Simulationsmodell wird also mit veränderten Parametern der betrachteten Steuergeräten erneut n mal simuliert. Für diese Simulation werden die im vorherigen Schritt bestimmten Streuungsparameter der Eingangsgrößen verwendet. Der zuvor erfolgte Abgleich der sich ergebenden Streuung mit real gemessenen Streuungen soll dabei sicherstellen, dass auch die sich aus den neuen Simulationen ergebenden Streuungen möglichst realitätsnah ergeben. Ergebnis ist wieder eine Verteilung der maximal auftretenden Rad-Drehmomente, welche mit den zuvor erzeugten Daten verglichen wird. Daraus erfolgt im Anschluss eine Bewertung des Datenstands hinsichtlich der zu erwartenden Belastungen für den Fahrzeugantriebstrang. Im ersten Beispiel wurden die Parameter des Steuergeräts des Porsche Stability Managements (PSM) verändert, siehe Bild 10. Dieser Datenstand beeinflusst das Regelverhalten des Automatischen Bremsen Differenzials (ABD), also die Höhe des durch die Bremsanlage aufgebauten Drehmoments bei Differenzdrehzahl zwischen linker und rechter Fahrzeugseite. Änderungen im Datenstand können sich beispielsweise durch geänderte Anforderungen an die Fahrdynamik ergeben.

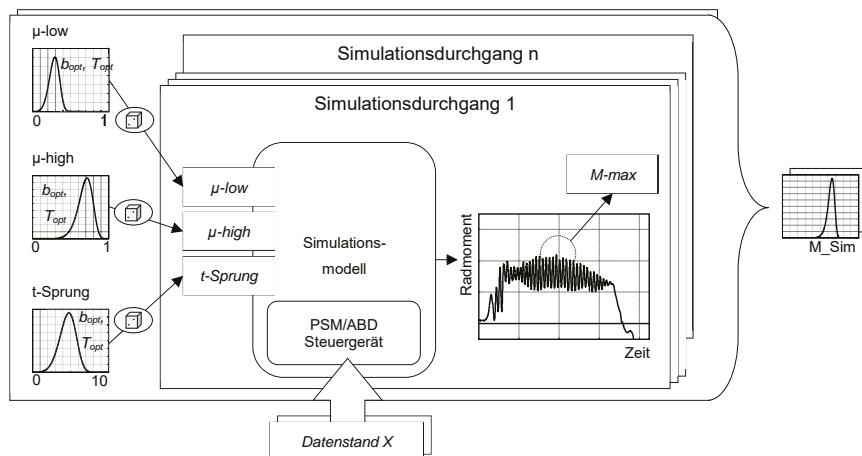


Bild 10: Änderung des Datenstands des PSM-Steuergeräts

Die Simulation wurde mit zwei veränderten Datenständen durchgeführt, die daraus resultierenden Verteilungen der maximal auftretenden Rad-Drehmomente sind in Bild 11 dargestellt. So führt Datenstand 1 zu einer Verlagerung zu geringfügig niedrigeren Rad-Drehmomenten, während Datenstand 2 im Durchschnitt höhere Werte mit geringerer Streubreite erzeugt.

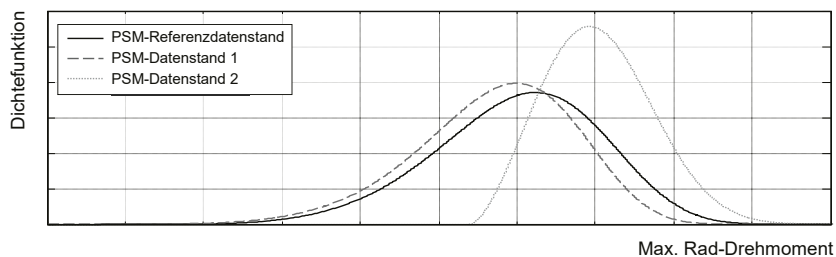


Bild 11: Dichtefunktionen max. Rad-Drehmoment je PSM-Datenstand

Aufgrund der deutlich höheren Belastungen bei Datenstand 2 muss nun geprüft werden, ob bisher durchgeführte Versuche wiederholt werden müssen. Im Gegensatz dazu macht Datenstand 1, aufgrund der geringfügig niedrigeren Belastungen, die Wiederholung der Versuche nicht nötig.

Im zweiten Beispiel wurden die Auswirkungen verschiedener Strategien der Allradverteilung untersucht. Dazu wurden wiederum zwei alternative Datenstände des Porsche Traction Management (PTM) -Steuergeräts eingebunden.

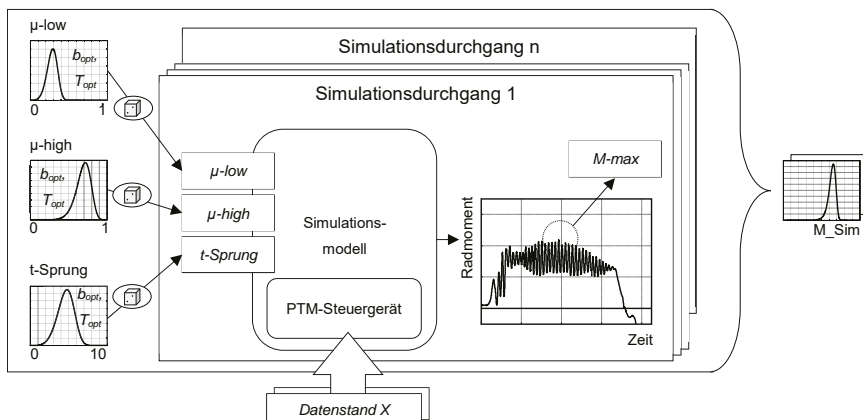


Bild 12: Änderung des Datenstands des PTM-Steuergeräts

Datenstand 1 führte im Vergleich zum Referenzdatenstand zu deutlich niedrigeren Rad-Drehmomenten bei einer ähnlichen Streuung. Datenstand 2 erzeugte hingegen deutlich höhere Rad-Drehmomente.

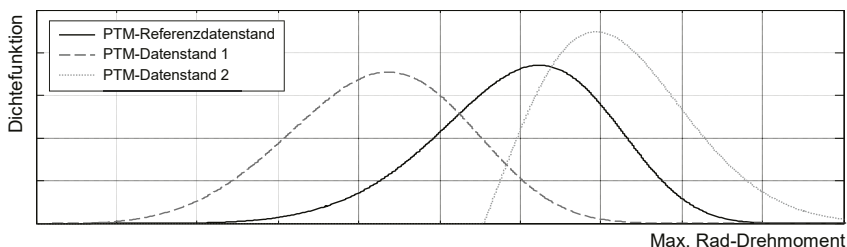


Bild 13: Dichtefunktionen max. Rad-Drehmoment je PTM-Datenstand

Somit führt Datenstand 1 zu einer höheren Lebensdauer des betrachteten Bauteils, hier kann sogar aus Gewichts- oder Kostengründen eine Verschlanung des Bauteils in Betracht gezogen werden. Bei Datenstand 2 müssen etwaige Versuche zur Freigabe des Bauteils wiederholt werden.

5. Zusammenfassung und Ausblick

In diesem Beitrag wurde ein Vorgehen entwickelt, um den Einfluss sich ändernder Steuergeräteparameter auf entstehende Belastungen im Fahrzeug-Antriebsstrang zu bewerten. Ein bestehendes Simulationsmodell wurde auf Grundlage einer Referenzmessung aus einem Fahrzeugversuch derart trainiert, dass es vergleichbare Messgrößenverläufe wie in der Realmessung generiert. Durch eine Monte-Carlo-Simulation wurden die Simulationsparameter des Modells variiert, damit die Ergebnisse einer Streuung unterliegen. Mit Hilfe einer zweiten, überlagerten Monte-Carlo-Simulation wurden die vorgegebenen Verteilungen der Simulationsparameter dahingehend optimiert, dass sich die Streuung der Simulationsergebnisse der Streuung der Ergebnisse aus den Fahrzeugmessungen annähert. Nach Abschluss dieser Initialisierungsphase wurden beispielhaft die Datenstände zweier Bauteil-Steuergeräte verändert, um deren Auswirkung auf entstehende Belastungen zu beurteilen. Daraus resultierte jeweils eine Verteilung der auftretenden Belastungen, was Grundlage für eine Beurteilung eines Datenstands ist.

Zukünftig sollten die betrachteten Bauteil-Steuergeräte über eine X-in-the-Loop Umgebung in das Simulationsmodell eingebunden werden, damit die zu untersuchenden Datenstände vergleichbar zum Fahrzeug auf die Steuergeräte aufgespielt werden können.

6. Literaturverzeichnis

- [1] H. Naunheimer, B. Bertsche und G. Lechner, Fahrzeuggetriebe, Berlin Heidelberg: Springer Verlag, 2007.
- [2] F. Küçükay und W. Pfau, „Extrembelastungen in Personenwagen-Antriebssträngen,“ *Automobiltechnische Zeitschrift* 91, 1989.
- [3] B. Bertsche und G. Lechner, Zuverlässigkeit im Fahrzeug- und Maschinenbau, Stuttgart: Springer Verlag, 2004.
- [4] J. Veith, Bewertung von Optionen unter der Coherent Market Hypothesis, Wiesbaden, Diss.: Deutscher Universitäts-Verlag, 2006.
- [5] T. Adamek, Statistik für Anwender, Berlin: Springer Spektrum, 2016.
- [6] H. Nahrstedt, Die Monte-Carlo-Methode, Wiesbaden: Springer Vieweg, 2015.
- [7] A. Romer, „Anwendungsspezifischer Zuverlässigkeitsnachweis auf Basis von Lastkollektiven und Vorwissen,“ Institut für Maschinenelemente der Universität Stuttgart, Diss., Stuttgart, 2014.

Berücksichtigung inhomogener Anforderungen in der Zuverlässigkeitsabsicherung am Beispiel eines Fahrwerksystems eines 1st-Tier Zulieferers der Nutzfahrzeugindustrie



Tobias Buschhaus, M.Sc., Institut für Qualitäts- und Zuverlässigkeitsmanagement GmbH, Wuppertal;
Dipl.-Ing. **Patrick Glöckner**, Dr.-Ing. **Michael Metzele**,
BPW Bergische Achsen Kommanditgesellschaft, Wiehl

Kurzfassung

Dieser Artikel stellt eine Methode dar, inhomogene Anforderungen in der Zuverlässigkeitsabsicherung zu berücksichtigen. Die Anwendbarkeit der Methode wird am Beispiel eines Fahrwerksystems eines 1st-Tier-Zulieferers aufgezeigt. Ein wichtiger Bestandteil der Methode ist die Bestimmung der Belastungen durch den sogenannten kritischen Kunden des Gesamtmarktes. Die Methode nutzt u.a. die Simulation, um unterschiedliche Märkte und Anforderungsprofile in gemeinsame Bewertungsgrößen zu überführen. Es wird weiterhin aufgezeigt, welche Mehrwerte für die Entwicklung bzw. den Hersteller durch optimale Kombination der verschiedenen Einflussfaktoren generiert werden können. Darüber hinaus wird ein Ausblick auf weitere Anwendungsmöglichkeiten gegeben.

1. Einleitung

Im Spannungsfeld verkürzter Entwicklungszeiten und betriebswirtschaftlich steigender Risiken sehen sich Nutzfahrzeug-Zulieferer (im weiteren Nfz-Zulieferer) vor der Herausforderung, Systeme zu entwickeln, die sowohl den Anforderungen verschiedener OEM als auch den Anforderungen stark unterschiedlicher Nutzer bzw. Märkte genügen. Hierbei stellt auch nach aktuellen Erkenntnissen die Zuverlässigkeit das entscheidende Kaufkriterium für den Spediteur dar. Aufgrund des steigenden Kostendrucks und der Gewichtssensibilität (zunehmend bedingt durch CO₂-Emissions-Beschränkungen) im Nfz-Bereich ist eine Überdimensionierung der Systeme keine zielführende Lösung, um die optimale Zuverlässigkeit über einen weiten Einsatzbereich gewährleisten zu können. In diesem Zielkonflikt rücken die anforderungsgerechte Entwicklung und Erprobung zunehmend in den aktuellen Fokus der Zulieferindustrie.

Der Stand der Technik in der Nfz-Industrie bezieht die Belastungsgröße „jährliche Fahrleistungsverteilung“ zur Zuverlässigkeitsabsicherung heran. Die vorgestellte Methode berücksichtigt hinsichtlich der Erprobungsauslegung weitere auf das System einwirkende Schädigungsparameter und Anforderungsprofile, um die Mindestzuverlässigkeit des Systems abzusichern. Des Weiteren wird aufgezeigt, wie die einzelnen unterschiedlichen Nutzungs- und Anforderungsprofile zusammengeführt werden können, so dass je Schädigungsparameter der auslegungsrelevante Zielwert für das 95% Kunden-/Fahrerkollektiv („kritischer Kunde“) bestimmt werden kann. Diese Konsolidierung des marktspezifischen Nutzerverhaltens sowie der Schädigungsparameter der Einzelmärkte zu einem Gesamtmarktmodell ist die wesentliche Voraussetzung einer Zuverlässigkeitsziel-orientierten und kostenoptimierten Erprobungsplanung. In der Nfz-Industrie gibt es sehr viele verschiedene Einsatzgebiete welche mit identischen Produkten bedient werden, so dass daraus ein breites Anforderungsspektrum resultiert. Der Vergleich der Anforderungen eines Trailers im innerstädtischen Verteilerverkehr in Europa mit einem Langstreckeneinsatz in Australien, den sogenannten Roadtrains, zeigt die Unterschiede bei den Anforderungen deutlich auf. Nicht nur die jährliche Fahrleistung variiert erheblich, sondern diverse andere Parameter unterscheiden sich wesentlich, wie beispielsweise die „Anzahl der jährlichen Lenkungen“, die Belastung durch bestimmte „Umweltfaktoren“ oder zahlreiche Belastungskollektive, die aufgrund der Straßenbeschaffenheit entstehen. Zum Zuverlässigkeitsnachweis des Produktes ist es daher notwendig, die beispielhaft genannten Parameter marktübergreifend zu bewerten [1].

Daher ist eine Methode entwickelt worden, die sowohl die verschiedenen Märkte kombiniert, als auch die verschiedenartigen Schädigungsparameter verknüpft, um so die marktübergreifenden auslegungsrelevanten Zielwerte (95% Kunden-/Fahrerprofil) je Schädigungsparameter zu bestimmen, gegen welche das System abgesichert werden soll.

2. Methode zur Berücksichtigung inhomogener Anforderungen

Eine relevante Auslegungsgröße ist gegeben durch den Schädigungsmechanismus und das probabilistische Ausfallverhalten (z.B. charakterisierbar durch die zweiparametrische Weibull-Verteilung β und T^*) und wird als Schädigungsparameter definiert. In der Realität wirken stets mehrere Einflüsse gleichzeitig auf ein System, so dass das Ausfallverhalten mehrdimensional beeinflusst wird und mehrere Schädigungsparameter definiert werden müssen. Die unterschiedlichen Schädigungsparameter können in zwei Kategorien unterschieden werden, die funktions- oder systembedingten und die umgebungs- oder einsatzbedingten. Die

Einteilung dient als Handlungshilfe zur Beschreibung der Märkte und des zugehörigen Nutzungsverhaltens. Die Tabelle 1 zeigt einige Beispiele für diese Einteilung.

Tabelle 1: Einteilung der Schädigungsparameter

Umgebungs- oder einsatzbedingte Schädigungsparameter
Luftfeuchtigkeit
Umgebungstemperatur
Witterungsbedingungen
Funktions- oder systembedingte Schädigungsparameter
Fahrleistung
Anzahl der Lenkungen
Anzahl der Start/Stop-Vorgänge

Es hat sich als zielführend herausgestellt, die Schädigungsparameter auf ein jährliches Intervall zu normieren. Die Zeit ist in der Regel immer noch die führende Auslegungsgenngroße in der Nfz-Industrie. Mit dieser Angabe der erwarteten Nutzungsdauer und der jährlichen Normierung der Schädigungsparameter können die für die Lebensdauer relevanten Belastungen der Schädigungsparameter ermittelt werden. Zur Quantifizierung von Schädigungsparametern können unterschiedliche Datenquellen genutzt werden, vergleichbar mit der Datenbasis der Belastungskollektive. So erfolgt die Quantifizierung beispielsweise auf Basis von Sensordaten, Felddaten oder Ergebnissen von Kundenbefragungen bzw. Marktforschungsstudien.

Die Beschreibung der Schädigungsparameter mit Hilfe von Zeitreihen ist eine weitere mögliche Methode, wie z.B. bei umweltbedingter Schädigung. Auf die Beschreibung sowie Anwendung von bedingt aufeinander folgenden Ereignisse mit Hilfe von Zeitreihen wird verzichtet und sei hier nur der Vollständigkeit halber erwähnt.

Eine weitere relevante Auslegungsgroße sind die Marktanteile. Durch eine unterschiedliche Ausprägung der Schädigungsparameter in den einzelnen Märkten können die zu erwartenden Beanstandungen sehr unterschiedlich ausfallen, was großen Einfluss auf den wirtschaftlichen Erfolg des Entwicklungsprojektes nehmen kann. Nur unter Einbeziehung der Marktanteile kann eine wirtschaftliche Risikobetrachtung realisiert werden. Abbildung 1 zeigt schematisch die Gegenüberstellung inhomogener Anforderungen für ein Entwicklungsprojekt.

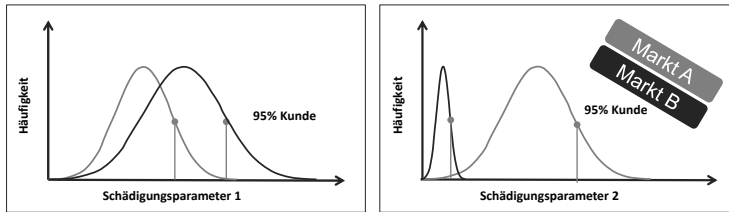


Bild 1: Inhomogene Markt-Anforderungen

Im Beispiel soll ein System für zwei Märkte (A und B) und zwei Schädigungsparameter (1 = jährliche Fahrleistung und 2 = jährliche Anzahl an Lenkungen) ausgelegt werden. Markt A hat einen Anteil von 70% und Markt B von 30%. Beide Schädigungsparameter hängen kausal zusammen und weisen eine Korrelation auf.

Aktuell ist es in der Praxis üblich, eine gemeinsame Führungs- bzw. Auslegungsgröße zu definieren (z.B. Nutzungszeit oder Fahrleistung) und einen konstanten Umrechnungswert (z.B. Mittelwert) zur Skalierung der weiteren Parameter zu nutzen. Dies kann dazu führen, dass die Ausprägungen der Parameter (z.B. Streuung) nicht vollständig berücksichtigt und somit unter Umständen die weiteren Parameter über- bzw. unterschätzt werden. Zusätzlich zeigt der Vergleich der marktübergreifenden Nutzung, dass die Korrelation zwischen zwei Schädigungsparametern global sehr inhomogen sein kann. Die Nutzung eines konstanten Wertes zur Umrechnung kann dazu führen, dass die vorhandenen Wechselwirkungen zwischen Marktnutzung und Schädigungsmechanismus nicht genügend abgebildet werden.

Die folgende Betrachtung des Beispiels erläutert diesen Zusammenhang näher. Markt A zeichnet sich durch eine niedrige Fahrleistung und große Anzahl an Lenkungen aus. Markt B weist eine große Fahrleistung aber eine geringe Anzahl an Lenkungen auf. Wird nur nach einem Markt ausgelegt, z.B. Markt A, könnten Schädigungsparameter in den anderen Märkten unterrepräsentiert sein (hier: für Markt B die Fahrleistung). Eine Auslegung des Systems auf den kritischen Teilmarkt je Schädigungsparameter, hier also eine große Fahrleistung und viele Lenkungen, hätte aber eine Überdimensionierung zur Folge.

Als Führungsgröße in diesem Modell wurde die Nutzungsdauer (Zeit) festgelegt. Die anderen Schädigungsparameter beziehen sich auf die Führungsgröße, werden aber nicht linear umgerechnet. D.h. ein zehn Jahre alter Trailer hat nicht zwingend 500.000 km zurückgelegt, sondern kann auch 1.000.000 km oder nur 280.000 km als Fahrleistung haben. Eine weitere wichtige Einschränkung ist, dass die Schädigungsparameter voneinander unabhängig sind, auch wenn dies nicht der Realität entspricht.

Diese Ausgangssituation und die Kombination aller Parameter und Einsatzgebiete bringen den gängigen Ansatz an seine Grenzen. Sind darüber hinaus die Marktanteile, wie im Beispiel, zudem sehr unterschiedlich, verstärkt sich zusätzlich das Potenzial zur Überdimensionierung des Systems und einhergehend der Reduzierung der Wirtschaftlichkeit des Geschäftsmodells. Die Bestimmung des markt- und schädigungsparameterübergreifenden 95%-Nutzers ist nötig, um den einsatzbedingten, wirtschaftlichen und rechtlichen Anforderungen zu genügen.

Im Folgenden wird die Erweiterung des bestehenden VDA-Ansatzes zur Bestimmung des kritischen Kunden vorgestellt. Der hier vorgestellte Ansatz berücksichtigt zusätzlich zu [3] die Auswirkung sowohl differierender Märkte und unterschiedlicher Marktanteile als auch Einflüsse verschiedener Schädigungsmechanismen. Abbildung 2 zeigt die Methode zur Bestimmung des marktübergreifenden kritischen Nutzers für einen Schädigungsparameter anhand von zwei exemplarischen Märkten.

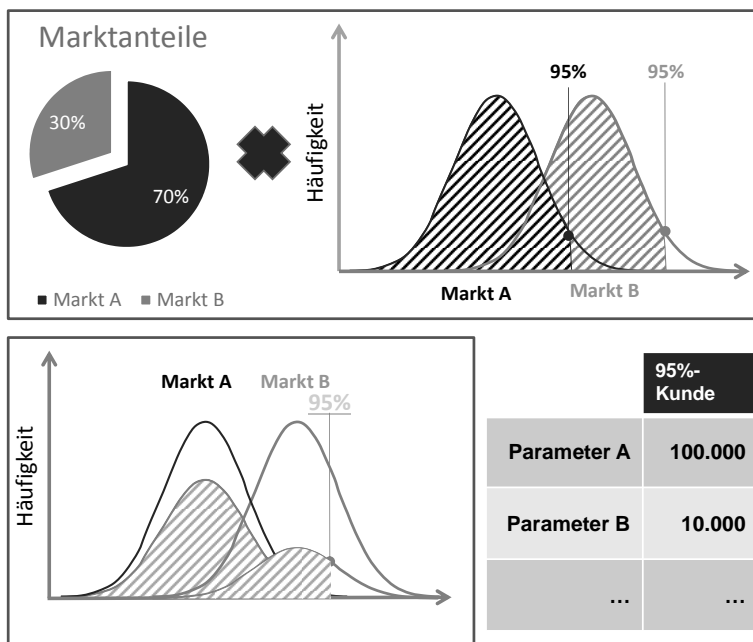


Bild 2: Methode zur Bestimmung des 95%-Kunden

Die Definition der Märkte und die Prognose der Marktanteile ist Ausgangspunkt der Erweiterung. Der Prognosehorizont der Marktanteile sollte sich nicht auf den SOP beschränken,

sondern die Abschätzung von zukünftigen Trends beachten. Der Ansatz bezieht so zukünftige Entwicklungen mit ein, wie z.B. die Erschließung neuer Märkte oder zukünftiger Technik-Trends. Nachdem die Märkte und ihre Anteile definiert sind, muss das spezifische Nutzungsverhalten je Markt beschrieben werden. Auf Basis von beispielsweise Belastungskollektiven, Messfahrten, Auswertungen von Statistik-Datenbanken oder Experteneinschätzungen wird das Nutzungsverhalten innerhalb jedes Marktes und für jeden einzelnen Schädigungsparameter bezogen auf eine geeignete und unabhängige Führungsgröße (z.B. km, Tage, Anzahl pro Überfahrt) definiert. Dies ist Voraussetzung für die folgenden Schritte.

Als Eingangsgröße für die Bestimmung des marktübergreifenden kritischen Nutzers liegen nun die Marktanteile und die Belastungskollektive je Schädigungsparameter und Markt vor. Aus der zufälligen und gewichteten Kombination der Eingangsgrößen mittels Monte-Carlo-Simulation (MCS) resultiert die marktgewichtete Grundgesamtheit der Führungsgröße. Unter Einbeziehung einer Spline-Interpolation wird eine Hüllkurve gebildet, deren Näherung für die weiteren Schritte nicht relevant ist. Die MCS wird genutzt, da diese verteilungsunabhängig ist. Weiter können mathematische Gesetze durch einfache Abfragen ersetzt werden, sodass die fehlende Normierung der gewichteten Eingangsgrößen ($P_{\text{einzel}}(\Omega) \neq 1$) nicht relevant ist.

Die Normierung der gewichteten Grundgesamtheit ($P_{\text{gesamt}}(\Omega) = 1$) folgt aus der Gesamtmarkt-Betrachtung. Das 95%-Quantil wird nach VDA Band 3.2 als auslegungsrelevanter kritischer Kunde empfohlen. Durch Bestimmung des 95%-Quantils der marktgewichteten Grundgesamtheit ist der marktübergreifende kritische Kunde für den Schädigungsparameter bestimmt. Dieses Vorgehen wird für alle definierten Schädigungsparameter durchgeführt. Das Resultat ist der markt- und schädigungsparameterübergreifende kritische Kunde, wie der Matrix in Abbildung 2 zu entnehmen ist.

Alternativ zu dieser Vorgehensweise hätte die Auslegung basierend auf dem 95%-Kunden des kritischsten Marktes erfolgen können. Solange keine sicherheitsrelevanten Fehlerbilder betrachtet werden, würde dies jedoch zu einer wirtschaftlich meist nicht tragbaren Überdimensionierung des Produktes führen. Ist aber ein bezüglich eines Schädigungsparameters besonders kritischer Markt mit geringer Stückzahl vorhanden, so wird dieser ggf. nur äußerst unzureichend in der Entwicklung berücksichtigt. Um in solchen Fällen eine Entscheidungshilfe zu generieren und das wirtschaftliche Risiko abzuschätzen, wird das Quantil des marktübergreifenden 95%-Kunden auf die Verteilung des Teilmarktes übertragen und somit die prozentuale Markterfüllung ermittelt. Hierdurch kann ein unterrepräsentierter Markt identifiziert werden und so die Basis für eine gezielte Risikoentscheidung gebildet werden. Diese Betrachtung kann z.B. die Ausgliederung eines Marktes in ein Sonderprojekt bewirken, um

das wirtschaftliche Risiko zu minimieren und die Kundenzufriedenheit langfristig zu gewährleisten. Die Anwendungsmöglichkeit des beschriebenen Ansatzes wird im Weiteren erläutert.

3. Berücksichtigung inhomogener Anforderungen in der Zuverlässigkeitsabsicherung

Zur Absicherung der Zuverlässigkeit wird die Success-Run-Methode (vgl. [2] und [3]) eingesetzt. Diese folgt der Logik, dass n Prüflinge den Prüfplan absolvieren müssen, um ein Mindestzuverlässigkeit $R(t)$ zur Konfidenz γ abzusichern. Es ergibt sich eine nachgewiesene Mindestzuverlässigkeit von

$$R(t) = (1 - \gamma)^{\frac{1}{n}}. \quad (\text{F1})$$

Die Mindestzuverlässigkeit $R(t)$ des Systems ist für den markt- und parameterübergreifenden kritischen Kunden nachgewiesen, wenn das System gegenüber jedem Schädigungsparameter die erforderliche Zuverlässigkeit aufweist. Verschiedene Prüfungen und Musterstände können durch Überführung der Anzahl der Prüflinge n in die Feld-Äquivalenzbauteile $n_{\text{Äqui}}$ gesamtheitlich berücksichtigt werden. Die Voraussetzung für diese Überführung ist die Berechnung des Lebensdauerverhältnisses L_V , welches den Lastfaktor zwischen den Bedingungen im Feld und in den Prüfungen beschreibt. Der Raffungsfaktor r berücksichtigt die Komprimierung der Belastung im Versuch gegenüber der Feldlast. Der Musterfaktor MF bezieht den jeweiligen Entwicklungsstand des Prüflings mit ein, kann aber auch genutzt werden um den Bezug zwischen Komponenten und System zu beschreiben. Die Bestimmung des Musterfaktors ist nicht fest vorgegeben und wird meist durch Experteneinschätzung bestimmt. Die Wechselwirkungen zwischen Komponenten und System müssen jedoch in der Prüfung abgebildet und für die Schädigungsmechanismen relevant sein. Die reale Prüfungsgröße $t_{\text{Lab; real}}$ ergibt sich direkt aus den Prüfkollektiven. Es ist zu berücksichtigen, dass nicht jedes Prüfkollektiv jeden Schädigungsparameter repräsentiert und der Schädigungsmechanismus durch Wechselwirkungen beeinflusst sein kann. Die Bezugsgröße im Feld t_{Feld} ist gegeben durch den marktübergreifenden 95%-Kunden. Die Einheiten der Prüfungsgröße und der Bezugsgröße im Feld sind nicht an die Einheit Zeit gebunden. Das Lebensdauerverhältnis L_V je Prüfkollektiv, Musterstand und Schädigungsparameter ist gegeben durch

$$L_V = \frac{t_{\text{Lab; real}} \cdot r \cdot MF}{t_{\text{Feld}}}. \quad (\text{F2})$$

Zur Bestimmung der Feld-Äquivalenzbauteile $n_{\ddot{A}qui}$ muss das Ausfallverhalten je Schädigungsparameter mit Hilfe des Formparameters der Weibull-Verteilung $\beta_{SchäPa}$ aufgenommen werden. Dieser wird häufig im Rahmen des Zuverlässigkeitsprozesses mit der Felddatenanalyse bestimmt. Mit Hilfe des Lebensdauerverhältnisses L_V und des Formparameters der Weibullverteilung $\beta_{SchäPa}$ ergibt sich das Feld-Äquivalenzbauteil zu

$$n_{\ddot{A}qui} = n \cdot L_V^{\beta_{SchäPa}}. \quad (F3)$$

Die einzelnen Äquivalenzbauteile der i Prüfungen summieren sich zur Gesamtzahl der Äquivalenzbauteile je Schädigungsparameter $n_{\ddot{A}qui;SchäPa}$ (siehe F4),

$$n_{\ddot{A}qui;SchäPa} = \sum_{k=1}^i n_{\ddot{A}qui;k} = \sum_{k=1}^i n_k \cdot L_V^{\beta_{SchäPa}}. \quad (F4)$$

Daraus folgt die nachgewiesene marktübergreifende Mindestzuverlässigkeit für den jeweiligen Schädigungsparameter bezogen auf den kritischen Kunden $R(t)_{\Sigma;SchäPa}$ durch Einsetzen von (F4) in (F1) zu

$$R(t)_{\Sigma;SchäPa} = (1 - \gamma)^{\frac{1}{n_{\ddot{A}qui;SchäPa}}} = (1 - \gamma)^{\frac{1}{\sum_{k=1}^i n_k \times \left(\frac{t_{Lab;real;k} \cdot MF_k}{t_{Feld}} \right)^{\beta_{SchäPa}}}}. \quad (F5)$$

Um den Nachweis der Mindestzuverlässigkeit für den markt- und parameterübergreifenden Schädigungsparameter erbringen zu können, ist es nötig, die oben beschriebene Vorgehensweise für alle Schädigungsparameter durchzuführen. Wenn die Anzahl der Prüflinge je Prüfung optimal festgelegt worden ist, resultiert eine abgesicherte Zuverlässigkeit für das System. Einzelne Prüfungen können hierbei je nach Prüfungsart mehrere Schädigungsparameter oder nur einzelne Parameter abprüfen. Zudem können bestimmte Schädigungsparameter in verschiedenen Prüfungen mit unterschiedlichen Lebensdauerverhältnissen berücksichtigt sein. Eine gesamtheitlich optimierte Auslegung der Zuverlässigkeitsabsicherung muss eine Aufteilung aller Prüflinge auf die Prüfungen unter den Gesichtspunkten minimaler Kosten, geringem Prüfaufwand und nachgewiesener Mindestzuverlässigkeit gewährleisten. Dies stellt ein klassisches Optimierungsproblem dar.

4. Anwendung der Methoden an dem Beispiel eines Fahrwerksystems eines 1st-Tier-Zulieferers

Die vorgestellte Methode wird nachfolgend am Beispiel der Zuverlässigkeitsabsicherung eines Fahrwerksystems für Nfz (siehe Abbildung 4) anhand fiktiver Beispielwerte dargestellt.



Bild 4: BPW -Fahrwerksystem für einen Nfz-Trailer

Das Fahrwerksystem gehört zu der Produktgruppe „Starrachse“ und besteht aus mehreren Komponenten. Die Abhängigkeiten zwischen Komponente, System und Prüfungen sind bekannt und sind im Nachfolgenden durch Raffungs- und Musterfaktoren quantifiziert.

Das System wird für die Anforderungen dreier inhomogener Märkte entwickelt. Die Abbildung 5 zeigt im ersten Quadranten die Märkte und die zugehörigen prognostizierten Marktanteile bei SOP.

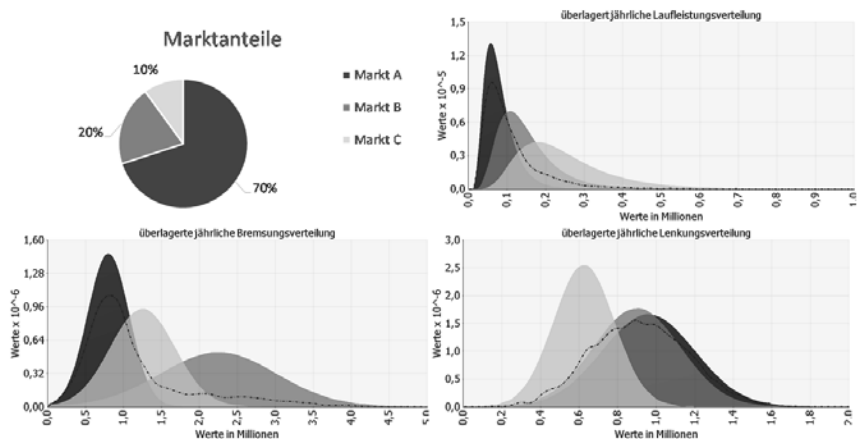


Bild 5: Marktanteile und die Gegenüberstellung der marktspezifischen und der gewichteten marktübergreifenden Nutzung

Ein Entwicklungsziel könnte es sein, den Markt B neu zu erschließen. Dieser neue Markt wurde hinsichtlich der Schädigungsparameter künstlich extrem verzerrt, um den Nutzen der Methode deutlich zeigen zu können. Marktforschungsstudien haben ergeben, dass ein Anteil von 20% der Systeme in diesem Markt abgesetzt werden können. Der Markt A soll zukünftig weiterhin eine dominierende Position einnehmen und ebenso soll der Markt C berücksichtigt werden. Ein wesentliches Entwicklungsziel ist es, eine Zuverlässigkeit von mindestens 90% für den Gesamtmarkt (bei einer Konfidenz von 90%) abzusichern. Diese Werte werden im VDA Band 3.2 als gängige Vorgaben genannt. Auf Basis von Felddatenanalysen und Expertenbefragungen wurden die Parameter "Kilometer-Fahrleistung", "Anzahl der Lenkungen" und "Anzahl der Bremsungen" bestimmt und als systemschädigend identifiziert. Die Gegenüberstellung der marktspezifischen und der mittels Hüllkurve simulierten marktübergreifenden Nutzung je Schädigungsparameter sind in den Quadranten 2, 3 und 4 der Abbildung 4 abgebildet. Es zeigt sich, dass die Anforderungen aus Markt A die marktübergreifende Nutzung wesentlich beeinflussen und Markt C durch seinen geringen Marktanteil den Gesamtmarkt-Nutzer nicht dominiert. Aus der simulierten Hüllkurve ergibt sich ein kritischer Gesamtmarkt-Nutzer, der durch folgende Parameter beschrieben wird:

- Üblicherweise anzunehmende Lebensdauer: 15 Jahre,
- Kilometer-Fahrleistung: 4.100.000 km,
- Anzahl der Lenkungen: 20.000.000,
- Anzahl der Bremsungen: 42.000.000.

Diese Parameter werden zur Zuverlässigkeitsabsicherung herangezogen. Vorab erfolgt jedoch der Abgleich der Markterfüllungen. Zur Bewertung der Markterfüllung werden die prozentualen Anteile zum Quantil des Gesamtmarkt-Nutzers aus der marktspezifischen Verteilung bestimmt und sind in der Tabelle 2 abgebildet.

Tabelle 2: Markterfüllung im Entwicklungsprojekt

	Markt A	Markt B	Markt C
Kilometer-Fahrleistung	99,78%	93,42%	66,53%
Anzahl der Lenkungen	94,02%	97,30%	99,99%
Anzahl der Bremsungen	99,99%	76,37%	99,99%

Das Ergebnis ist, dass für den Markt A bzgl. der Kilometer-Fahrleistung und der Anzahl der Bremsungen und für den Markt C bzgl. der Anzahl der Lenkungen und der Anzahl der Bremsungen eine 99,99% Markterfüllung vorliegt. Der Markt B weist für keinen Parameter eine 99,99% Markterfüllung auf. Angesichts der unzureichenden Markterfüllung könnte eine Maß-

nahme sein, den Markt B aus dem Entwicklungsprojekt auszugliedern. Da Markt B aber mit diesem Projekt neu erschlossen werden soll, benötigt es hier eine wirtschaftliche Risikoentscheidung: Ist das wirtschaftliche Risiko noch tragbar im Konsens zur wahrgenommenen Qualität und Zuverlässigkeit des Systems? Die Beantwortung dieser Fragestellung und die daraus folgenden Maßnahmen können vielfältig sein. So hätte der Hersteller beispielsweise die Möglichkeit, durch verkürzte Wartungsintervalle und eine verschärfte Feldbeobachtung das wirtschaftliche Risiko zu minimieren und mögliche Auffälligkeiten durch Früherkennung festzustellen. Wird diese Entscheidung zu Grunde gelegt, so sind folglich alle Märkte als auslegungsrelevant zu betrachten.

Die Ergebnisse der Erprobungsplanung sind in Abbildung 6 dargestellt.

Prüfungen		Prüfung A	Prüfung B	Felderprobung			
Anz. Prüflinge		2000	1500	600			
t_{Feld}		4.100.000 km	nachzuweisende Zuverlässigkeit		90%		
km Laufleistung	Raffungsfaktor	1,75	2	1	1,5		
	tLab real	80.000 km	100.000 km	120.000 km			
	MF	0,75	0,90	1,00			
	tLabÄqui	105.000 km	180.000 km	120.000 km			
	Lv	0,0256	0,0439	0,0293			
	$n_{\text{Äqui}}$	8,20	13,80	3,00	$\Sigma n_{\text{Äqui}}$	25,00	
	R(t)	75,51%	84,63%	46,47%	$R(t)_t$	91,20%	
	t_{Feld}	20.000.000	nachzuweisende Zuverlässigkeit		90%		
	Anzahl der Lenkungen	Raffungsfaktor		1,5	1	1,5	
		tLab real		800.000	1.500.000		
MF		0,80		1,00			
tLabÄqui		960.000		1.500.000			
Lv		0,0480		0,0750			
$n_{\text{Äqui}}$		15,77	12,32	$\Sigma n_{\text{Äqui}}$	28,10		
R(t)		86,42%	82,96%	$R(t)_t$	92,13%		
t_{Feld}		42.000.000	nachzuweisende Zuverlässigkeit		90%		
Anzahl der Brennsprünge		Raffungsfaktor		1	1,75		
		tLab real		2.250.000	3.500.000		
	MF	0,75		1,00			
	tLabÄqui	3.500.000		3.500.000			
	Lv	0,0703		0,0833			
	$n_{\text{Äqui}}$	19,20	7,76	$\Sigma n_{\text{Äqui}}$	26,96		
	R(t)	88,70%	74,31%	$R(t)_t$	91,81%		

Bild 6: Randbedingungen für die Erprobungsplanung

Zur Erprobungsplanung sind die Prüf- mit den Feldkollektive abgeglichen, das probabilistische Ausfallverhalten ermittelt und die Lebensdauerverhältnisse bestimmt worden. Es zeigt sich, dass nicht alle Prüfungen die verschiedenen Schädigungsparameter abdecken. Nur während der Felderprobung kommt es zur Kombination aller Schädigungsparameter. Die anderen Prüfungen werden dennoch zur Absicherung der Zuverlässigkeit herangezogen, um

frühestmöglich im Entwicklungsprozess eine Aussage bzgl. der Zuverlässigkeit zu generieren (z.B. durch den Aufbau eines Zuverlässigkeitswachstumsmodells). Es sollte eine Mindestanzahl an geforderten Prüfungen je Prüfstand festgelegt werden. Anschließend werden mit Hilfe der Anzahl der Prüflinge und (F5) die Äquivalenzbauteile je Schädigungsparameter und die abgesicherten Zuverlässigkeiten bestimmt. Für alle Schädigungsparameter wird bei der angegebenen Anzahl von Prüflingen die nachzuweisende Mindestzuverlässigkeit überschritten.

Da die nachzuweisende Mindestzuverlässigkeit überschritten wurde, besitzt der Versuchsplan das Potenzial, ressourceneffizient optimiert zu werden. In einem Entwicklungsprozess sind insbesondere die Entwicklungskosten und -zeit gern gewählte Optimierungsgrößen. Wenn die Erprobungsplanung unter diesen Voraussetzungen betrachtet wird, kann durch optimale Aufteilung der Prüflinge sowie der Prüfzeiten und -kosten ein Erprobungsplan erstellt werden, der allen Anforderungen des Entwicklungsprojektes Rechenschaft trägt.

5. Fazit

In der Zuverlässigkeitsabsicherung wird der kritische Kunde genutzt, um ein Entwicklungsziel zu definieren. Die Berücksichtigung von inhomogenen Anforderungen verschiedener Märkte stellt den bestehenden Ansatz vor Grenzen, da dieser die unterschiedlichen Führungsgrößen der Schädigungsparameter, die Marktanteile und die Verkettung der Schädigungsparameter nicht berücksichtigt. Die vorgestellte Methode schließt diese Aspekte mit in den bestehenden Ansatz zur Absicherung ein und vermeidet somit Überdimensionierung. Eine Entscheidungshilfe für wirtschaftliche Risiken kann in einer frühen Entwicklungsphase generiert werden, um eine gezielte Entscheidung bzgl. der Kundenzufriedenheit zu treffen. Die vorgestellte Erprobungsplanung bietet die Möglichkeit der Optimierung hinsichtlich Prüfdauer, -kosten und nachgewiesener Mindestzuverlässigkeit. Dies bietet Erweiterungspotential der Methode hinsichtlich zuverlässigkeitsoptimierter Prüfplanerstellung.

6. Literaturverzeichnis

- [1] Braasch, A.; Metzle, M.: Mit System in Feld und Zukunft – Einführung eines ganzheitlichen Zuverlässigkeitsmanagements, Sonderdruck aus der Fachzeitschrift QZ Qualität und Zuverlässigkeit, Carl Hanser Verlag, München, 2016, QZ-Archiv: 1699753.
- [2] Meyna, A.; Pauli, B.: Taschenbuch der Zuverlässigkeitstechnik – Quantitative Bewertungsverfahren, 2. Auflage, Carl Hanser Verlag, München, 2010, ISBN 978-3-446-41966-7.
- [3] VDA; QMC: VDA-Band 3.2 - Zuverlässigkeitssicherung bei Automobilherstellern und Lieferanten, 4.Aufgabe, Henrich Druck + Medien GmbH, Frankfurt am Main 2016

Auslegung von Zeitraffertests auf Basis numerischer Simulationen im Rahmen der Dentalimplantat-Erprobung



M.Sc. **Annika Müller**,
M.Sc. **Marcin Hinz**,
Univ.-Prof. Dr.-Ing. **Stefan Bracke**,
Bergische Universität Wuppertal

Kurzfassung

Während die Produktanforderungen bezüglich Qualität, Zuverlässigkeit und Lebensdauer kontinuierlich steigen, sinkt die Produktentwicklungszeit. Schnelle und effektive Testverfahren sind notwendig, um ein höheres, technologisches Produktniveau in kürzerer Zeit zu entwickeln:

Erste Informationen über die konstruktiven Produktschwachstellen und die Produktlebensdauer liefern in der frühen Entwicklungsphase numerische Ermüdungssimulationen. Zeitintensive Entwicklungsschleifen bestehend aus der prototypischen Produktion, der Produktprüfung und der konstruktiven Korrektur können mithilfe von numerischen Berechnungen im Vorfeld reduziert werden.

Im Rahmen dieses Papers wird ein Konzept vorgestellt, welches einen Beitrag zur effizienten Auslegung von Zeitraffererprobung von dentalen Implantaten auf Basis von Finite Elemente Simulationen und unter Anwendung von Zeitraffermethoden leistet.

1. Einleitung

Die Verweildauer eines Dentalimplantats im Kiefer eines Patienten wurde zu Beginn der dentalen Implantologie in den 1990er Jahren auf zehn Jahre geschätzt. Heute werden für die Lebensdauer der Implantate dreißig Jahre angenommen. Innerhalb der Nutzungsphase ist eine einwandfreie Produktfunktionalität gefordert: Das Versagen im Kiefer verursacht sowohl für den Patienten als auch für den Hersteller katastrophale Folgen. Daher wird die Ermüdungsfestigkeit eines dentalen Implantats vor der Vermarktung mithilfe einer genormten Ermüdungsprüfung [1] belegt. Dabei kann die Ermüdungsprüfung eines Prüfkörpers bis zu mehreren Wochen benötigen. Die Gesamttestzeit eines dentalen Implantats mit allen Derivaten dauert bis zu sechs Monaten.

Eine numerische Simulation, die nicht nur eine Ermüdungsberechnung, sondern auch zeitraffende Methoden (z.B. auf Basis physikalisch / chemischer Modelle wie dem Coffin-Manson

Ansatz) beinhaltet, erzielt mehr Informationen über Produktschwachstellen, die Produktlebensdauer und ermöglicht es, die Prüfmethode zu verbessern sowie die Prüfdauer zu verkürzen.

Durch die Anwendung dieses Konzepts wird die Anzahl an Prüflingen verringert, die Entwicklungskosten reduziert und die Entwicklungsphase verkürzt.

2. Grundlagen: Accelerated Testing

Zur Identifizierung von Bauteilschwachstellen, die in der Entwicklung und oder Produktion entstanden sind, dient das quantitative Testen. Dieses Testverfahren dient ausschließlich der Steigerung der Robustheit und Zuverlässigkeit des Bauteils. Dazu zählt auch das sogenannte Step-Stress-Testing (SST): ein Prüfling wird fortlaufend höheren Beanspruchungsniveaus unterzogen [6]. Zunächst wird das Bauteil eine definierte Zeit einer definierten konstanten Beanspruchung ausgesetzt. Versagt es nicht, wird es eine definierte Zeit einer höheren Beanspruchung ausgesetzt. Auf diese Art und Weise wird die Beanspruchung Stück für Stück bis zum Versagen erhöht. Normalerweise durchlaufen alle Prüflinge dasselbe Muster aus Beanspruchungsniveau und Zeit [6]. In Bild 1 sind zwei verschiedene Muster beispielhaft dargestellt.

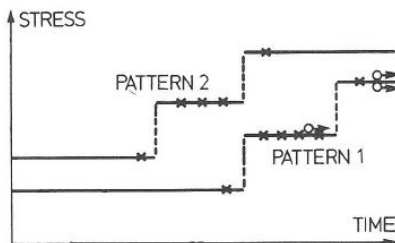


Bild 1: Step Stress Testing [6]

Während die Kreuze die Bauteile repräsentieren, die versagen, repräsentieren die Kreise die überlebenden Bauteile. Das schnelle Erreichen von Versagen ist der wesentliche Vorteil dieses Verfahrens. Allerdings garantieren sie keine genauen Abschätzungen der Produktlebensdauer. Darüber hinaus werden die meisten Bauteile unter konstanten und nicht unter schrittweisen Beanspruchungen betrieben. Hinzu kommt, dass es sich um ein komplexes Verfahren handelt, da hier ein kumulatives Schadensmodell betrachtet wird [6].

Anwendung findet das Verfahren in der Dentalimplantat Forschung: 2012 prüfte Freitas [7] zweiteilige Implantatsysteme aus Titan unter Wasser. In Bild 2 sind die drei Belastungsprofile, die er „mild“, „moderate“ und „aggressive“ nannte abgebildet.

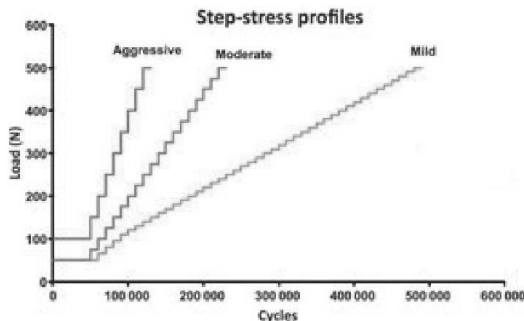


Bild 2: Lastprofile der Erprobung mithilfe des Step-Stress-Testing Verfahrens nach [7]

Das Verfahren verwendeten auch Silva 2009 [8], Baldassarri 2011 [9] und Borba [10]. Dabei variierten die Belastungsfrequenz, das Prüfmedium, die Umgebungstemperatur sowie der Implantatwerkstoff.

3. Grundlagen: Dentalimplantat-Erprobung

Dentale Implantate gehören zu den Medizinprodukten der Risikogruppe IIb, die ein erhöhtes Risiko für den Menschen darstellen. Sie werden gemäß der Norm DIN EN ISO 14801 [1] einer dynamischen Ermüdungsprüfung unter der Bedingung des „ungünstigsten Falles“ [1] (=Worst-Case-Szenario) unterzogen.

Das standardisierte Prüfverfahren findet im Labor (in-vitro) in warmer Luft ($20\pm 5^{\circ}\text{C}$) statt. Der Prüfaufbau für Implantatsysteme, die keine abgewinkelten Verbindungsteile enthalten, wird in Bild 3 schematisch dargestellt.

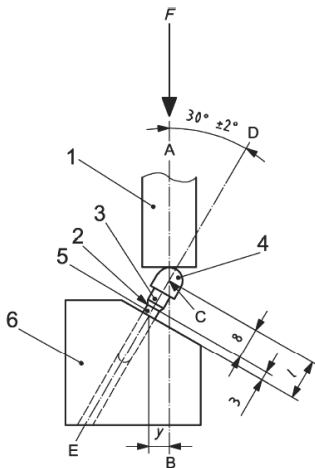


Bild 3: Schema des Prüfaufbaus für Systeme mit nicht abgewinkelten Verbindungsteilen [1]

Dabei muss der im Knochen zu verankernde Teil des dentalen Implantatkörpers (5) mit einer starren Einspannvorrichtung befestigt werden [1]. Wenn ein Einbettungsmaterial für die Probestkörperhalterung (6) verwendet wird, muss dieses einen Elastizitätsmodul von $>3\text{GPa}$ aufweisen [1]. Der Implantatprüfling muss so eingespannt werden, dass seine Achse einen Winkel von $30 \pm 2^\circ$ zur Belastungsrichtung des Prüfgerätes (1) bildet [1]. Die Kraftübertragung erfolgt querkraftfrei und uniaxial über den halbkugelförmigen Belastungsaufsatz (4) auf das genau definierte Belastungszentrum C. Um einen typischen Fall für den natürlichen Knochenabbau darzustellen, wurde zwischen dem nominellen Knochenrand (2) und der Probestkörperhalterung innerhalb der Norm ein Abstand von $3,0 \pm 0,5\text{mm}$ gewählt [1].

Während der Ermüdungsprüfung muss die Kraft F sinusförmig zwischen einem Nenn-Höchstwert und 10% dieses Wertes bei einer maximalen Frequenz von 15Hz wechseln [1]. Als geeignete Anfangsbelastung werden 80% der Belastung, die in einer statischen Prüfung mit derselben Prüfgeometrie zum Bruch führt, angegeben [1]. Die anschließenden Prüfungen werden mit niedrigeren Belastungen durchgeführt. Mindestens drei Prüflinge müssen diese Belastung $5 \cdot 10^6$ Zyklen versagensfrei überstehen. Wird das Belastungsniveau in Newton logarithmisch über die logarithmierte Anzahl an Lastzyklen aufgetragen, ergibt sich eine bauteilbezogene Wöhlerkurve, beispielhaft dargestellt in Bild 4.

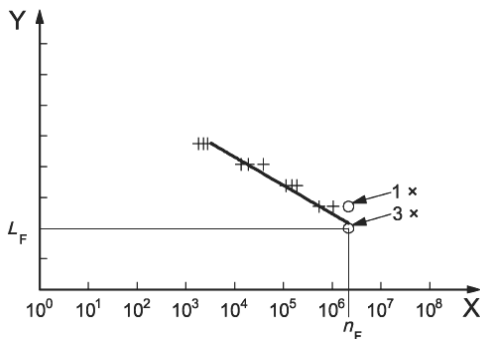


Bild 4: Beispiel einer Wöhlerkurve [1]

Die Plus-Zeichen stellen die gebrochenen und die Kreise die überlebenden Prüflinge dar. Die geforderten $5 \cdot 10^6$ Zyklen (n_F) überstehen drei Prüflinge bei der Ermüdungsgrenze L_F .

Seit Oktober 2014 liegt ein Entwurf [2] zum Ersatz dieser Norm vor. Neben einer detaillierten Beschreibung des „ungünstigsten Falles“ wird hier als Alternative zur klassischen Ermüdungsprüfung das Treppenstufenverfahren angegeben. Bei diesem Verfahren wird die Ermüdungsgrenzbelastung zunächst abgeschätzt oder über die statische Festigkeit berechnet, 10% ihres Wertes bilden die Stufentiefe d [2]. Die Prüfung beginnt mit der geschätzten Ermüdungsgrenzbelastung. Übersteht ein Prüfling die geforderte Anzahl an Lastwechseln von $5 \cdot 10^6$ Zyklen (von 2 bis maximal 15Hz) ist die Belastung um d zu erhöhen, bei Versagen um d zu verringern. Die Prüfung ist beendet wenn mindestens vier Prüflinge versagt und mindestens vier überlebt haben. Zuletzt wird eine statistische Auswertung vorgenommen, bei der die Ermüdungsgrenzbelastungen zur Berechnung der Überlebenswahrscheinlichkeiten von 10%, 50% und 90% dienen.

4. Ansatz

Der Ansatz zur Auslegung von Zeitraffertests auf Basis numerischer Simulationen besteht aus sechs aufeinanderfolgenden Schritten, schematisch dargestellt in Bild 5. Im Folgenden wird das Konzept anhand des Anwendungsgebietes der dentalen Implantologie erläutert. Im Fokus steht ein einphasiges Zahnimplantat aus Titan, welches nach einem Zahnverlust als künstliche Zahnwurzel dient.

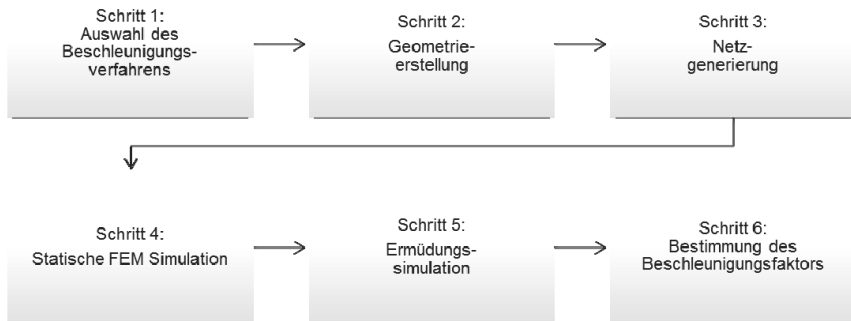


Bild 5: Ablauf des Ansatzes Schritt für Schritt

4.1 Schritt 1 - Auswahl des Beschleunigungsverfahrens

Zur Reduzierung der Erprobungsdauer werden Prüfungen unter verschärften Beanspruchungen durchgeführt [11]. Für die sogenannte Raffung wird im Rahmen dieser Studie die Temperatur als Beschleunigungsvariable ausgewählt. Zur Verifikation der numerischen Ergebnisse kann eine Erwärmung einfach realisiert werden. Darüber hinaus wird eine Raffung durch eine kontinuierliche zyklische Be- und Entlastung des Implantats, ohne Unterbrechungen, simuliert.

4.2 Schritt 2 - Geometrieerstellung

Zunächst wird die Geometrie eines handelsüblichen einphasigen Dentalimplantats in einer gewöhnlichen CAD-Software Umgebung erstellt. Zahnimplantate werden mit einer Länge von 7-13mm und einem Durchmesser von 3-4,8mm aus Titan hergestellt [17]. Neben einer hervorragenden Biokompatibilität und Korrosionsbeständigkeit weist Titan eine hohe spezifische Festigkeit bei geringer Dichte auf [3]. Außerdem ermöglicht die Anwendung von Titan problemlos medizinische Weiterbehandlungen, wie beispielsweise das Röntgen [3].

In Bild 6 ist das speziell für diese Studien konstruierte einphasige Dentalimplantat dargestellt. Die Gesamtlänge beträgt 13mm. Der Kopf des Implantats ist 2,3mm hoch und hat einen maximalen Außendurchmesser von 4,8mm. Auf diesen Kopf, der aus dem Zahnfleisch herausragt, wird der keramische Zahnersatz geschraubt. Der Gewindedurchmesser am Rumpf des Implantats misst 3,2mm. Dieser untere Teil wird in den Kieferknochen des Patienten geschraubt und wächst im Idealfall fest ein. Die hier angewendete Modellierung beinhaltet we-

der die Verbindungsschraube noch den keramischen Zahnersatz, da sie zu Kontaktproblemen und der Nutzung eines weiteren Spezialwerkstoffes führen, welche an dieser Stelle den Rahmen des Möglichen übersteigen [17].

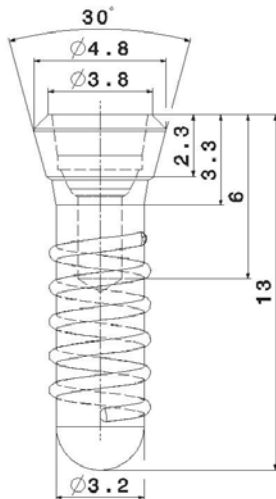


Bild 6: Skizze des einphasigen Dentalimplantats für diese Studie [17]

4.3 Schritt 3 – Netzgenerierung

Die erstellte CAD-Geometrie muss zunächst in einem kompatiblen Dateiformat abgespeichert werden, um in der Berechnungssoftware verwendet werden zu können [17]. Während der Umwandlung können Informationsverluste auftreten. Zur Vernetzung wird aber ein vollständiges und geschlossenes Volumenmodell benötigt, daher müssen fehlende Flächen erzeugt, überflüssige entfernt sowie unpassende Flächen entweder repariert oder entfernt und neu erstellt werden. Liegt im Anschluss ein geschlossenes Flächenmodell vor, kann daraus ein Volumenmodell erzeugt werden. Während der Vernetzung wird es in beschreibbare Teilgebiete, die Elemente, zerlegt. Im Rahmen dieser Studie werden quadratische Tetraederelemente, die auch parabolisch oder Elemente zweiter Ordnung genannt werden, verwendet [17]. Sie sind besonders für komplexe Geometrien mit vielen verschiedenen Radien und Winkeln geeignet. Zu Verifizierung der Netzgröße wird in Anlehnung an die Studie von Pessoa [4] eine Konvergenzstudie der maximal auftretenden Von-Mises Vergleichsspannung durchgeführt. Dazu wird das Implantat mit quadratischen Tetraederelementen der folgenden Elementgrößen vernetzt: 1,25mm, 1,0mm, 0,75mm, 0,5mm, 0,25mm, 0,125mm und 0,1mm.

Bei kleineren Elementgrößen steigt die Anzahl der Elemente: bei einer Elementgröße von 0,125mm beträgt die Elementanzahl 65482, bei 0,1mm 105651. Für die Konvergenzstudie wird die maximale Von-Mises Vergleichsspannung resultierend aus der Belastung mit drei verschiedenen Kräften (100N, 200N und 300N) herangezogen. Das Konvergenzkriterium wird in Anlehnung an die Studie von Pessoa [4] auf 3% Änderung der maximalen Von-Mises Vergleichsspannung von einer Vernetzungsstufe zur nächst feineren gewählt. In Bild 7 ist die resultierende Vergleichsspannung in Abhängigkeit der Belastung über die Elementanzahl aufgetragen.

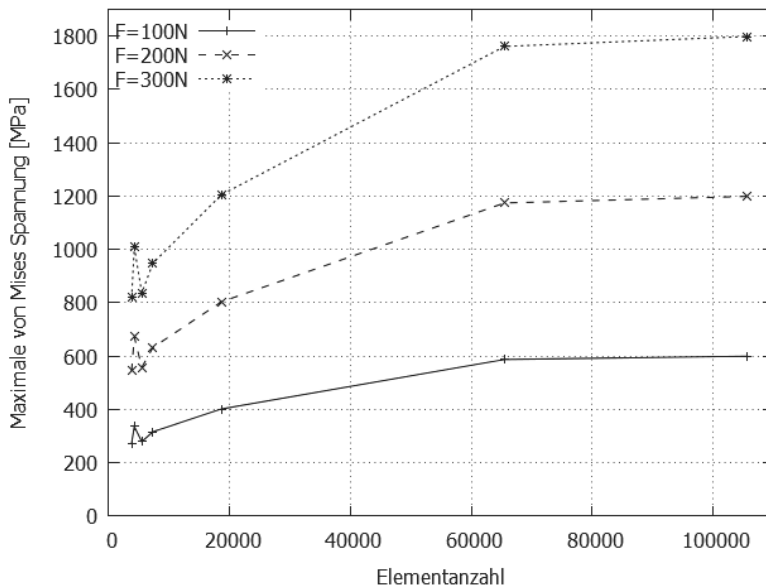


Bild 7: Konvergenzstudie

Die relative Abweichung der maximalen Von-Mises Vergleichsspannung von 65482 auf 105651 Elemente beträgt im Mittel 2,1% und erfüllt damit das Konvergenzkriterium.

4.4 Schritt 4 - Statische FEM Simulation

Im Rahmen der statischen Berechnung wird der Spannungszustand des Implantats in einer handelsüblichen Berechnungssoftware bestimmt [17]. Die einwirkenden Kräfte und Lagerungsbedingungen werden auf Basis der Erprobungsnorm [1] und weiteren Literaturquellen [4, 12, 13] angenommen.

Außerdem wird angenommen, dass es sich um ein rotationssymmetrisches Bauteil handelt: nur das Außengewinde des Implantats ist nicht vollständig symmetrisch. Darüber hinaus ist sowohl die Insertionstiefe als auch -richtung maßgeblich vom Knochenstatus sowie der individuellen Kieferbeschaffenheit des Patienten (wie beispielsweise die Anzahl und Lage der Nachbarzähne, Verlauf der Nervenbahnen etc.) abhängig.

Gemäß der Prüfnorm [1] wird das Implantatsystem im 30° Winkel belastet. Bezüglich der Kraftübertragung vom keramischen Zahnersatz bzw. des halbkugelförmigen Belastungsaufsatzes auf das dentale Implantat wird die Annahme getroffen, sie finde auf den Außenflächen des Implantatkopfes statt. Diese Flächen sind in Bild 8 rot eingefärbt.

Als Lagerung wird eine feste Einspannung gewählt, weil idealerweise davon auszugehen ist, dass das Implantat fest in den Kiefer des Patienten einwächst. Der nominelle Knochenschwund von 3mm wird bei der Implementierung berücksichtigt. Die Flächen, die fest eingespannt werden, sind in Bild 8 blau dargestellt.

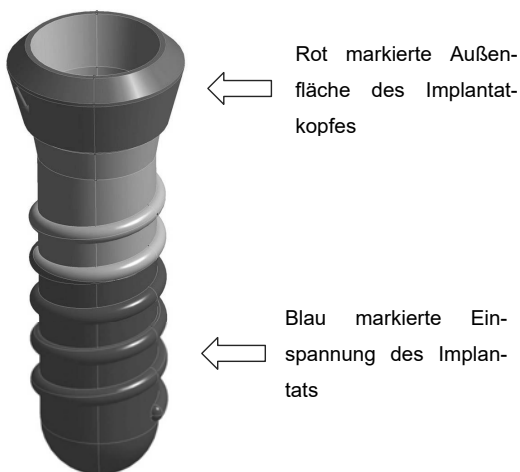


Bild 8: Anwendung der gewählten Randbedingungen auf das einphasige Dentalimplantat dieser Studie

Um einen möglichst präzisen Überblick über die Verteilung der Spannungen im Implantat zu erhalten, wird das Implantat, zunächst bei Raumtemperatur, mit Kräften von 100N bis 400N in 25N Schritten belastet. In Bild 9 ist die resultierende Von-Mises Vergleichsspannung exemplarisch dargestellt.



Bild 9: Von-Mises Vergleichsspannung von $F=250\text{N}$ bei Raumtemperatur

Eindeutig zu erkennen ist das Auftreten der maximalen Spannung in Höhe von $1498,6\text{MPa}$ am Rand der Einspannung. Hier tritt das höchste Moment auf.

Neben einer rein mechanischen Belastung des Bauteils, die bisher erörtert wurde, besteht auch die Möglichkeit Bauteile thermisch zu belasten. Temperaturänderungen können Spannungen und Formänderungen verursachen [5]. Wird die freie Wärmeausdehnung behindert, entstehen im inneren des Bauteils Spannungen [5]. Für die Wahl der Temperatur ist zu beachten, dass es sich bei Titan, ähnlich wie bei Eisen, um einen Werkstoff handelt, der eine Phasenumwandlung durchläuft: unterhalb von 882°C liegt Titan als sogenanntes α -Titan mit einer hexagonalen Gitterkonfiguration vor. Oberhalb der Umwandlungswandlungstemperatur entsteht das kubisch-raumzentrierte β -Titan [3]. Daher wird für die Belastung bestehend aus Kraft und Temperatur, eine Temperatur in Höhe von 220°C gewählt. Die Temperaturübertragung erfolgt instationär über die Außenflächen des Implantatkopfes. Die resultierende Von-Mises Spannung aus Temperaturbelastung und Kraft ist exemplarisch in Bild 10 dargestellt.

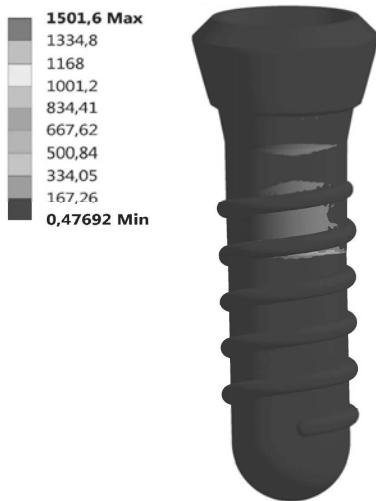


Bild 10: Von-Mises Vergleichsspannung bei einer Belastung von $F=250\text{N}$ und $T=220^\circ\text{C}$

Resultierend aus der Temperaturbelastung mit 220°C auf den Kopf des Implantats sowie einer Kraft in Höhe von 250N tritt ein höheres Beanspruchungsniveau auf: die maximale Spannung beträgt $1501,6\text{MPa}$ und tritt am Rand der Einspannung auf.

4.5 Schritt 5 - Ermüdungssimulation

Im fünften Schritt des Ansatzes wird die FEM basierte Ermüdungssimulation mithilfe einer handelsüblichen Software durchgeführt. Als Input dient die statische Berechnung des Spannungszustandes, der zur Bestimmung der Belastbarkeit des Implantats zyklisch aufgeprägt wird [17].

Als Berechnungsansatz wird der High Cycle Fatigue Ansatz ausgewählt [17]. Er gilt ab einer Belastung von 10^5 Zyklen [14] und ist damit für die Implantaterprobung, deren Grenzwertspielzahl nach DIN EN ISO 14801 mit $5 \cdot 10^6$ Zyklen definiert ist, heran zu ziehen [1]. Gemäß der Ermüdungsprüfung wird die Last sinusförmig wiederholt bis das Bauteil versagt. Dabei wird eine konstante, reinwechselnde Amplitude periodisch aufgeprägt. Da die Mittelspannung einen Einfluss auf die Belastbarkeit des Bauteils ausübt, wird der in der Technik meist [15] genutzte Ansatz zur Mittelspannungskorrektur verwendet: der Goodman Ansatz. Hierbei werden die Grenzwerte der Spannungsamplitude über einen linearen Ansatz durch die Wechselfestigkeit, die Zugfestigkeit und die Mittelspannung beschrieben [16].

Auf Basis der maximalen Vergleichsspannung wird die minimale Lebensdauer jedes einzelnen Elements berechnet. Ausgegeben wird die Anzahl an Lastzyklen, die das Implantat ohne Versagen überlebt. Die Bestimmung der Belastbarkeit basiert auf der Berechnung elastischer und empirischer Spannungen versus der Werkstoff Wöhlerkurve. Mithilfe dieses Ansatzes wird die Produktschwachstelle eindeutig identifiziert. In Bild 11 wird das Resultat der Belastung mit 250N und 220°C als Beispiel dargestellt. Ausgehend von der statischen Berechnung, versagt das Implantat, wie erwartet, am Rand der Einspannung.

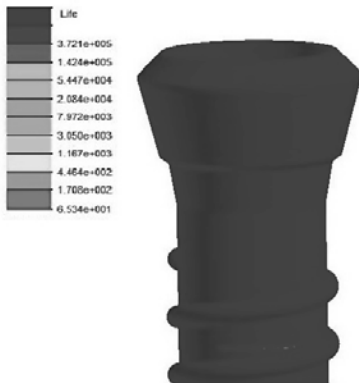


Bild 11: Resultierende Lebensdauer [Wiederholungen] der Belastung aus $F=250\text{N}$ und $T=220^\circ\text{C}$

4.6 Schritt 6 - Bestimmung des Beschleunigungsfaktors

Im Rahmen dieser Studie ist zur Raffung der Dental-Implantaterprobung die Temperatur als Beschleunigungsvariable untersucht worden. Zunächst wurde die Belastbarkeit, von 100N bis 400N in Raumtemperatur und im Anschluss zusätzlich zu einer Temperaturbelastung in Höhe von 220°C, numerisch berechnet. Das Ergebnis ist grafisch in Bild 12 dargestellt. Hier ist die Lebensdauer in ertragbaren Anzahl an Lastwiederholungen über die Kraft in Abhängigkeit der Temperatur aufgetragen.

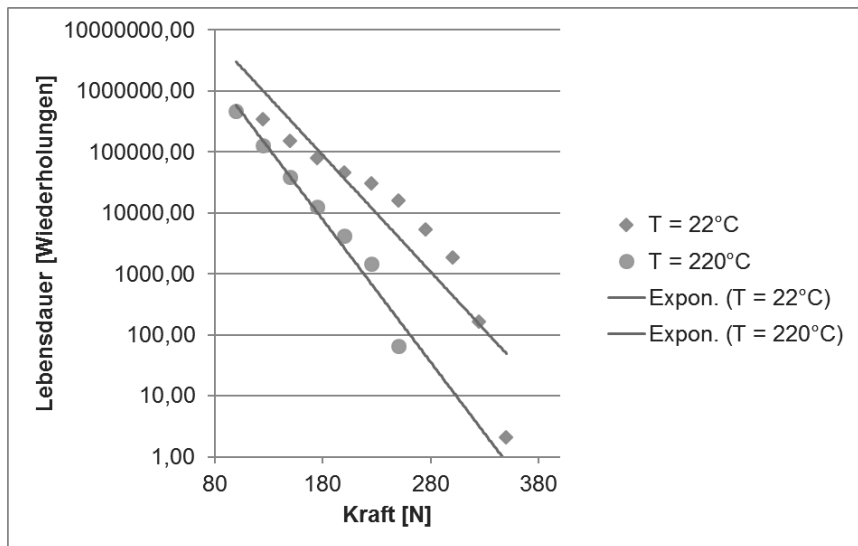


Bild 12: Ergebnis der numerischen Ermüdungsberechnung

Die blauen Rauten stellen die Lebensdauer in ertragbaren Wiederholungen bis zum Versagen des Implantats bei einer Umgebungstemperatur von 22°C dar. Die roten Punkte repräsentieren die Belastung aus Last und Temperatur in Höhe von 220°C. Die Geraden basieren auf einer Exponentialverteilung. Die Verschiebung von der blauen zur roten Geraden visualisiert den Beschleunigungsfaktor.

Das Ergebnis ist plausibel, obwohl beide Punktverläufe einen Knick aufweisen, der durch weitere Untersuchungen, wie eine feinere Aufteilung der Kraftintervalle und die Anwendung weiterer Temperaturen, genauer analysiert werden muss. Auch die beiden Ausreißer bei 250N und 350N müssen genauer betrachtet werden.

5. Zusammenfassung und Ausblick

Das vorliegende Paper stellt das Konzept sowie die Ergebnisse der numerischen Studie im Hinblick auf FEM Simulation von beschleunigter Zeitraffererprobung, die der Optimierung der Erprobungspläne dienen soll. Das Verfahren ist rein numerischer Natur, daher auch auf andere technische Produkte umsetzbar.

Die Simulation bietet eine hervorragende Alternative zur herkömmlichen Erprobung in den frühen Entwicklungsphasen, wobei anzumerken ist, dass die Erprobung von der Simulation

nicht abgelöst werden soll. Vielmehr stellt sie eine Erweiterung des Testings mit einer hohen Flexibilität bezüglich der Einsatzgebiete und Applikation von Randbedingungen dar.

Das vorgestellte Konzept ist einfach zu realisieren und wurde mit Hilfe von kommerzieller Software umgesetzt. Für die Umsetzung dieser Studie wurde ein realitätstreues Implantat in einem gängigen CAD Programm (Catia V5) erstellt und mit Third-Party Software vernetzt und berechnet. Die Lastapplikation erfolgte auf Basis der DIN EN ISO 14801 [1].

Diese Studie wird mit weiteren Berechnungen fortgeführt. Hierzu werden neben anderen Lasten (Lastkollektiven) weitere Zeitraffermodelle sowie Bauteilgeometrien simuliert. Im Hinblick auf die Diskreditierung der Bauteile sollen Netze mit anderen Netzelementen überprüft werden. Besonderer Focus wird dabei auf Elemente höherer Ordnung gelegt. Darüber hinaus sollen dynamische Belastungen sowie Bauteile bestehend aus mehreren Körpern (Multi-Body-Simulation) erforscht werden.

Ein wesentlicher Punkt der weiteren Forschungsarbeit stellt die Validierung der numerischen Ergebnisse dar. Hierzu ist eine Reihe an Experimenten unter Laborbedingungen geplant. Anzumerken ist, dass es zur Zeit der Forschungsarbeit keine standardmäßigen Prüfstände zur beschleunigten Erprobung von Zahnimplantaten auf dem Markt verfügbar sind.

Schließlich soll die Methode mittels Simulation und Validierung anhand gezielter Erprobung von anderen, technischen Produkten umgesetzt werden.

6. Literaturangaben

- [1] Deutsches Institut für Normung e.V.: Zahnheilkunde - Implantate
- Dynamische Ermüdungsprüfung für enossale dentale Implantate.
DIN EN ISO 14801:2008, Beuth Verlag, Berlin
- [2] Deutsches Institut für Normung e.V.: Zahnheilkunde - Implantate
- Dynamische Ermüdungsprüfung für enossale dentale Implantate.
DIN EN ISO 14801:2014, Beuth Verlag, Berlin
- [3] Peters, M. und C. Leyens: „Titan und Titanlegierungen“. Weinheim: Wiley-VCH 2002
- [4] Pessoa, R. S., et al. "Influence of Implant Connection Type on the Biomechanical Environment of Immediately Placed Implants—CT-Based Nonlinear, Three-Dimensional Finite Element Analysis." *Clinical implant dentistry and related research* 12.3 (2010) S. 219-234
- [5] Schlecht, B. „Maschinenelemente 1: Festigkeit, Wellen, Verbindungen, Federn, Kuppungen“. München: Pearson Education Deutschland GmbH 2007
- [6] Nelson, W. B. "Accelerated testing: statistical models, test plans, and data analysis". New York: John Wiley & Sons 1990
- [7] Freitas-Júnior, A. C. et al. "Reliability and failure modes of internal conical dental implant connections." *Clinical oral implants research* 24.2 (2013) S. 197-202
- [8] Silva, N., et al. "Reliability of one-piece ceramic implant." *Journal of Biomedical Materials Research Part B: Applied Biomaterials* 88.2 (2009) S. 419-426
- [9] Baldassarri, M. et al. "Reliability and failure modes of implant-supported zirconium-oxide fixed dental prostheses related to veneering techniques." *Journal of dentistry* 39.7 (2011) S. 489-498
- [10] Borba, M. et al. "Step-stress analysis for predicting dental ceramic reliability." *Dental Materials* 29.8 (2013) S. 913-918
- [11] Verband der Automobilindustrie. "VDA 3 Teil 2 Qualitätsmanagement in der Automobilindustrie—Zuverlässigkeitssicherung bei Automobilherstellern und Lieferanten—Zuverlässigkeits-Methoden und-Hilfsmittel." Frankfurt 2000
- [12] Sannino, G. et al. "Influence of finish line in the distribution of stress trough an all ceramic implant-supported crown: A 3D Finite Element Analysis." *ORAL & implantology* 2.2 2009 S. 14-27
- [13] Amid, R. et al. "Effect of microthread design of dental implants on stress and strain patterns: a three-dimensional finite element analysis." *Biomedizinische Technik/Biomedical Engineering* 58.5 (2013) S. 457-467

- [14] Gudehus, H. und H. Zenner. "Leitfaden für eine Betriebsfestigkeitsrechnung". Verlag Stahleisen GmbH: Düsseldorf 1999
- [15] Haibach, E. „Betriebsfestigkeit - Verfahren und Daten zur Bauteilberechnung." Springer Science & Business Media: Berlin Heidelberg 2006
- [16] Radaj, D. und M. Vormwald. „Ermüdungsfestigkeit: Grundlagen für Ingenieure". Springer-Verlag: Berlin Heidelberg New York 2007
- [17] A. Müller, M. Hinz and S. Bracke: "Optimization of the dental implant testing based on FEM simulation of fatigue and accelerated life." 26th European Safety and Reliability Conference (ESREL), Glasgow, 25. – 29.09.2016, Glasgow, United Kingdom. European Safety and Reliability Association, ESRA (2016)

Planung eines Zuverlässigkeits-DoE im Spannungsfeld zwischen Test-Design und Erfolgswahrscheinlichkeit



Dipl.-Ing. **Matthias Stohrer**, Dipl.-Ing. **Kevin Lucan**,
Prof. Dr.-Ing. **Bernd Bertsche**,
Institut für Maschinenelemente, Universität Stuttgart

Kurzfassung

In diesem Beitrag wird ein Vorgehen zur systematischen und fallbezogenen Planung eines Zuverlässigkeits-DoE für Produkte mit weibullverteilten Lebensdauern vorgestellt. Kern des Vorgehens ist das Ergebnis einer Parameterstudie, in der mit Hilfe von Monte-Carlo-Simulationen der Einfluss von Stichprobenumfang, Weibullformparameter, Zensierung, Faktorenanzahl und Testdesign auf die Genauigkeit geschätzter Modellparameter untersucht wird. Das vorgestellte Konzept der „Erfolgswahrscheinlichkeit“ ermöglicht die quantitative Bewertung der Genauigkeit der Schätzung von Modellparametern mit Hilfe eines einzelnen Kennwerts und ist damit Grundlage für eine Anwendung des Vorgehens in der Praxis. Mit Hilfe der Parameterstudie können bereits vor Durchführung erster Versuche Testdesigns ausgeschlossen werden, die ein hohes Risiko beinhalten, das gewünschte Ergebnis nicht zu erhalten. Mit einer Fallstudie wird die praktische Anwendung des Vorgehens erläutert.

1. Einleitung

Die statistische Versuchsplanung (engl. design of experiments, DoE) gilt als Methode zur effizienten Planung und Auswertung von Versuchsreihen mit normalverteilten Zielgrößen [1, 2]. Bei Anwendung auf Lebensdauerversuche technischer Produkte wird in der englischsprachigen Literatur von „Reliability-DoEs“, zu Deutsch Zuverlässigkeits-DoEs (Zuv-DoE) gesprochen [3]. Zuv-DoEs bieten die Möglichkeit, effizient den Einfluss von mehreren Faktoren wie Designparametern, Herstellungsverfahren, Belastungsarten und deren Wechselwirkungen auf die Produktzuverlässigkeit in Lebensdauerversuchen zu ermitteln. Dafür wird untersucht, ob eine Änderung des Niveaus der Eingangsfaktoren zu einer signifikanten Veränderung der Lebensdauer des Produkts führt. Daraus gewonnene Erkenntnisse finden Verwendung bei der Entwicklung zuverlässigerer Produkte im Sinne von Design for Reliability (DFR) und der Bestimmung von Raffungsmodellen für eine beschleunigte Produktabsicherung [4]. In der Literatur finden sich Versuchspläne für unterschiedliche Zielsetzungen [1, 2]. Im Fokus dieser Untersuchung sind 2-stufige teil- und vollfaktorielle Versuchspläne, die für Screenings und die Ermittlung linearer Zusammenhänge genutzt werden.

2 Eigenschaften von Zuverlässigkeits-DoE

Bei Zuv-DoEs muss berücksichtigt werden, dass Lebensdauerdaten als Zielgröße i.d.R. lognormal-, exponential- oder weibullverteilt, nicht aber normalverteilt sind. Zudem können Lebensdauerdaten aus Versuchen vollständig oder zensiert sein. In diesem Fall können Standardverfahren wie ANOVA oder lineare Regression, die auf normalverteilte Daten in herkömmlichen DoEs anwendbar sind, nicht mehr für die Ergebnisanalyse herangezogen werden [5]. In dieser Untersuchung wird daher das in [4] vorgestellte Verfahren der Maximum Likelihood Estimation zur Analyse von Zuv-DoEs angewendet.

3 Ansatz zur Planung eines Zuverlässigkeits-DoEs

Wichtiger Bestandteil der statistischen Versuchsplanung ist die Bestimmung des Stichprobenumfanges in Abhängigkeit von Trennschärfe, Alpha-Niveau, Streuung der Grundgesamtheit und gewünschter zu erkennender Effekthöhe. Dabei ist die Trennschärfe ein Maß für die Fähigkeit eines Tests, richtigerweise zu erkennen, dass die Nullhypothese falsch ist. Bei normalverteilter Zielgröße und bekannter Standardabweichung kann hierfür ein Gauß-Test als Hypothesentest zur Überprüfung des angesetzten Stichprobenumfangs angesetzt werden [6]. Da aber Lebensdauerdaten in der Regel nicht normalverteilt sind, lässt sich der Gaußtest auf sie nicht anwenden. Der Grundzusammenhang gilt allerdings auch für weibullverteilte Lebensdauerdaten. Ist die Effekthöhe E zwischen unterer und oberer Faktorstufe gering, liegen die Stichproben dicht beieinander und ihre Vertrauensbereiche überlappen sich deutlich, siehe Bild 1 links.

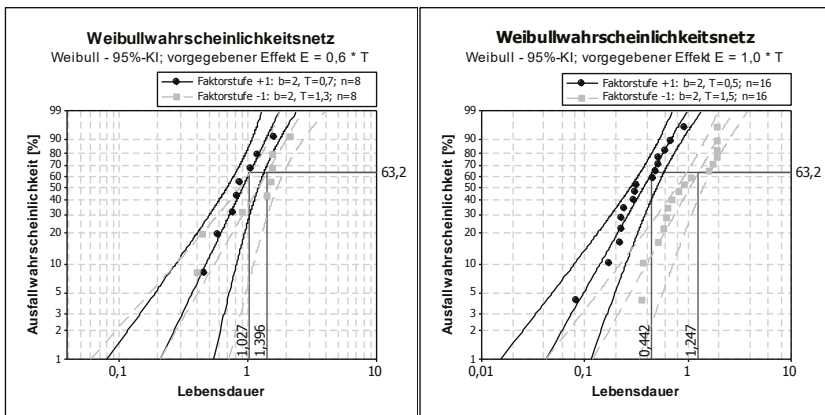


Bild 1: Einfluss von Effekthöhe und Stichprobenumfang auf die Trennschärfe

Die Trennschärfe steigt mit Effekthöhe E und Stichprobenumfang n , siehe Bild 1 rechts. Je größer der Stichprobenumfang, desto genauer können Effekte, d.h. die Differenz zwischen Lageparameter (hier: char. Lebensdauer T) auf oberer Faktorstufe FS_o ($x = 1$) und unterer Faktorstufe FS_u ($x = -1$), erkannt werden. Der notwendige Stichprobenumfang ist umso größer, je kleiner die Effekte sind, die erkannt werden sollen. Eine größere Streuung der Lebensdauern, charakterisiert durch einen kleinen Formparameter b , führt ebenfalls zu einer Erhöhung des notwendigen Stichprobenumfangs. Ziel dieser Arbeit ist, diesen Zusammenhang für weibullverteilte Daten zu quantifizieren und damit die Bestimmung des Stichprobenumfangs bei der Planung von Zuv-DoEs zu ermöglichen.

3.1 Simulation von Versuchsplänen

Für die rechnergestützte Untersuchung verschiedener Versuchspläne und Lebensdauermodelle (LM) wurde ein Simulationsprogramm in Matlab aufgebaut. Es ermöglicht die virtuelle Durchführung von Lebensdauerversuchen nach einem Versuchsplan für ein vorgegebenes LM, an dessen Ende die Parameter des geschätzten LM mit der Vorgabe verglichen werden. Für die Berücksichtigung der Zufallsstreuung bei der Bewertung wird der Versuchsplan 1.000-mal durchsimuliert. Die Implementierung von Latin Hypercube Sampling (LHS) zur Erzeugung zufälliger Ausfallzeiten für die Faktorstufenkombinationen (FSK) des Versuchsplans sorgt dabei für eine gute Abdeckung aller möglichen Ausfallzeiten. Eine automatisierte statistische Zuv-DoE Analyse ermöglicht eine zeiteffiziente Ermittlung der Modellparameter nach gleichbleibenden Kriterien.

3.1.1 Lebensdauermodell

Der Grundaufbau des Lebensdauermodells (LM) entspricht dem Ansatz von [4, 7] und wird nachfolgend vereinfacht beschrieben. Für die Modellierung von Effekten wird anstelle der char. Lebensdauer T in der Weibullverteilung ein lineares LM T_{LM} hinterlegt. Da T_{LM} eine Lebensdauer darstellt, die keine negativen Werte annehmen kann, wird eine logarithmische Transformation angewendet. Ein Beispiel für eine zweiparametrische Weibullverteilung mit einem Zwei-Faktoren-Lebensdauermodell ist in Gleichung (1) abgebildet. Dabei geben die Modellparameter a_1 bis a_3 in Kombination mit den Faktorstufeneinstellungen x_1 und x_2 die Effekthöhe E der Haupteffekte und Wechselwirkungen vor.

Für die Schätzung der Parameter des LM wird die Methode der Maximum-Likelihood-Schätzung verwendet. Dieses Schätzverfahren bietet den Vorteil, dass sowohl rechts- als

auch intervall-zensierte Daten berücksichtigt werden können. Durch die multiplikative Kombination der Likelihoodfunktionen können bei der Auswertung alle erdenklichen Ausgangsszenarien berücksichtigt werden, vgl. Gleichung (2).

$$f(t) = \frac{b}{e^{a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1 x_2}} \left(\frac{t}{e^{a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1 x_2}} \right)^{b-1} e^{-\left(\frac{t}{e^{a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1 x_2}} \right)^b} \quad (1)$$

$$L(b, a_0, a_1, \dots, a_k) = \prod_{i=1}^n \left[\frac{b}{T_{LM}} \left(\frac{t_i}{T_{LM}} \right)^{b-1} e^{-\left(\frac{t_i}{T_{LM}} \right)^b} \right] * \prod_{i=1}^n \left[e^{-\left(\frac{t_i}{T_{LM}} \right)^b} \right] \quad (2)$$

3.1.2 Automatisierte statistische Zuv-DoE Analyse

Für die Überprüfung der Signifikanz der Effekte wird ein Likelihood-Ratio-Quotienten-Test verwendet. Es wird definiert, dass alle Haupteffekte und Wechselwirkungen von Faktoren mit einem P-Wert von größer als 0,05 als nicht signifikant bewertet werden.

Die Herausforderung bei einer automatisierten, rein auf der Statistik beruhenden Auswertung liegt in der Definition eines Prozesses, den das Simulationsprogramm bei jedem Durchlauf selbstständig und auf dieselbe Weise durchläuft. Dieses ist nachfolgend beschrieben, Bild 2.

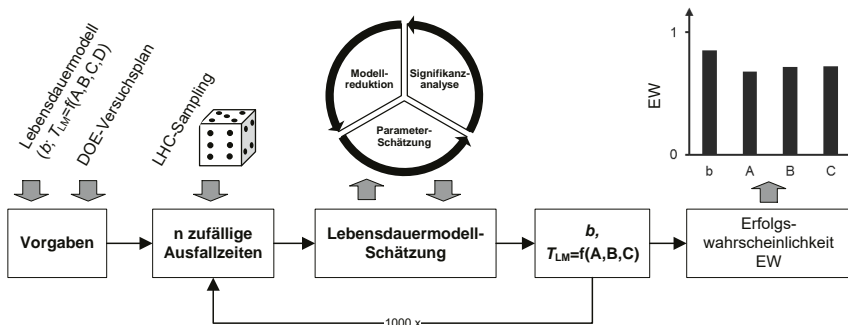


Bild 2: Ablaufschema der automatisierten statistischen Zuv-DoE Analyse

Zunächst wird die nicht signifikante Wechselwirkung mit der geringsten Signifikanz aus dem Modell gestrichen. Anschließend werden die verbleibenden Modellparameter erneut geschätzt und die nächste nicht signifikante Wechselwirkung gestrichen. Dies wird so lange wiederholt, bis nur noch signifikante Wechselwirkungen im Modell enthalten sind. Als

nächstes werden alle Faktoren, die nicht Teil einer Wechselwirkung sind, auf die Signifikanz ihres Haupteffekts untersucht und gegebenenfalls aus dem LM entfernt. Faktoren, die Teil einer Wechselwirkung sind, müssen, auch wenn ihr Haupteffekt nicht signifikant ist, im Modell enthalten bleiben [1]. Sollte nach dem Streichen eines Haupteffektes eine Wechselwirkung bei der nächsten Modellschätzung nicht signifikant werden, muss diese gestrichen werden, bevor weitere Effekte auf ihre Signifikanz untersucht werden können. Am Ende des iterativen Vorgehens steht ein LM mit geschätzten Parametern für Haupteffekte und Wechselwirkungen, die, basierend auf der Stichprobe, alle signifikant sind. Das geschilderte Vorgehen wird bei jedem Simulationsdurchlauf nach demselben Schema durchlaufen. Es ergeben sich somit 1.000 LM, die mit dem vorgegebenen LM zu vergleichen sind.

3.2 Parameterstudie als Planungsbasis

Basis des Planungsansatzes ist eine Parameterstudie, in der die Güte der Schätzung der Modellparameter eines vorgegebenen LM für verschiedene Kombinationen von Effekthöhen E , Formparameter b und Lebensdauerversuche je Faktorstufe eines Faktors n_{FS} ermittelt wird, vgl. Tabelle 1.

Tabelle 1: Definition des Untersuchungsraums der Parameterstudie

	Untersuchungsbereich							
b	1,5	2	2,5	3	3,5	-	-	-
E	$0,2 \cdot T$	$0,4 \cdot T$	$0,6 \cdot T$	$0,8 \cdot T$	$1,0 \cdot T$	-	-	-
n_{FS}	4	8	12	16	20	24	28	32

Zur Sicherstellung der Übertragbarkeit der Ergebnisse auf beliebige LM wird definiert:

- Die char. Lebensdauer T des LM im Zentralpunkt ($x = 0$) wird auf $T = 1$ normiert. Damit ergibt sich für alle Modelle nach Gleichung (2) die Vorgabe $a_0 = \ln(1) = 0$.
- Die Effekthöhe E wird auf die normierte charakteristische Lebensdauer T im Zentralpunkt bezogen. Ein Effekt von $0,4 \cdot T$ bedeutet damit, dass die Differenz zwischen char. Lebensdauer auf FS_o ($x = 1$) und FS_u ($x = -1$) $0,4 \cdot T$ beträgt.

Für die Parameterstudie wird ein 2-stufiger vollfaktorieller Versuchsplan für $k = 2$ Faktoren verwendet. Um die Genauigkeit der Schätzung der Haupteffekte der Faktoren A und B nicht durch WW zu beeinflussen (vgl. Abschnitt 3.1.2), wird die Zwei-Faktor-Wechselwirkung (2FWW) im LM gestrichen. Des Weiteren wird nur für Faktor A die Effekthöhe E entsprechend der Stufen in Tabelle 1 variiert. Faktor B wird hingegen null gesetzt. Dadurch kann ermittelt werden, bis zu welchem Mindeststichprobenumfang Pseudo-Effekte für Faktor B

erkannt werden, die durch die Zufallsstreuung der weibullverteilten Lebensdauern als vermeintlich signifikant bewertet werden. Das LM für die Parameterstudie lautet damit

$$\ln(T_{LM}) = a_1 x_1 + 0 \cdot x_2. \quad (3)$$

Der Zusammenhang zwischen a_1 in Gleichung (7) und vorgegebener Effekthöhe E ist

$$E = T_{LM}(x_1 = +1) - T_{LM}(x_1 = -1) = e^{a_1} - e^{-a_1} = -2\sinh(a_1). \quad (4)$$

Der Parameter a_1 errechnet sich damit zu

$$a_1 = \operatorname{arcsinh}\left(\frac{E}{-2}\right). \quad (5)$$

Für die Parameterkombination $b = 1,5$ und $E = 0,2 \cdot T$ ergibt sich damit

$$f(t) = \frac{1,5}{T_{LM}} \left(\frac{t}{T_{LM}}\right)^{1,5-1} e^{-\left(\frac{t}{T_{LM}}\right)^{1,5}}, \quad T_{LM} = e^{\operatorname{arcsinh}\left(\frac{0,2}{-2}\right) \cdot x_1 + 0 \cdot x_2}. \quad (6)$$

Der 2-stufige vollfaktorielle (VF) Versuchsplan für $k = 2$ Faktoren und Reduktionsstufe $p = 0$ enthält $m = 2^k(k - p) = 4$ Faktorstufenkombinationen. Je Faktor liegen davon für beide Faktoren $2^k(k - p - 1) = 2$ auf der unteren und $2^k(k - p - 1) = 2$ auf der oberen Faktorstufe. Zur Erreichung der Vorgabe von n_{FS} Versuchen je Faktorstufe müssen daher mehrere Replikationen n_R des Versuchsplans durchgeführt werden. Es gilt

$$n_R \geq \frac{2 \cdot n_{FS}}{m}. \quad (7)$$

Der Gesamtstichprobenumfang n errechnet sich zu

$$n = n_R \cdot m. \quad (8)$$

3.3 Erfolgswahrscheinlichkeit

Den Autoren ist bei der Ermittlung geeigneter Versuchspläne zur Bestimmung von signifikanten Haupteffekten oder Wechselwirkungen wichtig, dass neben den tatsächlichen durch

das LM vorgegebenen Effekten auch deren Vorzeichen und absoluter Wert mit hinreichender Genauigkeit bestimmt wird. Für die Bewertung der Genauigkeit der Schätzung einzelner Modellparameter mit einem bestimmten Versuchsplan wird daher die in [8] vorgestellte Erfolgswahrscheinlichkeit EW herangezogen. Sie gibt an, wieviel Prozent aller 1.000 Schätzungen eines Modellparameters a_i oder b innerhalb einer Toleranzbandbreite T_B um den Parameter des vorgegebenen LM liegen. Ein hoher Wert für EW ist danach ein Indiz für einen Versuchsplan, der robust gegenüber der Zufallsstreuung ist.

Für die Bewertung der Schätzung des Weibull-Formparameters b wird im Folgenden $T_B = \pm 0,1 \cdot b$, Gleichung (9). Für die Bewertung der geschätzten LM-parameter a_i wird eine Toleranzbandbreite von $\pm 10\%$ der normierten char. Lebensdauer T vorgegeben, Gleichung (10). Dadurch wird sichergestellt, dass in einem Ranking signifikanter Effekte eines LM keine Faktoren mit einem Effektunterschied $\geq 0,2 \cdot T$ fälschlicherweise in der falschen Reihenfolge angeordnet werden.

$$EW(b|T_B = \pm 0,1 \cdot b) \quad (9)$$

$$EW(a_i|T_B = \pm 0,1) \quad (10)$$

3.4 Ergebnisse der Parameterstudie

Alle möglichen Parameterkombinationen von b und E aus Tabelle 1 wurden mit dem in Kap. 3.2 vorgestellten VF Versuchsplan für alle in Tabelle 1 angegebenen n_{FS} nach dem in Kap. 3.1 vorgestellten Simulationsprogramm 1.000 mal simuliert.

Die Ergebnisse für $EW(a_i|T_B = \pm 0,1)$ sind in Tabelle 2 dargestellt. Folgende Zusammenhänge bestätigen sich:

- Je kleiner der Formparameter b , desto geringer ist die Erfolgswahrscheinlichkeit EW bei konstanter Lebensdauerversuchsanzahl je Faktorstufe und Effekthöhe E .
- Je kleiner der Effekt, der erkannt werden soll, desto geringer ist die Erfolgswahrscheinlichkeit bei gleichbleibendem Formparameter und konstanter Lebensdauerversuchsanzahl je Faktorstufe.

Tabelle 2: Erfolgswahrscheinlichkeit für a_i in Abhängigkeit von n_{FS} und Effekthöhe E

E [%·T]	n_{FS}	$EW(a_i T_B = \pm 0,1) [\%]$				
		b=1,5	b=2,0	b=2,5	b=3,0	b=3,5
20	4	< 5	< 5	< 5	8	11
20	8	< 5	< 5	< 5	9	14
20	12	< 5	< 5	6	15	28
20	16	< 5	< 5	12	23	41
20	20	< 5	< 5	17	36	52
20	24	< 5	7	23	40	60
20	28	< 5	12	29	51	64
20	32	< 5	17	36	57	73
40	4	< 5	9	17	26	31
40	8	< 5	15	31	41	54
40	12	6	25	43	57	60
40	16	14	36	50	60	67
40	20	20	46	58	71	75
40	24	32	50	63	69	81
40	28	37	55	64	76	83
40	32	40	57	68	79	85
60	4	8	17	26	31	38
60	8	16	31	40	48	56
60	12	25	40	51	59	66
60	16	34	45	54	63	72
60	20	39	51	62	73	81
60	24	41	55	67	72	80
60	28	45	59	69	78	86
60	32	49	60	72	81	87
80	4	14	23	28	37	41
80	8	27	35	44	49	59
80	12	34	41	51	60	66
80	16	39	47	57	65	74
80	20	43	57	68	73	81
80	24	44	57	68	76	84
80	28	46	60	74	82	89
80	32	50	65	76	84	89
100	4	18	29	33	40	48
100	8	30	40	49	56	62
100	12	35	49	56	69	71
100	16	41	54	63	72	80
100	20	47	63	73	79	85
100	24	50	61	76	82	88
100	28	53	68	78	87	92
100	32	55	70	80	88	93

Die Erfolgswahrscheinlichkeit für die richtige Schätzung des Formparameters innerhalb der vorgegebenen Toleranzbandbreite $EW(b|T_B = \pm 0,1 \cdot b)$ ist indes nur von der Anzahl an Versuchen je Faktorstufe n_{FS} und damit vom Gesamtstichprobenumfang abhängig, Bild 3.

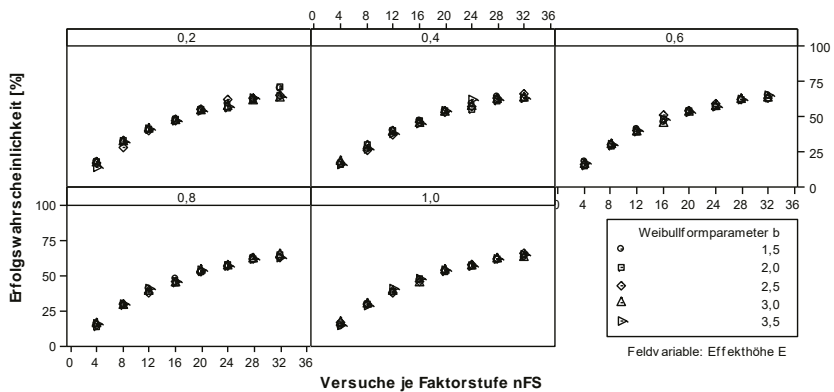


Bild 3: Erfolgswahrscheinlichkeit für Formparameter b in Abhängigkeit von Formparameter, Effekthöhe und Versuchsanzahl je Faktorstufe

3.5 Systematischen Planung eines 2F-Zuv-DoE

Auf Grundlage des Ergebnisses der Parameterstudie kann die systematische Planung eines beliebigen 2F-Zuv-DoEs in drei Schritten erfolgen, Bild 4.

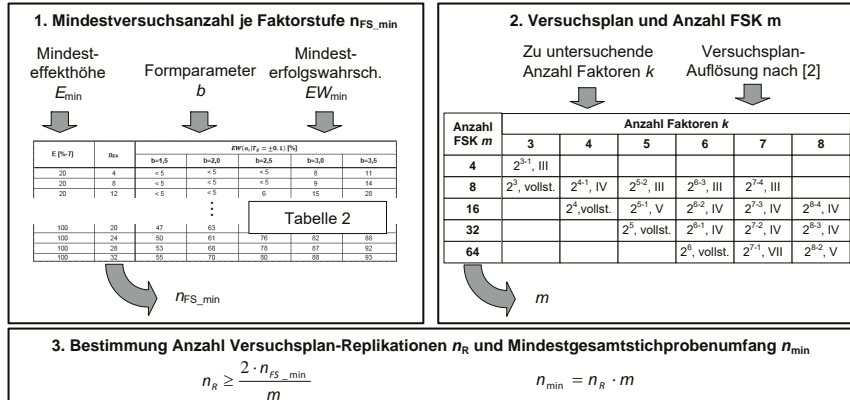


Bild 4: Planung eines 2F-Zuv-DoE in drei Schritten

Im ersten Schritt ist festzulegen, welche auf die normierte char. Lebensdauer bezogene Mindesteffekthöhe mit Hilfe des Versuchs erkannt werden soll. Zusätzlich wird ein Erfah-

rungs- oder Schätzwert für den Formparameter b benötigt. In Abhängigkeit der gewünschten Erfolgswahrscheinlichkeit kann damit aus Tabelle 2 die Mindestversuchsanzahl je Faktorstufe n_{FS_min} bestimmt werden.

Der zweite Schritt besteht aus der Auswahl eines geeigneten 2F-Versuchsplans in Abhängigkeit der zu untersuchenden Anzahl an Faktoren und der Versuchsplanauflösung. Sollen 2FWW sicher erkannt werden, empfiehlt sich eine Auflösung größer III [2]. Faktoranzahl und Auflösung bestimmen die Faktorstufenanzahl m .

Im dritten Schritt kann mit n_{FS_min} und m die Anzahl an Versuchsplan-Replikationen n_R sowie der Mindestgesamtstichprobenumfang n_{min} errechnet werden.

4. Fallstudie

Für vier Faktoren A, B, C und D soll mit einem Zuv-DoE ein Faktorscreening durchgeführt werden. Ziel ist, Effekte $E \geq 0,6$ -T mit $EW_{min} = 70\%$ zu erkennen. Bei Annahme von $b = 2,5$ ergibt sich $n_{FS_min} = 32$ (Tabelle 2). Da mögliche 2FWW erkannt werden sollen, wird ein VF Versuchsplan mit $m = 2^4 = 16$ FSK gewählt. Daraus ergibt sich $n_R = 4$ und $n_{min} = 64$.

Für die Überprüfung der für diesen Versuchsplan in Tabelle 2 prognostizierten Erfolgswahrscheinlichkeit für vorgegebene Effekthöhen wird ein fiktives LM mit den in Bild 5 links angegebenen Effekthöhen vorgegeben. Mit diesem Modell wird der gewählte Versuchsplan nach dem in Bild 2 gezeigten Vorgehen 1.000 Mal simuliert. Die sich ergebenden Istwerte für die EW der einzelnen Faktoren liegen dabei sehr nah an den EW -Planungswerten, siehe Bild 5 rechts. Abweichungen resultieren aus vermeintlich als signifikant beurteilten WW, die zu zusätzlichen Modelltermen führen und die restlichen Modellterme beeinflussen.

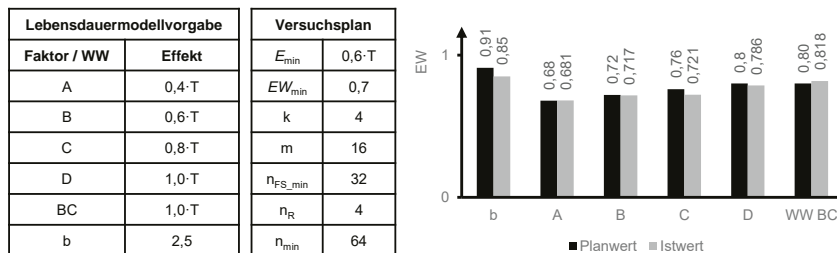


Bild 5: Lebensdauermodellvorgabe und Ergebnis für VF Versuchsplan der Fallstudie

5. Optimierungspotentiale

64 bis zum Ausfall gefahrene Lebensdauerversuche erfordern immense Testkapazitäten.

Es stellt sich daher die Frage nach Optimierungspotentialen für den Versuchsplan.

Zur Verkürzung von Lebensdauerversuchszeiten ist die Zeitzensierung ein möglicher Ansatz. Angewendet auf die Zuv-DoEs führt sie allerdings zu unausgewogenen Versuchsplänen, bei denen für manche FSK auf geringem Belastungsniveau kein einziger Ausfall zu verzeichnen ist. Entsprechend schlecht sind die Schätzer für die Modellparameter, Bild 6.

Lebensdauermodellvorgabe		Versuchsplan	
Faktor / WW	Effekt	E_{\min}	0,6-T
A	0,4-T	EW_{\min}	0,7
B	0,6-T	k	4
C	0,8-T	m	16
D	1,0-T	$n_{FS, \min}$	32
b	2,5	n_R	4
$t_{\text{Zensierung}}$	1-T	n_{\min}	64

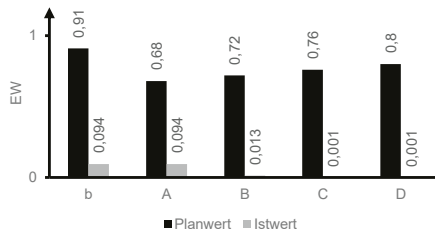


Bild 6: Lebensdauermodellvorgabe und Ergebnis für zeitzensierten VF Versuchsplan

Ein anderer Optimierungsansatz sind teilfaktorielle (TF) Versuchspläne. Da die EW bei konstantem b und E_{\min} nur von n_{FS} abhängt, ergibt sich bei gleicher Faktoranzahl k keine Stichprobenreduktion gegenüber VF Plänen. Jedoch können mit TF Plänen bei gleichem n mehr Faktoren gleichzeitig untersucht werden, siehe Bild 7. Ein weiterer praktischer Vorteil von TF Plänen liegt in der geringeren FSK- und damit Variantenanzahl, die zu geringeren Kosten für Grenzmuster und reduzierten Rüstzeiten im Versuch führen.

Lebensdauermodellvorgabe		Versuchsplan	
Faktor / WW	Effekt	E_{\min}	0,6-T
A	0,4-T	EW_{\min}	0,7
B	0,6-T	k	6
C	0,7-T	p	2
D	0,8-T	m	16
E	0,9-T	$n_{FS, \min}$	32
F	1,0-T	n_R	4
b	2,5	n_{\min}	64

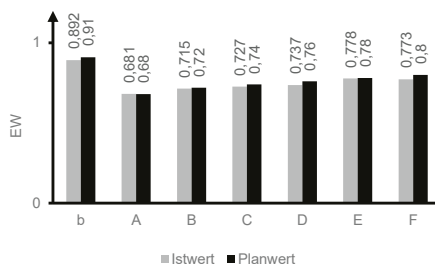


Bild 7: Lebensdauermodellvorgabe und Ergebnis für zeitzensierten TF Versuchsplan

6. Zusammenfassung

Die vorliegende Arbeit beschreibt ein Vorgehen zur systematischen Planung von 2-faktoriellen Zuverlässigkeits-DoEs für weibullverteilte Zielgrößen. Grundlage des Vorgehens und Voraussetzung für eine quantitative Bewertung des zu erwartenden Versuchsergebnisses sind eine standardisierte Parameterstudie sowie die Einführung der Erfolgswahrscheinlichkeit.

Ergebnis der Untersuchung ist, dass die Schätzung des Formparameters b ausschließlich vom Gesamtstichprobenumfang abhängt, während die Schätzung einzelner Lebensdauermodellparameter von Formparameter, Effekthöhe und Anzahl an Lebensdauerversuchen je Faktorstufe beeinflusst wird.

Für genaue Parameterschätzer sind große Stichprobenumfänge und damit lange Gesamtversuchsdauern notwendig. Eine Zeitzensierung führt allerdings zu sehr schlechten Parameterschätzungen und birgt damit kein Optimierungspotential. Teilkonfaktorielle Pläne bieten insbesondere bei großen Anzahlen zu untersuchender Faktoren ($k > 5$) die Möglichkeit, den Gesamtstichprobenumfang zu reduzieren. Hierbei ist zu beachten, dass 2FWW ab Versuchsplan-Auflösung III nicht mehr ohne Alias-Effekte erkannt werden können.

7. Danksagung

Die Autoren danken Jörg Waller für seine Unterstützung bei der Optimierung des Simulationsprogramms und der Auswertung der großen Datenmengen.

Literaturangaben

- [1] Siebertz, K., Bebbler, D. T. v. u. Hochkirchen, T.: Statistische Versuchsplanung. Design of Experiments (DOE). Heidelberg, Dordrecht [u.a.]: Springer 2010
- [2] Kleppmann, W.: Taschenbuch Versuchsplanung. Produkte und Prozesse optimieren. München [u.a.]: Hanser 2006
- [3] Wu, C.-F. J. u. Hamada, M.: Experiments. Planning, analysis, and parameter design optimization. Wiley series in probability and statistics. New York: Wiley 2000
- [4] ReliaSoft Corporation: Reliability DOE for Life Tests. ReliaSoft's Experiment Design and Analysis Reference. http://reliawiki.org/index.php/Reliability_DOE_for_Life_Tests, abgerufen am: 14.02.2017
- [5] ReliaSoft Corporation: An Example of Using Reliability DOE for Life Testing. Reliability Hotwire Nr. 88, 2008. <http://www.weibull.com/hotwire/issue88/hottopics88.htm>, abgerufen am: 14.02.2017
- [6] Steland, A.: Basiswissen Statistik. Kompaktkurs für Anwender aus Wirtschaft, Informatik und Technik. SpringerLink : Bücher. Berlin, Heidelberg: Springer Berlin Heidelberg 2007
- [7] Guo, H. u. Mettas, A.: Design of experiments and data analysis. 2010 Proc. Ann. Reliability and Maintainability Symp. (RAMS). IEEE 2010
- [8] Dazer, M., Stohrer, M., Kemmler, S. u. Bertsche, B.: Planning of reliability life tests within the accuracy, time and cost triangle. 2016 IEEE Accelerated Stress Testing & Reliability Conference (ASTR), S. 1–9